NEW SOLITARY WAVES FOR THIN-FILM FERROELECTRIC MATERIAL EQUATION ARISING IN DIELECTRIC MATERIALS

Jalil MANAFIAN*/** , **Walla Rahim JUADIH***** , **Amitha Manmohan RAO****** , **Baharak ESLAMI*******
Natavan ALLAHVERDIYEVA****** , **Parvin MUSTAFAYEVA*********

*Department of Applied Mathematics, Faculty of Mathematical Sciences, University of Tabriz, Tabriz, Iran
**Natural Sciences Faculty, Lankaran State University, 50, H. Aslanov str., Lankaran, Azerbaijan
***Computer Science Department, College of Education for Pure Sciences, University of Thi-Qar, Nasiriya, Iraq.
****Department of Mathematics \& Statistics, N.S.S College of Commerce & Economics,
University of Mumbai, Mumbai, Maharashtra, India
*****Department of Physics, Payame Noor University, P.O. Box 19395-4697, Tehran, Iran
******Sumgait State University, Sumgait, Azerbaijan
*******Ganja State University, Haydar Aliyev Ave., 429, Ganja, Azerbaijan

j_manafianheris@tabrizu.ac.ir, wallarahim85@gmail.com, amitha.m.rao@gmail.com, Bkeslami@pnu.ac.ir
natavan.sdu@gmail.com, pervin.mustafayeva.81@mail.ru

Abstract: In this paper, the thin-film ferroelectric material equation (TFFME), which enables the propagation of solitary polarisation in thin-film ferroelectric materials is investigated, will be expressed through the non-linear evolution models. Ferroelectrics are dielectric materials that explain wave propagation non-linear demeanors. The non-linear wave propagation form is administrated by TFFME. To investigate the characterisations of new waves and solitonic properties of the TFFME, the modified exponential Jacobi technique and rational $\exp(-\phi(\eta))$-expansion technique are used. Plenty of alternative responses may be achieved by employing individual formulas; each of these solutions is offered by some plain graphs. The validity of such schemes and solutions may be exhibited by assessing how well the relevant schemes and solutions match up. The effect of the free variables on the manner of acting of reached plots to a few solutions in the exact forms was also explored depending upon the nature of non-linearities. The descriptive characteristics of the reached results are presented and analysed by some density, two- and three-dimensional figures. We believe that our results would pave the way for future research generating optical memories based on non-linear solitons.

Key words: thin-film ferroelectric material equation, ferroelectrics are dielectric materials, modified exponential Jacobi method,
rational $\exp(-\phi(\eta))$-expansion method, soliton solution

1. INTRODUCTION

Ferroelectric thin films have become a potential nominee in the field of ultraviolet photodetectors detection due to their wide bandgap and unique photovoltaic aspects. Additionally, ferroelectric thin films perform excellent dielectric, piezoelectric, pyroelectric, acousto-optic effects, etc. [1]. Xiao et al. [2] showed that the growth of ferroelectric layer on the original perovskite grains can reduce the formation of grain boundaries and hence minimise the recombination of electrons and holes at grain boundaries. The solitary wave dynamics of the thin-film ferroelectric material equation (TFFME) were investigated by Xiao et al. [2]. Umoh et al. [3] used nanoelectronic devices based on oxide films to require materials for exhibiting combined properties such as ferroelectricity, ferromagnetism and ferroelasticity at the same phase. An analysis of the thermal and ferroelectric properties of using a thin film in a transverse field extended for a higher spin within the quantum Monte Carlo method was provided by Tarnaoui et al. [4]. Ferroelectric thin films have demonstrated great potential in electrocaloric solid-state refrigeration on account of large adiabatic temperature changes [3]. Yang et al. [5] demonstrated a

ferroelectric tunnel junction, whose conductivity varies linearly and symmetrically by judiciously combining ferroelectric domain switching and oxygen vacancy migration. The order parameter of ferrotoroidic order has been generated by a head-to-tail configuration of magnetic moment and has been theoretically proposed that one-dimensional dimerised and antiferromagnetic (AFM)-like spin chain hosts ferrotoroidicity [6]. Through the Landau–Ginzburg–Devonshire equation, the governing behaviour of the polarisation field in ferroelectric material was derived, and ferroelectric material is subjected to a standing electric field [7]. The modified simple equation method and the Riccati-Bernoulli Sub-ODE method were utilised for TFFME, which plays a vital role in optics to waves propagate through ferroelectric materials [8]. The optical soliton solutions of the thin-film ferroelectric materials equation through the Paul–Painlevé approach have been obtained [9]. In this paper, we are concerned with a wave polarisation for ferroelectric materials to the TFFME in one-dimensional form as follows:

$$\frac{m}{T^2}\frac{\partial^2 u}{\partial t^2} - [(p_2 - 2\mu)u + p_4 u^3 + p_6 u^5] - J\Delta u = 0, \qquad (1)$$

where mass and charge density are m and T, p_2, p_4, p_6 showing temperature and pressure. Also, J is related to the space inhomo-

geneity coefficient and μ is the reciprocal of the electric susceptibility [7]. In the last decades, researchers have developed numerous methods. Gu et al. [10] investigated the generalised (2 + 1)-dimensional shallow water wave equation, which enables a unidirectional propagation of shallow water waves by using the bilinear method and semi-inverse variational principle scheme. Some novel analytical solutions, including 2-lump-type, the interaction between 2-lump and one kink, and other forms were obtained for the (3 + 1)-dimensional Burger system [11]. The modified auxiliary equation approach and the generalised projective Riccati equation method were used for the first time to solve the Zoomeron equation [12]. In paper [13], experts applied the bilinear method on models arising in variable coefficient Caudrey–Dodd–Gibbon–Kotera–Sawada equation. In paper [14], the third-order evolution equation was investigated by the Hirota bilinear method, which arises the propagation of long waves over shallow water. Rao et al. [15] studied the extended homoclinic breather wave solutions to the non-linear vibration and dispersive wave systems. The solutions of (3 + 1)-dimensional Burgers system via Lie symmetry analysis have been investigated by Alimirzaluo et al. [16]. In paper [17], Xiao at al. studied the inverse scattering transform for the coupled modified Korteweg–de Vries equation with non-zero boundary conditions. The cubic B-splines and linear triangular elements were used for a test problem including the motion of a single solitary solution of Benjamin–Bona–Mahony (BBM) equation coupled BBM-system by finite element method [18]. M-soliton and N-soliton solutions have been discovered for variable coefficient-generalised non-linear wave equation arising in liquid with gas bubbles [19]. Hirota's bilinear operator has been used for the generalised Hietarinta equation [20] and to study the variable coefficients of generalised shallow water wave equation [21]. The exact solutions to the generalised non-linear Schrodinger equation by means of the extended sinh–Gordon equation expansion method, $\tan(\Gamma(\varpi))$-expansion method and the improved $\cos(\Gamma(\varpi))$ function method were obtained [22]. In paper [23], Singh and co-workers focussed on describing the evolution of water waves with higher order temporal dispersion by characterising the dynamics of lump and soliton waves on different spatially varying backgrounds to an integrable (3 + 1)-dimensional non-linear model. Non-linear partial differential equations arise in many different branches of social and basic sciences and engineering. They have gained prominence in recent years due to their crucial role in a variety of domains involving complicated physical processes, from control theory and electrical circuits to wave propagation. Especially, they appear in electrodynamics of complex medium, electrical networks, signal and image processing, electrodynamics, including porous flow, surface water flow, land sliding, faulting, circled fuel reactor, wave motion and distribution, transmission lines. In order to compute these solutions and better understand the fundamental characteristics of physical structures in varied contexts, several authors have employed a variety of techniques. As a result, the analytical methods have been developed and it has been shown that no single technique can be used to solve all types of non-linear problems with precision. Therefore, many different methods have emerged, some of which are sub-equation methods. High-dimensional partial differential equations have attracted academics' curiosity greatly in recent years. They also appear in modelling many phenomena in biology, chemistry, physics, engineering, mechanics, economy and many different branches [24, 25, 26, 27]. Soliton theory is a very efficient and competent way to describe non-linear features. Soliton theory has two basic routes to study and explain non-linear features. Solitons have the most remarkable properties of particles and waves simultaneously that reflect non-linear features in a well-organised and competent way. To study nature by framing non-linear evolution equations, along with their soliton solutions is immediate and unquestionable. Solitons keep their velocities, shapes and amplitudes unchanged even after interacting with others due to their perfectly elastic interaction. In this paper, some solutions including soliton, bright soliton, singular soliton, periodic wave solutions by the modified exponential Jacobi technique and rational $\exp(-\phi(\eta))$-expansion technique were also obtained. These good results show that the auxiliary methods are a powerful mathematical tool to handle non-linear integrable equations from nature. The complex integrable Kuralay governing system was offered to study the new auxiliary equation method for discovering multiple types of solitons [28]. Faridi et al. [29] studied the non-linear integrable model, namely the generalised Kadomtsev–Petviashvili modified equal width–Burgers equation, which utilised a weakly non-linear restoring forces, dispersion, small damping and non-linear media with dissipation to narrate the long wave propagation in chemical theory. Faridi et al. [30] analysed the dimensional elliptic non-linear Schrödinger equation under the influence of three different fractional operators and found the generalised fractional soliton solutions and propagation of magnetohydrodynamics fluid in sort of solition. The propagation of optical pulses in optical fibres and plasma has been examined for the Chen–Lee–Liu dynamical equation using the extended direct algebraic technique [31]. The combo of $\frac{G'}{G^2}$-expansion method and new extended direct algebraic method were used to find soliton solutions like periodic patterns with anti-peaked crests and anti-troughs, singular solution and mixed complex solitary shock solution to the fractional TFFME [32]. The solitonic patterns of the considered model were successfully surveyed by using two integrated analytical techniques, new extended direct algebraic and expansion method, to investigate the system of cold bosonic atoms in zig-zag optics lattices [33]. Chen and Li [34] constructed the optical soliton solutions of the well-known TFFME, which describes the propagation of polarisation in thin-film materials. The bifurcation, phase portrait and travelling wave solution of time-fractional TFFME with beta fractional derivative were studied [35]. The Zoomeron model was applied to various types of solitons arising in fluid mechanics, laser optics and non-linear physics [36]. New optical soliton solutions for the coupled conformable Fokas–Lenells equation with spatio-temporal dispersion were obtained via Atangana's derivative operator [37]. The resonant non-linear Schrödinger equation with Kerr law non-linearity considering inter-modal dispersion and spatio-temporal was investigated using the new extended direct algebraic method [38]. The variational iterative method with the Laplace transform was used to solve non-linear evolution problems of a simple pendulum and mass-spring oscillator, which represents the Duffing equation [39]. The natural decomposition method and Laplace decomposition method were studied to solve the second-order Painlevé equation [40]. An optimal Galerkin-homotopy asymptotic method was applied to the non-linear second-order boundary value problems (BVPs) [41]. An analytical analysis to solve the fractional differential equations has been offered by Manafian and Allahverdiyeva [42]. Different forms of optical soliton solutions to the Kudryashov's quintuple self-phase modulation were obtained by Li et al. [43]. New soliton solutions to the Van der Waals model through the improved $\exp(-\Omega)$-expansion method (IEFM) and

extended sinh-Gordon equation expansion method (EShGEEM) were attained [44]. Researchers worked on the network governance step-by-step method [46] and the neural network method [47]. Based on the invariant subspace method, the Lie symmetries including Riemann–Liouville and Erdelyi–Kober fractional derivatives of the time-fractional form of the Gardner equation have been studied [48]. Numerical analysis of bioconvective heat and mass transfer across a non-linear stretching sheet with hybrid nanofluids was investigated by Aneja et al. [49]. Inspired by the previous work, the motivation of the paper is to investigate the solitons and other forms of solutions by the modified exponential Jacobi technique and rational $\exp(-\phi(\eta))$-expansion technique. The outline of the paper is as follows. In Section 2, we transform the thin-film ferroelectric material (TFFM) equation to a non-linear equation of the non-linear ordinary differential equation. Furthermore, in Sections 3–6, different forms of solitary wave solutions have been established by the modified exponential Jacobi method (MEJM) and rational $\exp(-\phi(\eta))$-expansion method. Finally, the conclusions are provided in Section 7.

2. TRANSFORMING PDE TO ODE

For Eq. (1), x, t show the longitudinal and transverse coordinates. Using the next wave transformation, $u(x, t) = u(\eta), \eta = x - \lambda t$, where λ is arbitrary constants to be determined through the method's steps, leads to the ODE as follows:

$$\frac{\partial^2 u}{\partial t^2} = \lambda^2 \frac{d^2 u}{d\eta^2}, \quad J\Delta u = J\frac{\partial^2 u}{\partial x^2} = J\frac{d^2 u}{d\eta^2}, \quad 4cm \tag{2}$$

By substituting $\frac{d^2 u}{d\eta^2} = u''$ and simplifying Eq. (2), we get,

$$\left(\frac{m\lambda^2}{T^2} - J\right)u'' - [(p_2 - 2\mu)u + p_4 u^3 + p_6 u^5 = 0, \tag{3}$$

where wave speed is denoted by λ. By using the balance principle to the terms of Eq. (4) between ψ'' and ψ^5 we get $k + 2 = 5k$, which leads to $k = 1/2$. By utilising the transformation $u(\eta) = \sqrt{\psi(\eta)}$ in Eq. (3) once, with respect to η and zero-integration constant, leads to

$$\left(\frac{m\lambda^2}{T^2} - J\right)\left(-\frac{1}{4}\psi'^2 + \frac{1}{2}\psi\psi''\right) - (p_2 - 2\mu)\psi^2 - p_4\psi^3 - p_6\psi^4 = 0. \tag{4}$$

Deponing up the balance principle to the terms of Eq. (4) between ψ'' and ψ^4, we get $k = 1$.

3. THE MEJM

This part introduces the general properties of a new MEJM, which has been proposed by Aldhabani et al. [45]. The necessary steps for using this method are summarised as follows. Handling the inquired into model through the MEJM gets the following steps as mentioned earlier:
Step 1.

$$\mathcal{S}_1(F, F_x, F_t, F_{xx}, F_{tt}, \dots) = 0, \tag{5}$$

where $\mathcal{S}$ is a polynomial of F and its partial derivatives.
Step 2. Firstly, by utilising travelling wave transformation,

$$\xi = x - \lambda t, \tag{6}$$

where λ is the non-zero arbitrary value, allows to diminish Eq. (61) to an ODE of $F = F(\xi)$ in the below form

$$\mathcal{S}_2(F, F', -\lambda F', F'', \lambda^2 F'', \dots) = 0. \tag{7}$$

Step 3. The generated solutions of Eq. (61) are:

$$F(\xi) = A_0 + \sum_{i=1}^a A_i \left(\frac{\Omega'(\xi)}{\Omega(\xi)}\right)^i + \sum_{i=1}^a B_i \left(\frac{\Omega(\xi)}{\Omega'(\xi)}\right)^i, \tag{8}$$

where $\Omega'(\xi) = d\Omega(\xi)/d\xi$ and

$$\Omega(\xi) = \frac{\sigma_1 cn(\xi,k) + \sigma_2 sn(\xi,k)}{\sigma_3 cn(\xi,k) + \sigma_4 sn(\xi,k)}, \quad 5.9\ cm \tag{9}$$

where $cn(\xi, k)$ and $sn(\xi, k)$ are the Jacobi elliptic functions of index, k, and we obtain a few interesting and important relationships as follows:

$$sn^2(\xi, k) = 1 - cn^2(\xi, k), \tag{10}$$

$$dn^2(\xi, k) = 1 - sn^2(\xi, k),$$

$$\frac{d}{d\xi}cn(\xi, k) = -sn(\xi, k)dn(\xi, k),$$

$$\frac{d^2}{d\xi^2}cn(\xi, k) = -cn(\xi, k)dn^2(\xi, k) + kcn(\xi, k)sn^2(\xi, k),$$

$$\frac{d^3}{d\xi^3}cn(\xi, k) = -k^2 dn(\xi, k)sn^3(\xi, k) + 4k^2 cn^2(\xi, k)sn(\xi, k)dn(\xi, k) + sn(\xi, k)dn^3(\xi, k),$$

$$sn(\xi, 0) = \sin(\xi), \quad cn(\xi, 0) = \cos(\xi), \quad dn(\xi, 0) = 1,$$

$$sn(\xi, 1) = \tanh(\xi), \quad cn(\xi, 1) = sech(\xi), \quad dn(\xi, 1) = sech(\xi).$$

By utilising the balance tenet on Eq. (63), we can discover the value of a.

Step 4. Substituting Eq. (9) in Eq. (63) and collecting the coefficients of disparate orders in terms of $cn(\xi, k)$, $sn(\xi, k)$ and $dn(\xi, k)$ make a set of non-linear algebraic equations.

Step 5. In the next step, we solve the non-linear algebraic equations and get the needed results.

4. APPLICATION OF MEJM

It can be seen that the above governing differential equation is highly non-linear, and such non-linearity imposes some difficulties in the development of exact analytical techniques to generate closed-form solutions for the equation. Therefore, a modified exponential Jacobi scheme is used in this work. The MEJM, which is an analytical scheme for providing analytical solutions to non-linear ordinary differential equations, is adopted. Upon constructing the transformation and a new function, the following categories of solutions can be expressed:
The set of categories of solutions:

4.1. Set I

$$J = -\frac{4T^2 A_1^2 p_6 - 3\lambda^2 m}{3T^2}, \quad A_0 = \left(-1 + \sqrt{-k^2 + 1}\right)A_1,$$

$$p_2 = -\frac{4}{3}A_1^2 p_6\left(k^2 + 3\sqrt{1-k^2} - 2\right) + 2\mu, \tag{11}$$

$$A_1 = A_1, \quad B_1 = 0, \quad p_4 = -\frac{8}{3}\left(-1+\sqrt{-k^2+1}\right)A_1 p_6,$$
$$\sigma_2 = \sigma_3 = 0,$$

$$\psi_1 = \frac{A_1\left(\left(-1+\sqrt{-k^2+1}\right)(sn(\xi,k))^2 - dn(\xi,k)\right)}{sn(\xi,k)cn(\xi,k)},$$

$$u_1(x,t) = \left\{ \frac{\frac{9\,p_4+\sqrt{-192\,k^2 p_2 p_6+384\,k^2 p_6-192\,p_2 p_6+81\,p_4{}^2+384\,p_6}}{16 p_6(k^2+1)}\left(\left(-1+\sqrt{-k^2+1}\right)(sn(\xi,k))^2 - dn(\xi,k)\right)}{sn(\xi,k)cn(\xi,k)} \right\}^{\frac{1}{2}}, \tag{12}$$

$$\xi = x - \frac{\sqrt{3}T}{3m}\sqrt{m\left(\frac{\left(9\,p_4+\sqrt{-192\,k^2 p_2 p_6+384\,k^2 p_6-192\,p_2 p_6+81\,p_4{}^2+384\,p_6}\right)^2}{64\,p_6(k^2+1)^2}+3J\right)}\,t.$$

Some subgroups for relation (12):
Supposing k = 0 in Eq. (12) provides

$$u_2(x,t) = \left\{ \frac{\frac{9\,p_4+\sqrt{-192\,p_2 p_6+81\,p_4{}^2+384\,p_6}}{16 p_6}}{\sin(\xi)\cos(\xi)} \right\}^{\frac{1}{2}}, \tag{13}$$

$$\xi = x - \frac{\sqrt{3}T}{3m}\sqrt{m\left(\frac{\left(9\,p_4+\sqrt{-192\,p_2 p_6+81\,p_4{}^2+384\,p_6}\right)^2}{64\,p_6}+3J\right)}\,t.$$

Supposing k = 1 in Eq. (12) yields

$$u_3(x,t) =$$
$$\left\{ \frac{\frac{9\,p_4+\sqrt{-192 p_2 p_6+384 p_6-192\,p_2 p_6+81\,p_4{}^2+384\,p_6}}{32 p_6}\left(-(\tanh(\xi))^2 - sceh(\xi)\right)}{\tanh(\xi)sech(\xi)} \right\}^{\frac{1}{2}}, \tag{14}$$

$$\xi = x - \frac{\sqrt{3}T}{3m}\sqrt{m\left(\frac{\left(9\,p_4+\sqrt{-384\,p_2 p_6+81\,p_4{}^2+768\,p_6}\right)^2}{256\,p_6}+3J\right)}\,t.$$

The effect of analysis periodic solution when plots of u are given as in Fig. 1 with the following amounts:

$$p_2 = 1, p_4 = 2, p_6 = 3, m = 2, T = 3,$$
$$\mu = 2, J = 2, k = 0, \tag{15}$$
$$u = \sqrt{\frac{1}{\sin\left(3\sqrt{3}t-x\right)\cos\left(3\sqrt{3}t-x\right)}}. \tag{16}$$

for Eq. (16). We investigate the behaviour of general periodic and the periodic received from the mentioned technique, which is presented in Fig. 1. From the graph, it is ostensible that the periodic structure exhibits a stable propagation for the generalised non-local non-linearity as offered in Fig. 1. Also, the effect of analysis of the periodic solution when plots of u are given in Fig. 2 with the following amounts

$$p_2 = 1, p_4 = 2, p_6 = 3, m = 2,$$
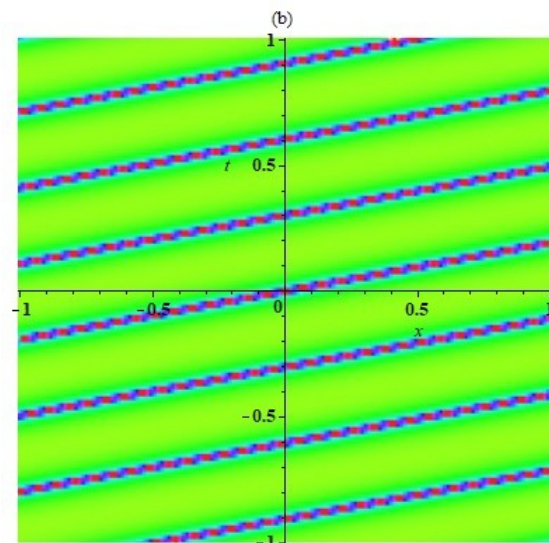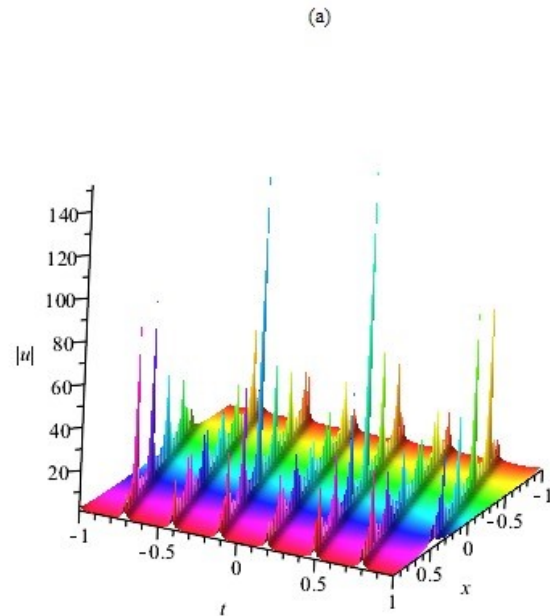$$T = 3, \mu = 2, J = 2, k = 1, \tag{17}$$
$$u = 1/24\sqrt{-\frac{(108+36\sqrt{41})(\tanh(\zeta)sech\zeta - sech(\zeta))}{\tanh(\zeta)sech(\zeta)}},$$

$$A_1 = \frac{9\,p_4+\sqrt{-192\,k^2 p_2 p_6+384\,k^2 p_6-192\,p_2 p_6+81\,p_4{}^2+384\,p_6}}{16 p_6(k^2+1)}.$$

As a result, the exact soliton solution is given by

$$\zeta = \frac{\sqrt{3}\sqrt{6\left(18+6\sqrt{41}\right)^2+27648t}}{96} - x. \tag{18}$$
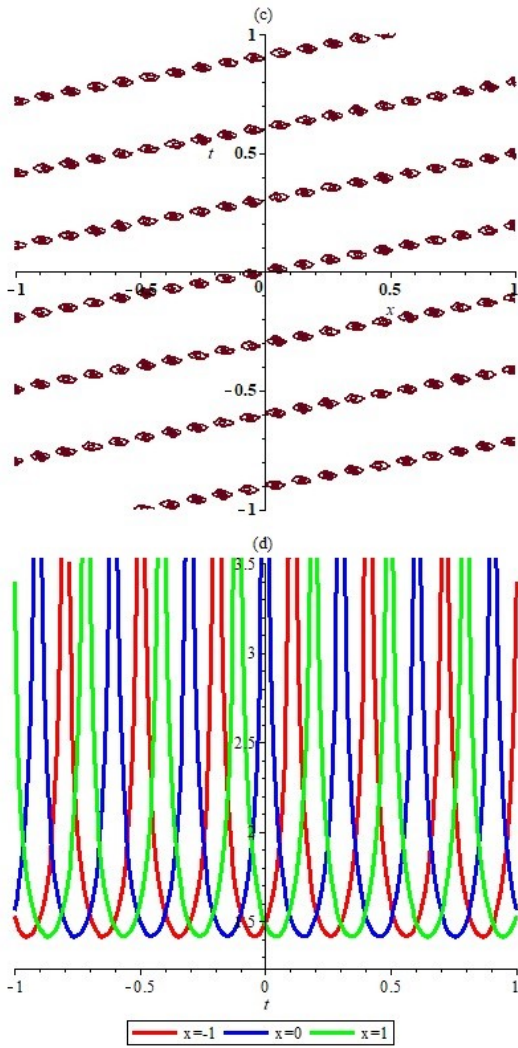
for Eq. (18).

(a)

(b)

Fig. 1. Plots of real solution (16) (a [3D plot], b [density plot], c [contour plot], d [2D plot]) for Graph u_1
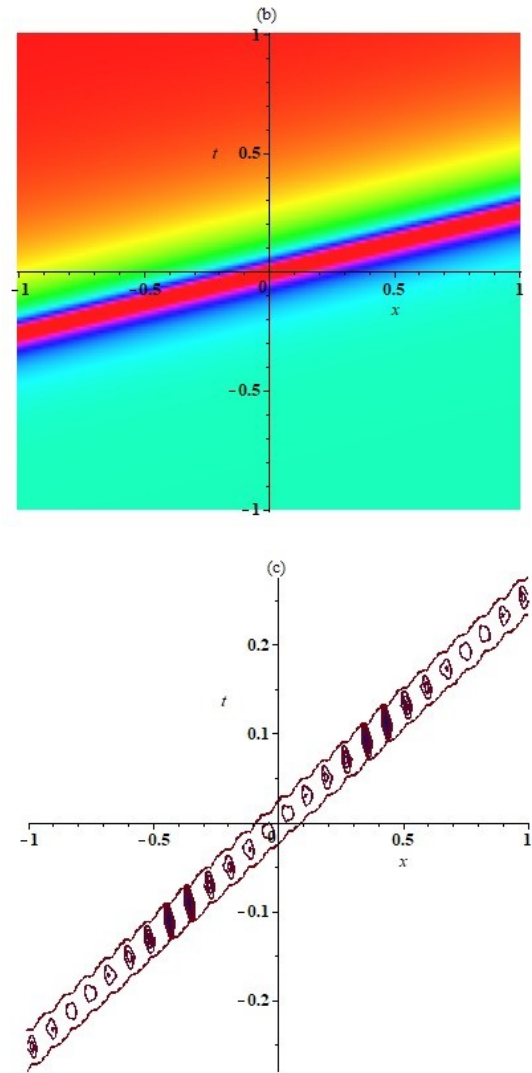

Fig. 2. Plots of real solution (18) (a [3D plot], b [density plot], c [contour plot], d [2D plot]) for Graph u_2

4.2. Set II

$$J = -1/3 \frac{-3k^4\lambda^2 m + 4T^2 B_1^2 p_6}{T^2 k^4},$$

$$A_0 = \frac{(-1+\sqrt{-k^2+1})B_1}{k^2}, \quad A_1 = 0, \tag{19}$$

$$p_2 = -2/3 \frac{2B_1^2 p_6(k^4-10k^2+10)-3k^6\mu+2\sqrt{-k^2+1}(-3k^4\mu+5k^2B_1^2 p_6-10B_1^2 p_6)+6k^4\mu}{k^4(k^2+2\sqrt{-k^2+1}-2)},$$

$$\psi_1 = \frac{B_1}{dn(\xi,k)}\left(\frac{(-1+\sqrt{-k^2+1})dn(\xi,k)}{k^2} - sn(\xi,k)cn(\xi,k)\right).$$

As a result, the exact soliton solution is given by

$$u_1(x,t) =$$

$$\left\{\frac{B_1}{dn(\xi,k)}\left(\frac{(-1+\sqrt{-k^2+1})dn(\xi,k)}{k^2} - sn(\xi,k)cn(\xi,k)\right)\right\}^{\frac{1}{2}}, \tag{20}$$

$$\xi = x - \frac{\sqrt{3}\sqrt{m(3Jk^4+4B_1^2 p_6)T}}{3mk^2}\, t.$$

Some subgroups for relation (20):
Supposing k = 1 in Eq. (20) provides

$$p_2 = 2/3 \frac{-2(1+\sqrt{-k^2+1})(-3k^4\mu+5k^2B_1^2 p_6-10B_1^2 p_6)-(3k^4\mu-2k^2B_1^2 p_6+10B_1^2 p_6)k^2}{k^4(-k^2+2\sqrt{-k^2+1}+2)},$$

$$\psi_1 = \frac{B_1}{dn(\xi,k)}\left(\frac{(1+\sqrt{-k^2+1})dn(\xi,k)}{k^2} - sn(\xi,k)cn(\xi,k)\right).$$

As a result, the exact soliton solution is given by

$$u_1(x,t) =$$

$$\left\{\frac{B_1}{dn(\xi,k)}\left(\frac{(1+\sqrt{-k^2+1})dn(\xi,k)}{k^2} - sn(\xi,k)cn(\xi,k)\right)\right\}^{\frac{1}{2}}, \tag{23}$$

$$\xi = x - \frac{\sqrt{3}\sqrt{m(3Jk^4+4B_1^2 p_6)T}}{3mk^2}\, t.$$

Some subgroups for relation (23):
Supposing k = 1 in Eq. (23) provides

$$u_2(x,t) =$$

$$\left\{\frac{B_1}{sech(x-\lambda t)}\left(sech(x-\lambda t) - \tanh(x-\lambda t)sech(x-\lambda t)\right)\right\}^{\frac{1}{2}}$$

$$\lambda = \frac{\sqrt{3}\sqrt{m(3J+4B_1^2 p_6)T}}{3m}. \tag{24}$$

4.4. Set IV

$$J = -1/3 \frac{4T^2 A_1^2 p_6 - 3\lambda^2 m}{T^2}, \quad A_0 = 2A_1, \quad B_1 = k^2 A_1, \tag{25}$$

$$p_2 = -4/3 A_1^2 k^2 p_6 + \frac{20A_1^2 p_6}{3} + 2\mu,$$

$$p_4 = -16/3 A_1 p_6, \quad \sigma_2 = \sigma_3 = 0, \quad A_1 = -4\frac{2\mu-p_2}{p_4(k^2-5)},$$

$$B_1 = B_1, \quad p_4 = -8/3 \frac{(-1+\sqrt{-k^2+1})B_1 p_6}{k^2}, \quad \sigma_2 = \sigma_3 = 0,$$

$$u_2(x,t) = \left\{B_1\left[-1 - \tanh\left(x - \frac{\sqrt{3}\sqrt{m(4B_1^2 p_6+3J)}T}{3m}t\right)\right]\right\}^{\frac{1}{2}}. \tag{21}$$

4.3. Set III

$$J = -1/3 \frac{-3k^4\lambda^2 m + 4T^2 B_1^2 p_6}{T^2 k^4},$$

$$A_0 = \frac{(-1+\sqrt{-k^2+1})B_1}{k^2}, \quad A_1 = 0, \tag{22}$$

$$B_1 = B_1, \quad p_4 = -8/3 \frac{(1+\sqrt{-k^2+1})B_1 p_6}{k^2}, \quad \sigma_2 = \sigma_3 = 0,$$

$$\psi_1 = -4\frac{(2\mu-p_2)((sn(\xi,k))^4 k^2+2\,sn(\xi,k)cn(\xi,k)dn(\xi,k)-1)}{p_4(k^2-5)sn(\xi,k)cn(\xi,k)dn(\xi,k)},$$

$$\xi = x - 1/3\frac{\sqrt{3}T}{m}\sqrt{m\left(3\frac{2\mu-p_2}{k^2-5}+3J\right)}\, t.$$

As a result the exact soliton solution is given by

$$u_1(x,t) =$$

$$\left\{-4\frac{(2\mu-p_2)((sn(\xi,k))^4 k^2+2\,sn(\xi,k)cn(\xi,k)dn(\xi,k)-1)}{p_4(k^2-5)sn(\xi,k)cn(\xi,k)dn(\xi,k)}\right\}^{\frac{1}{2}}, \tag{26}$$

$$\xi = x - \frac{\sqrt{3}T}{3m}\sqrt{m\left(3\frac{2\mu-p_2}{k^2-5}+3J\right)}\, t.$$

Some subgroups for relation (26):
Supposing k = 0 in Eq. (26) provides

$$u_2(x,t) = \left\{4\frac{(2\mu-p_2)(2\sin(x-\lambda t)\cos(x-\lambda t)-1)}{5p_4\sin(x-\lambda t)\cos(x-\lambda t)}\right\}^{\frac{1}{2}},$$

$$\lambda = \frac{\sqrt{3}T}{3m}\sqrt{m\left(3J-3\frac{2\mu-p_2}{5}\right)}. \tag{27}$$

Supposing k = 1 in Eq. (26) supplies

$$u_3(x,t) =$$

$$\left\{\frac{(2\mu-p_2)((\tanh(x-\lambda t))^4+2\tanh(x-\lambda t)sech^2(x-\lambda t)-1)}{p_4\tanh(x-\lambda t)sech^2(x-\lambda t)}\right\}^{\frac{1}{2}}, \tag{28}$$

$$\lambda = \frac{\sqrt{3}T}{3m}\sqrt{m\left(3J-3\frac{2\mu-p_2}{4}\right)}.$$

The effect of analysis periodic solution when plots of u are given in Fig. 3 with the following amounts

$$p_2 = 5, p_6 = 1, m = 2, T = 3,$$

$$\mu = 2, J = 2, k = 0, \tag{29}$$

$$u = \frac{\sqrt{10}}{10}\sqrt{2\,\frac{\sqrt{15}\left(-1-\sin(3/5\,\sqrt{110}t-x)\right)}{\sin(3/5\,\sqrt{110}t-x)\left(\left(\cos(3/10\,\sqrt{110}t-x)\right)^2+\left(\sin(3/10\,\sqrt{110}t-x)\right)^2\right)}}, \tag{30}$$

for Eq. (30). We investigate the behaviour of general periodic and periodic received from the mentioned technique, which is presented in Fig. 3. From the graph, it is ostensible that the periodic structure exhibits a stable propagation for the generalised non-

local non-linearity as offered in Fig. 3. Also, the effect of analysis periodic solution when plots of u are given in Fig. 4 with the following amounts

$$p_2 = 5, p_6 = 1, m = 2, T = 3, \mu = 2, J = 2, k = 1, \tag{31}$$

$$u = 1/2\sqrt{\frac{\sqrt{3}\left(\left(\tanh(9/4\,\sqrt{2}t-x)\right)^4-2\tanh(9/4\,\sqrt{2}t-x)\left(sech(9/4\,\sqrt{2}t-x)\right)^2-1\right)}{\tanh(9/4\,\sqrt{2}t-x)\left(sech(9/4\,\sqrt{2}t-x)\right)^2}}, \tag{32}$$
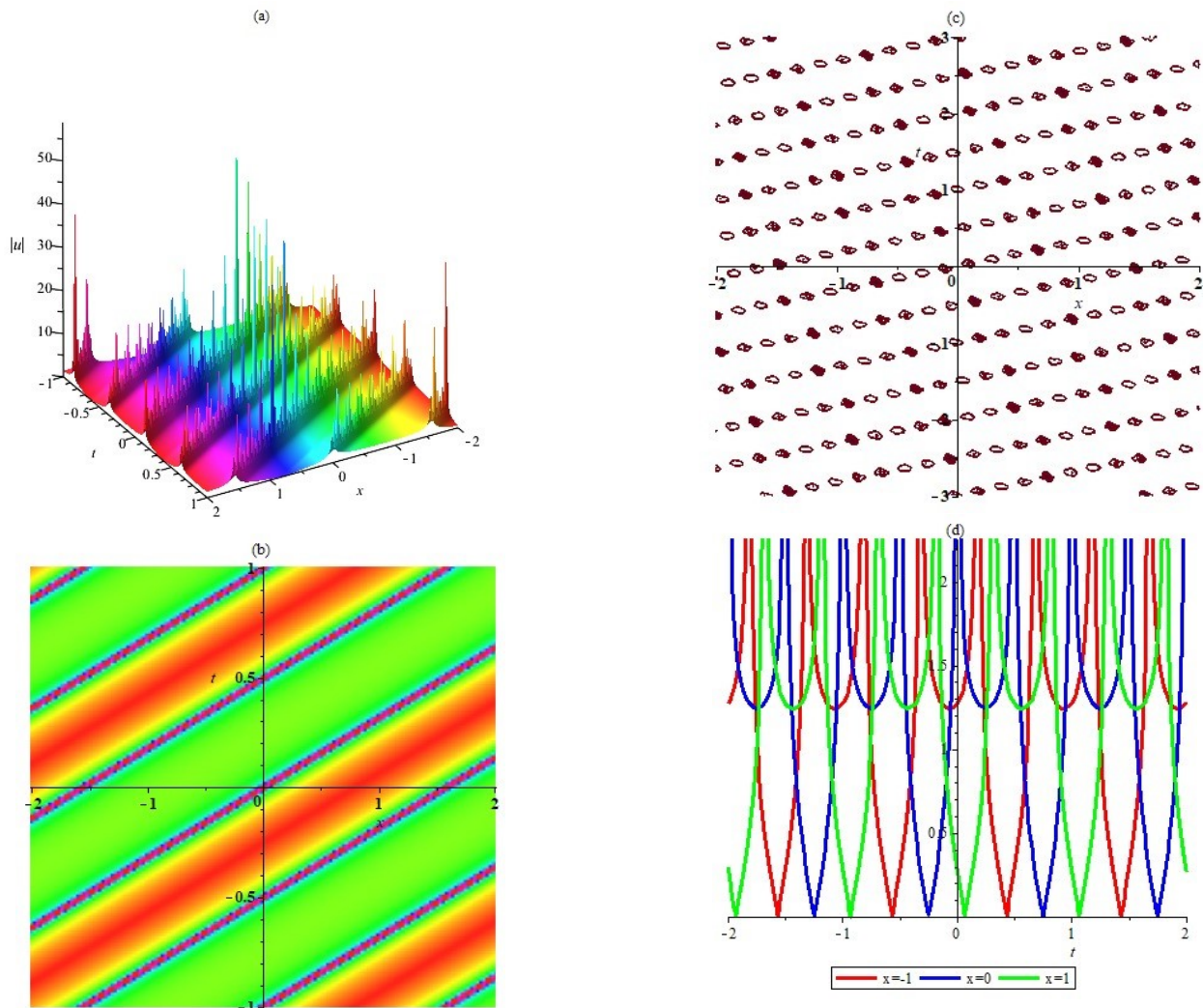
for Eq. (32).

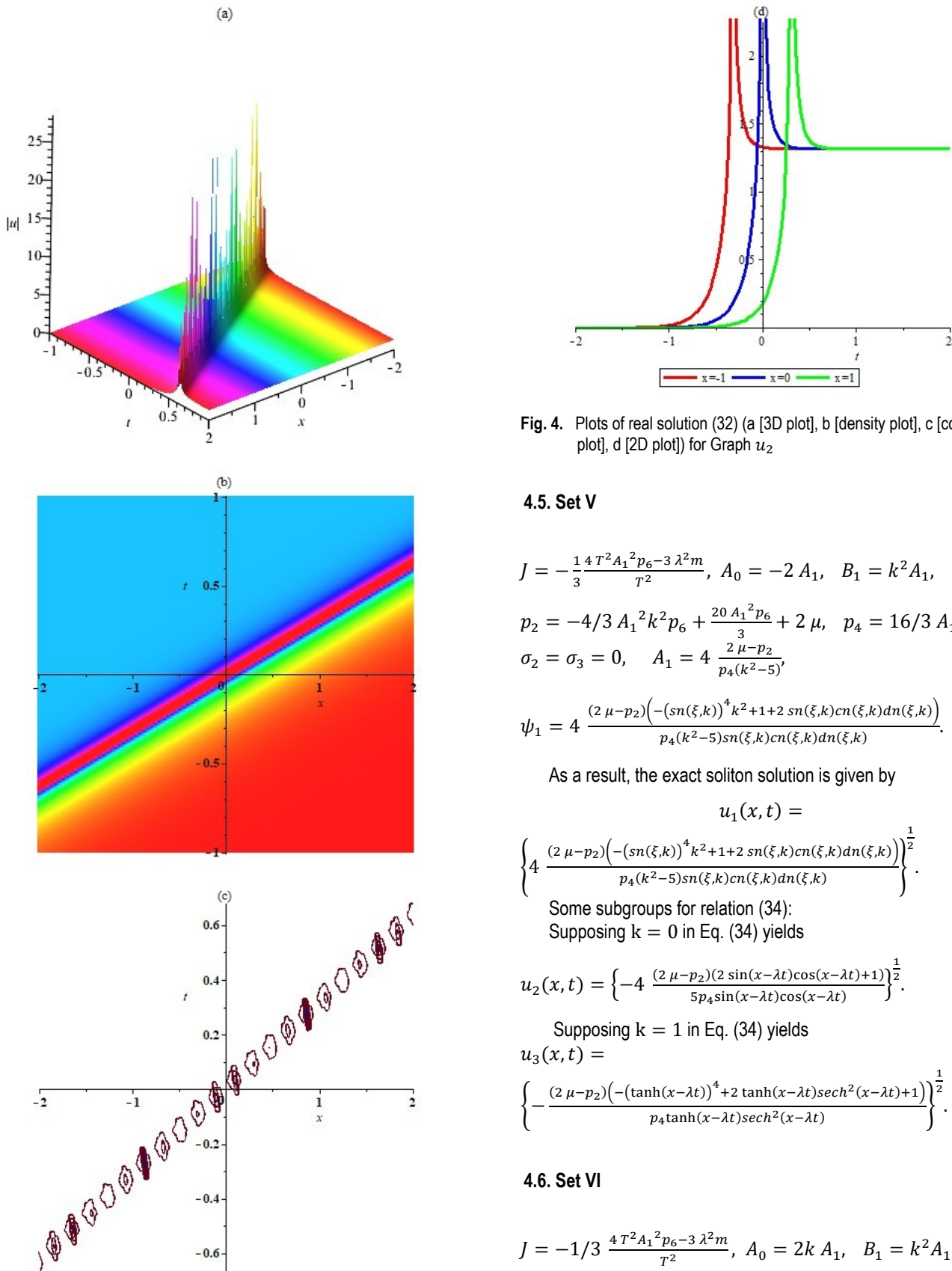**Fig. 3.** Plots of real solution (30) (a [3D plot], b [density plot], c [contour plot], d [2D plot]) for Graph u_1.

(a)

(d)

Fig. 4. Plots of real solution (32) (a [3D plot], b [density plot], c [contour plot], d [2D plot]) for Graph u_2

(b)

4.5. Set V

$$J = -\frac{1}{3}\frac{4\,T^2A_1{}^2p_6 - 3\,\lambda^2 m}{T^2},\ \ A_0 = -2\,A_1,\ \ B_1 = k^2A_1, \quad (33)$$

$$p_2 = -4/3\,A_1{}^2k^2p_6 + \frac{20\,A_1{}^2p_6}{3} + 2\,\mu,\quad p_4 = 16/3\,A_1p_6,$$

$$\sigma_2 = \sigma_3 = 0,\quad A_1 = 4\,\frac{2\,\mu - p_2}{p_4(k^2-5)},$$

$$\psi_1 = 4\,\frac{(2\,\mu - p_2)\left(-(sn(\xi,k))^4k^2 + 1 + 2\,sn(\xi,k)cn(\xi,k)dn(\xi,k)\right)}{p_4(k^2-5)sn(\xi,k)cn(\xi,k)dn(\xi,k)}.$$

As a result, the exact soliton solution is given by

$$u_1(x,t) =$$

$$\left\{4\,\frac{(2\,\mu - p_2)\left(-(sn(\xi,k))^4k^2 + 1 + 2\,sn(\xi,k)cn(\xi,k)dn(\xi,k)\right)}{p_4(k^2-5)sn(\xi,k)cn(\xi,k)dn(\xi,k)}\right\}^{\frac{1}{2}}. \quad (34)$$

Some subgroups for relation (34):
Supposing k = 0 in Eq. (34) yields

$$u_2(x,t) = \left\{-4\,\frac{(2\,\mu - p_2)(2\sin(x-\lambda t)\cos(x-\lambda t)+1)}{5p_4\sin(x-\lambda t)\cos(x-\lambda t)}\right\}^{\frac{1}{2}}. \quad (35)$$

Supposing k = 1 in Eq. (34) yields
$$u_3(x,t) =$$

$$\left\{-\frac{(2\,\mu - p_2)\left(-(\tanh(x-\lambda t))^4 + 2\tanh(x-\lambda t)sech^2(x-\lambda t)+1\right)}{p_4\tanh(x-\lambda t)sech^2(x-\lambda t)}\right\}^{\frac{1}{2}}. \quad (36)$$

4.6. Set VI

$$J = -1/3\,\frac{4\,T^2A_1{}^2p_6 - 3\,\lambda^2 m}{T^2},\ \ A_0 = 2k\,A_1,\ \ B_1 = k^2A_1, \quad (37)$$

$$p_2 = 4/3\,A_1{}^2p_6(5\,k^2 - 1) + 2\,\mu,\quad p_4 = -16/3k\,A_1p_6,$$

$$\sigma_2 = \sigma_3 = 0,\quad A_1 = 4\,\frac{(2\,\mu - p_2)k}{p_4(k^2-5)},$$

$$\psi_1 = 4\,\frac{(2\,\mu - p_2)k\left((sn(\xi,k))^4k^2 + 2\,ksn(\xi,k)cn(\xi,k)dn(\xi,k)-1\right)}{p_4(k^2-5)sn(\xi,k)cn(\xi,k)dn(\xi,k)}.$$

(c)

As a result, the exact soliton solution is given by

$$u_1(x,t) =$$

$$\left\{4\,\frac{(2\,\mu-p_2)k\big(\big(sn(\xi,k)\big)^4 k^2 + 2\,k sn(\xi,k)cn(\xi,k)dn(\xi,k)-1\big)}{p_4(k^2-5)sn(\xi,k)cn(\xi,k)dn(\xi,k)}\right\}^{\frac{1}{2}}. \quad (38)$$

Some subgroups for relation (38):
Supposing $k = 1$ in Eq. (38) provides

$$u_2(x,t) =$$

$$\left\{-\frac{(2\,\mu-p_2)\big(\big(\tanh(x-\lambda t)\big)^4 + 2\,\tanh(x-\lambda t)sech^2(x-\lambda t)-1\big)}{p_4\tanh(x-\lambda t)sech^2(x-\lambda t)}\right\}^{\frac{1}{2}}. \quad (39)$$

The effect of analysis soliton solution when plots of u are given in Fig. 5 with the following amounts

$$p_2 = 5, p_6 = 1, m = 2, T = 3, \mu = 2, J = 2, k = 1, \quad (40)$$

$$u =$$

$$\frac{\big(\tanh(\sqrt{3}\sqrt{7}t-x)\big)^4 - 2\,\tanh(\sqrt{3}\sqrt{7}t-x)\big(sech(\sqrt{3}\sqrt{7}t-x)\big)^2 - 1}{\tanh(\sqrt{3}\sqrt{7}t-x)\big(sech(\sqrt{3}\sqrt{7}t-x)\big)^2\big(\big(sech(\sqrt{3}\sqrt{7}t-x)\big)^2 + \big(\tanh(\sqrt{3}\sqrt{7}t-x)\big)^2\big)},$$

$$(41)$$

for Eq. (41).

(a)

(b)

(c)

(d)
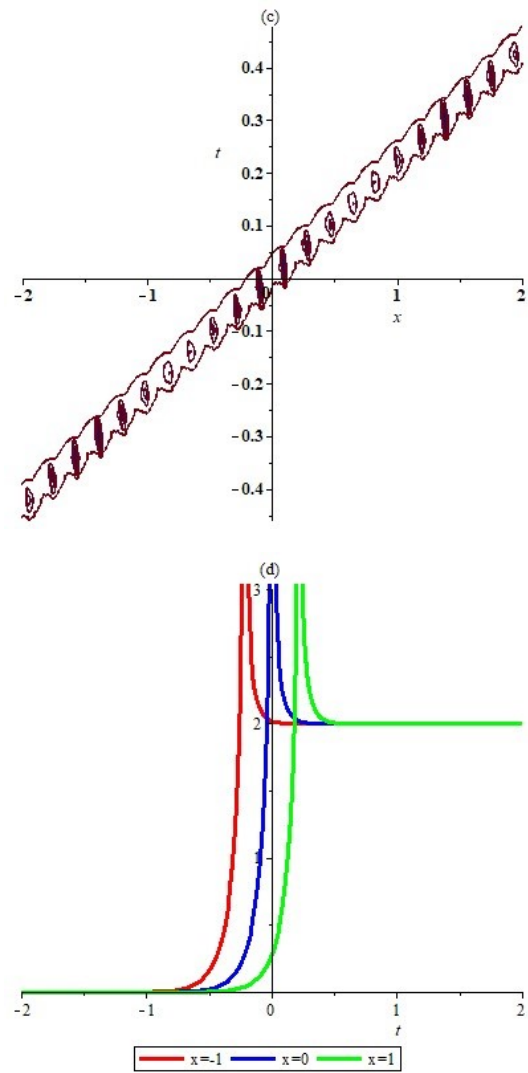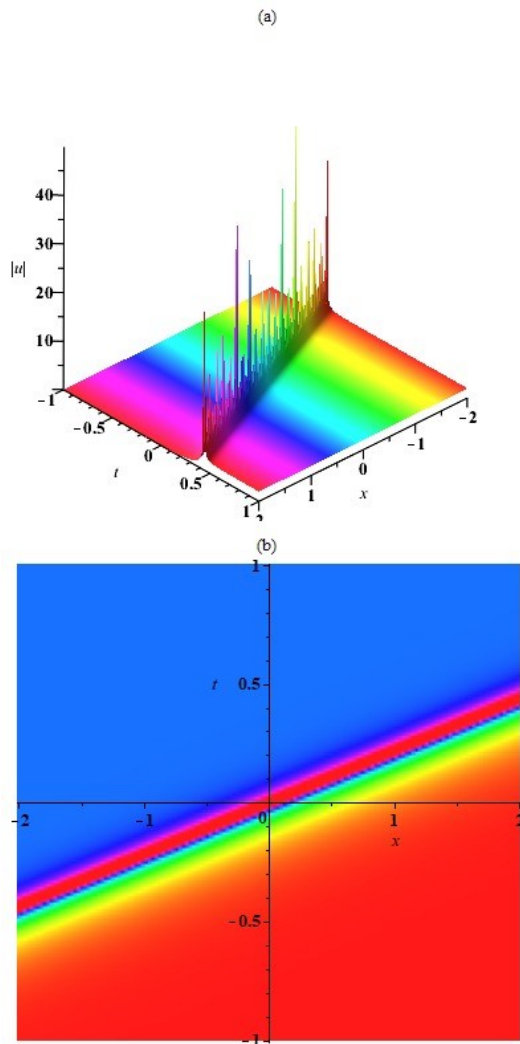
Fig. 5. Plots of real solution (39) (a [3Dplot], b [density plot], c [contour plot], d [2D plot]) for Graph u_2

4.7. Set VII

$$J = -1/3\,\frac{4\,T^2 A_1{}^2 p_6 - 3\,\lambda^2 m}{T^2}, \quad A_0 = -2k\,A_1, \quad B_1 = k^2 A_1, (42)$$

$$p_2 = 4/3\,A_1{}^2 p_6(5\,k^2 - 1) + 2\,\mu, \quad p_4 = 16/3k\,A_1 p_6,$$

$$\sigma_2 = \sigma_3 = 0, \quad A_1 = -4\,\frac{(2\,\mu-p_2)k}{p_4(k^2-5)},$$

$$\psi_1 = -4\,\frac{(2\,\mu-p_2)k\big(\big(sn(\xi,k)\big)^4 k^2 - 2\,k sn(\xi,k)cn(\xi,k)dn(\xi,k)-1\big)}{p_4(k^2-5)sn(\xi,k)cn(\xi,k)dn(\xi,k)}.$$

As a result, the exact soliton solution is given by

$$u_1(x,t) =$$

$$\left\{-4\,\frac{(2\,\mu-p_2)k\big(\big(sn(\xi,k)\big)^4 k^2 - 2\,k sn(\xi,k)cn(\xi,k)dn(\xi,k)-1\big)}{p_4(k^2-5)sn(\xi,k)cn(\xi,k)dn(\xi,k)}\right\}^{\frac{1}{2}}. \quad (43)$$

Some subgroups for relation (43):
Supposing $k = 1$ in Eq. (43) supplies

$$u_2(x,t) =$$

$$\left\{\frac{(2\,\mu-p_2)\big(\big(\tanh(x-\lambda t)\big)^4 - 2\,\tanh(x-\lambda t)sech^2(x-\lambda t)-1\big)}{p_4\tanh(x-\lambda t)sech^2(x-\lambda t)}\right\}^{\frac{1}{2}}. \quad (44)$$

4.8. Set VIII

$$J = -1/3 \frac{4 T^2 A_1^2 p_6 - 3 \lambda^2 m}{T^2}, \; A_0 = 2ik A_1,$$

$$B_1 = -k^2 A_1, \; i = \sqrt{-1}, \tag{45}$$

$$p_2 = -4/3 A_1^2 p_6 (4 k^2 + 1) + 2 \mu,$$

$$p_4 = 16/3 ik A_1 p_6, \quad \sigma_2 = \sigma_3 = 0, \quad A_1 = \frac{-4 ik(2 \mu - p_2)}{p_4(4 k^2 + 1)},$$

$$\psi_1 = \frac{-4 ik(2 \mu - p_2)\left(-(sn(\xi,k))^4 k^2 + 2 iksn(\xi,k)cn(\xi,k)dn(\xi,k) + 2 k^2 (sn(\xi,k))^2 - 1\right)}{p_4(4 k^2 + 1)sn(\xi,k)cn(\xi,k)dn(\xi,k)}.$$

As a result, the exact soliton solution is given by

$$u_1(x,t) = \left\{ \frac{-4 ik(2 \mu - p_2)\left(-(sn(\xi,k))^4 k^2 + 2 iksn(\xi,k)cn(\xi,k)dn(\xi,k) + 2 k^2 (sn(\xi,k))^2 - 1\right)}{p_4(4 k^2 + 1)sn(\xi,k)cn(\xi,k)dn(\xi,k)} \right\}^{\frac{1}{2}}. \tag{46}$$

Some subgroups for relation (46):
Supposing k = 1 in Eq. (46) yields

$$u_2(x,t) = \left\{ \frac{-4 i(2 \mu - p_2)\left(-(\tanh(x - \lambda t))^4 + 2 i\tanh(x - \lambda t)sech^2(x - \lambda t) + 2 (\tanh(x - \lambda t))^2 - 1\right)}{5 p_4 \tanh(x - \lambda t)sech^2(x - \lambda t)} \right\}^{\frac{1}{2}}. \tag{47}$$

4.9. Set IX

$$J = -\frac{1}{3} \frac{4 T^2 A_1^2 p_6 - 3 \lambda^2 m}{T^2},$$
$$A_0 = 2\sqrt{1 - k^2} A_1,$$
$$B_1 = -k^2 A_1, \; i = \sqrt{-1}, \tag{48}$$

$$p_2 = -\frac{4}{3} A_1^2 p_6 (4 k^2 - 5) + 2 \mu,$$
$$p_4 = -\frac{16}{3} \sqrt{1 - k^2} A_1 p_6,$$
$$\sigma_2 = \sigma_3 = 0,$$
$$A_1 = -4 \frac{\sqrt{-k^2 + 1}(2 \mu - p_2)}{p_4(4 k^2 - 5)},$$

$$\psi_1 = -4 \frac{\sqrt{1 - k^2}(2 \mu - p_2)\left(-(sn(\xi,k))^4 k^2 + 2 \sqrt{1 - k^2} sn(\xi,k)cn(\xi,k)dn(\xi,k) + 2 k^2 (sn(\xi,k))^2 - 1\right)}{p_4(4 k^2 - 5)sn(\xi,k)cn(\xi,k)dn(\xi,k)}.$$

As a result, the exact soliton solution is given by

$$u_1(x,t) = \left\{ \frac{4\sqrt{1 - k^2}(2\mu - p_2)\left(2\sqrt{1 - k^2} sn(\xi,k)cn(\xi,k)dn(\xi,k) - sn^4(\xi,k)k^2 + 2k^2 sn^2(\xi,k) - 1\right)}{p_4(5 - 4 k^2)sn(\xi,k)cn(\xi,k)dn(\xi,k)} \right\}^{\frac{1}{2}}. \tag{49}$$

Some subgroups for relation (49):
Supposing k = 0 in Eq. (49) provides

$$u_2(x,t) = \left\{ 4 \frac{(2 \mu - p_2)(2 \sin(x - \lambda t)\cos(x - \lambda t) - 1)}{5 p_4 \sin(x - \lambda t)\cos(x - \lambda t)} \right\}^{\frac{1}{2}}. \tag{50}$$

4.10. Set X

$$J = -\frac{4 T^2 A_1^2 p_6 - 3 \lambda^2 m}{3 T^2}, \; A_0 = \left(-1 + \sqrt{1 - k^2}\right) A_1$$

$$B_1 = 0, \; p_2 = -\frac{4}{3} A_1^2 p_6 \left(k^2 + 3 \sqrt{1 - k^2} - 2\right) + 2 \mu, \tag{51}$$

$$p_4 = -8/3 \left(-1 + \sqrt{-k^2 + 1}\right) A_1 p_6, \quad \sigma_1 = \sigma_4 = 0,$$

$$A_1 = -2 \frac{2 \mu \sqrt{-k^2 + 1} - p_2 \sqrt{-k^2 + 1} - 2 \mu + p_2}{p_4 \left(k^2 + 3 \sqrt{-k^2 + 1} - 2\right)},$$

$$\psi_1 = -2 \frac{\left(2 \mu \sqrt{-k^2 + 1} - p_2 \sqrt{-k^2 + 1} - 2 \mu + p_2\right)\left(\left(-1 + \sqrt{-k^2 + 1}\right)sn(\xi,k)cn(\xi,k) + dn(\xi,k)\right)}{p_4 \left(k^2 + 3 \sqrt{-k^2 + 1} - 2\right)cn(\xi,k)sn(\xi,k)}.$$

As a result, the exact soliton solution is given by

$$u_1(x,t) = \left\{ -2 \frac{\left(2 \mu \sqrt{-k^2 + 1} - p_2 \sqrt{-k^2 + 1} - 2 \mu + p_2\right)\left(\left(-1 + \sqrt{-k^2 + 1}\right)sn(\xi,k)cn(\xi,k) + dn(\xi,k)\right)}{p_4 \left(k^2 + 3 \sqrt{-k^2 + 1} - 2\right)cn(\xi,k)sn(\xi,k)} \right\}^{\frac{1}{2}}. \tag{52}$$

Some subgroups for relation (52):
Supposing k = 0 in Eq. (52) yields

$$u_2(x,t) = \left\{ -2 \frac{(2 \mu - p_2 - 2 \mu + p_2)}{p_4 \cos(x - \lambda t)\sin(x - \lambda t)} \right\}^{\frac{1}{2}}. \tag{53}$$

Supposing $k = 1$ in Eq. (52) provides

$$u_2(x,t) = \left\{ 2\, \frac{(-2\,\mu + p_2)(-\tanh(x-\lambda t)\,sech(x-\lambda t) + sech(x-\lambda t))}{p_4\, sech(x-\lambda t)\tanh(x-\lambda t)} \right\}^{\frac{1}{2}}. \tag{54}$$

4.11. Set XI

$$J = -1/3\, \frac{-3\,k^4\lambda^2 m + 4\,T^2 B_1{}^2 p_6}{T^2 k^4}, \quad A_0 = \frac{\left(-1+\sqrt{-k^2+1}\right)B_1}{k^2}, \quad A_1 = 0,\ \sigma_1 = \sigma_4 = 0, \tag{55}$$

$$p_2 = -2/3\, \frac{2\,B_1{}^2 p_6(k^4 - 10\,k^2 + 10) - 3\,k^4\mu\,(k^2 - 2) + 2\,\sqrt{-k^2+1}\left(-3\,k^4\mu + 5\,k^2 B_1{}^2 p_6 - 10\,B_1{}^2 p_6\right)}{k^4\left(k^2 + 2\,\sqrt{-k^2+1} - 2\right)},$$

$$p_4 = -\frac{8}{3}\, \frac{\left(-1+\sqrt{1-k^2}\right)B_1 p_6}{k^2}, \quad B_1 = -2\, \frac{k^2\left(-\sqrt{1-k^2}(k^2-4)(p_2-2\mu) - 6k^2\mu + 3p_2 k^2 + 8\mu - 4\,p_2\right)}{p_4\left(k^4 + 5(k^2-2)\sqrt{1-k^2} - 10k^2 + 10\right)},$$

$$\psi_1 = 2\, \frac{(p_2 - 2\,\mu)\left(\sqrt{-k^2+1}(k^2-4) - 3\,k^2 + 4\right)\left(sn(\xi,k)cn(\xi,k)k^2 + dn(\xi,k)\sqrt{-k^2+1} - dn(\xi,k)\right)}{p_4\left(k^4 + 5\,(k^2-2)\sqrt{-k^2+1} + 10\right)dn(\xi,k)}.$$

As a result, the exact soliton solution is given by

$$u_1(x,t) = \left\{ \frac{2(p_2 - 2\,\mu)\left(\sqrt{1-k^2}(k^2-4) - 3\,k^2 + 4\right)\left(sn(\xi,k)cn(\xi,k)k^2 + dn(\xi,k)\sqrt{1-k^2} - dn(\xi,k)\right)}{p_4\left(k^4 + 5\,(k^2-2)\sqrt{1-k^2} + 10\right)dn(\xi,k)} \right\}^{\frac{1}{2}}. \tag{56}$$

Some subgroups for relation (56):
Supposing k = 1 in Eq. (56) yields

$$u_2(x,t) = \left\{ 2/11\, \frac{(p_2 - 2\,\mu)(\tanh(x-\lambda t)\,sech(x-\lambda t) - sech(x-\lambda t))}{p_4\, sech(x-\lambda t)} \right\}^{\frac{1}{2}}. \tag{57}$$

4.12. Set XII

$$J = -1/3\, \frac{-3\,k^4\lambda^2 m + 4\,T^2 B_1{}^2 p_6}{T^2 k^4}, \quad A_0 = \frac{\left(1+\sqrt{-k^2+1}\right)B_1}{k^2}, \quad A_1 = 0,\ \sigma_1 = \sigma_4 = 0, \tag{58}$$

$$p_2 = -2/3\, \frac{-2\,B_1{}^2 p_6(k^4 - 10\,k^2 + 10) + 3\,k^6\mu + 2\,\sqrt{-k^2+1}\left(-3\,k^4\mu + 5\,k^2 B_1{}^2 p_6 - 10\,B_1{}^2 p_6\right) - 6\,k^4\mu}{k^4\left(-k^2 + 2\,\sqrt{-k^2+1} + 2\right)},$$

$$p_4 = -8/3\, \frac{\left(1+\sqrt{-k^2+1}\right)B_1 p_6}{k^2}, \quad B_1 = 2\, \frac{k^2\left(-\sqrt{-k^2+1}(k^2-4)(p_2-2\,\mu) + 6\,k^2\mu - 3\,p_2 k^2 - 8\,\mu + 4\,p_2\right)}{p_4\left(-k^4 + 5\,(k^2-2)\sqrt{-k^2+1} + 10\,k^2 - 10\right)},$$

$$\psi_1 = \frac{2(2\,\mu - p_2)\left(\sqrt{1-k^2}(k^2-4) + 3k^2 - 4\right)\left(sn(\xi,k)cn(\xi,k)k^2 + dn(\xi,k)\sqrt{-k^2+1} + dn(\xi,k)\right)}{p_4\left(-k^4 + 5\sqrt{1-k^2}k^2 + 10k^2 - 10\sqrt{1-k^2} - 10\right)dn(\xi,k)}.$$

As a result, the exact soliton solution is given by

$$u_1(x,t) = \left\{ \frac{2(2\,\mu - p_2)\left(\sqrt{1-k^2}(k^2-4) + 3k^2 - 4\right)\left(sn(\xi,k)cn(\xi,k)k^2 + dn(\xi,k)\sqrt{1-k^2} + dn(\xi,k)\right)}{p_4\left(-k^4 + 5\sqrt{1-k^2}k^2 + 10k^2 - 10\sqrt{1-k^2} - 10\right)dn(\xi,k)} \right\}^{\frac{1}{2}}. \tag{59}$$

Some subgroups for relation (59):
Supposing k = 1 in Eq. (59) supplies

$$u_2(x,t) = \left\{ -2\, \frac{(p_2 - 2\,\mu)(\tanh(x-\lambda t)\,sech(x-\lambda t) + sech(x-\lambda t))}{p_4\, sech(x-\lambda t)} \right\}^{\frac{1}{2}}. \tag{60}$$

5. THE $\exp(-\phi(\eta))$-EXPANSION METHOD

Handling the investigated model through the rational $\exp(-\phi(\eta))$-expansion method gets the following steps as mentioned earlier:
Step 1.

$$\mathcal{S}_1(\psi, \psi_x, \psi_t, \psi_{xx}, \psi_{tt}, \dots) = 0, \tag{61}$$

where $\mathcal{S}$ is a polynomial of ψ and its partial derivatives.

Step 2. Firstly, by utilising travelling wave transformation

$$\eta = x - \lambda t + \theta_0, \quad 8.5 cm \tag{62}$$

where λ is the non-zero arbitrary value, allows to diminish Eq. (61) to an ODE of $\psi = \psi(\eta)$ in the below form,

$$\mathcal{S}_2(\psi, \kappa\psi', \omega\psi', \kappa^2\psi'', \omega^2\psi'', \dots) = 0. \tag{63}$$

Step 3. The generated solutions of (61) are:

$$\psi(\eta) = \frac{\sum_{i=0}^{N} A_i(e^{-\phi(\eta)})^i}{\sum_{i=0}^{M} B_i(e^{-\phi(\eta)})^i}, \tag{64}$$

where $A_i, B_i (0 \leq j \leq N, M)$, are the parameters to be determined $A_N, B_M \neq 0$, and, $\phi = \phi(\eta)$ satisfying the ODE given below

$$\phi' = w_1 e^{\phi} + e^{-\phi} + w_2, \quad \phi' = \frac{d\phi}{d\eta}. \tag{65}$$

The particular solutions of Eq. (65) will be read as:

Solution-1: When $w_1 \neq 0$ and $w_2^2 - 4w_1 > 0$, therefore we attain

$$\phi(\eta) = \ln\left(-\frac{\sqrt{w_2^2 - 4w_1}}{2w_1}\tanh\left(\frac{\sqrt{w_2^2 - 4w_1}}{2}(\eta + E)\right) - \frac{w_2}{2w_1}\right).$$

Solution-2: When $w_1 \neq 0$ and $w_2^2 - 4w_1 < 0$, therefore we attain

$$\phi(\eta) = \ln\left(\frac{\sqrt{-w_2^2 + 4w_1}}{2w_1}\tan\left(\frac{\sqrt{-w_2^2 + 4w_1}}{2}(\eta + E)\right) - \frac{w_2}{2w_1}\right).$$

Solution-3: When $w_1 = 0$, $w_2 \neq 0$, and $w_2^2 - 4w_1 > 0$, therefore we attain

$$\Phi(\eta) = -\ln\left(\frac{w_2}{\exp(w_2(\eta + E)) - 1}\right).$$

Solution-4: When $w_1 \neq 0$, $w_2 \neq 0$, and $w_2^2 - 4w_1 = 0$, therefore we attain

$$\phi(\eta) = \ln\left(-\frac{2w_2(\xi + E) + 4}{w_2^2(\eta + E)}\right).$$

Solution-5: When $w_1 = 0$, $w_2 = 0$, and $w_2^2 - 4w_1 = 0$, therefore we attain $\phi(\eta) = \ln(\eta + E)$, where $A_j (0 \leq j \leq$

$$u_1(x, t) =$$

$$\left\{1/8 B_0 \frac{4 A_1 p_6 w_2 - 3 B_0 p_4}{p_6} + \frac{A_1}{B_0} \times \frac{1}{-24 \frac{B_0 p_4 A_1 p_6}{(4 A_1 p_6 w_2 - 3 B_0 p_4)(4 A_1 p_6 w_2 + 3 B_0 p_4)}\tanh\left(3/8 \frac{B_0 p_4}{A_1 p_6}(\eta + E)\right) - 32 \frac{w_2 p_6^2 A_1^2}{(4 A_1 p_6 w_2 - 3 B_0 p_4)(4 A_1 p_6 w_2 + 3 B_0 p_4)}}\right\}^{\frac{1}{2}} \tag{69}$$

when $\eta = x - \frac{T}{B_0}\sqrt{1/3 \frac{3 J B_0^2 + 4 A_1^2 p_6}{m}} t$.

The set of categories of solutions with $N = 2, M = 1$:

$N), B_j (0 \leq j \leq M)$, E, w_2 and w_1 are also the constants to be explored later.

Step 4. Balancing the non-linear ODE can obtain the values M and N.

Step 5. By solving the algebraic equations, we can get to the mentioned values.

6. APPLICATION EFM

In this section, the innovative soliton wave solutions for the model under investigation are constructed through an analytical approach outlined.

The set of categories of solutions with $N = 1, M = 0$:

6.1. Set I

$$\mu = 1/32 \frac{16 p_2 p_6 - 3 p_4^2}{p_6}, \quad A_0 = 0,$$

$$A_1 = -\frac{B_0(JT^2 - \lambda^2 m)}{T}\sqrt{-3(4 JT^2 p_6 - 4 \lambda^2 m p_6)^{-1}}, \tag{66}$$

$$w_1 = 0, \quad w_2 = \sqrt{-3(4 JT^2 p_6 - 4 \lambda^2 m p_6)^{-1}} p_4 T.$$

As a result (Group 3), the kink soliton solution is given by

$$u_1(x, t) = \left\{\frac{3(JT^2 - \lambda^2 m)(4 JT^2 p_6 - 4 \lambda^2 m p_6)^{-1} p_4}{\exp\left(\sqrt{-3(4 JT^2 p_6 - 4 \lambda^2 m p_6)^{-1}} p_4 T(x - \lambda t + E)\right) - 1}\right\}^{\frac{1}{2}}. \tag{67}$$

6.2. Set II

$$\lambda = \frac{T}{B_0}\sqrt{1/3 \frac{3 J B_0^2 + 4 A_1^2 p_6}{m}}, \quad \mu = 1/32 \frac{16 p_2 p_6 - 3 p_4^2}{p_6},$$

$$A_0 = 1/8 \frac{4 A_1 p_6 w_2 - 3 B_0 p_4}{p_6}, \tag{68}$$

$$w_1 = \frac{16 A_1^2 p_6^2 w_2^2 - 9 B_0^2 p_4^2}{64 A_1^2 p_6^2}, \quad w_2^2 - 4w_1 = \frac{9 B_0^2 p_4^2}{16 A_1^2 p_6^2}.$$

As a result by (Group 1), the soliton solution is concluded by

6.3. Set I

$$\lambda = \frac{T}{w_2}\sqrt{1/4 \frac{4 J p_6 w_2^2 + 3 p_4^2}{m p_6}}, \quad \mu = \frac{16 p_2 p_6 - 3 p_4^2}{32 p_6},$$

$$A_0 = A_2 = w_1 = 0, \quad B_0 = \frac{w_2(4 A_1 p_6 + 3 B_1 p_4)}{3 p_4}, \tag{70}$$

$$u_1(x,t) = \left\{ 3\, \frac{A_1 \mathrm{e}^{-\phi(\eta)} p_4}{3\, B_1 \mathrm{e}^{-\phi(\eta)} p_4 + 4\, A_1 p_6 w_2 + 3\, B_1 p_4 w_2} \right\}^{\frac{1}{2}},$$

$$\eta = x - \frac{T}{w_2} \sqrt{1/4\, \frac{4\, J p_6 w_2{}^2 + 3\, p_4{}^2}{m p_6}} t. \tag{71}$$

6.4. Set II

$$\lambda = \frac{T}{B_1} \sqrt{\frac{3\, J B_1{}^2 w_1 + 4\, A_1{}^2 p_6 + 3\, A_1 B_1 p_4}{3 m w_1}},\ \mu = \frac{16\, p_2 p_6 - 3\, p_4{}^2}{32 p_6},\ A_0 =$$

$$\sqrt{\frac{4\, A_1{}^2 p_6 w_1 + 3\, A_1 B_1 p_4 w_1}{4 p_6}}, \tag{72}$$

$$A_2 = B_0 = 0,\quad w_2 =$$

$$1/2\, \frac{(8\, A_1 p_6 + 3\, B_1 p_4) w_1}{p_6}\, \frac{1}{\sqrt{\frac{A_1 w_1 (4\, A_1 p_6 + 3\, B_1 p_4)}{p_6}}},$$

$$u_2(x,t) =$$

$$\left\{ \frac{\mathrm{e}^{\phi(\eta)}}{B_1} \left(\sqrt{1/4\, \frac{4\, A_1{}^2 p_6 w_1 + 3\, A_1 B_1 p_4 w_1}{p_6}} + A_1 \mathrm{e}^{-\phi(\eta)} \right) \right\}^{\frac{1}{2}}, \tag{73}$$

where $\eta = x - \frac{T}{B_1} \sqrt{1/3\, \frac{3\, J B_1{}^2 w_1 + 4\, A_1{}^2 p_6 + 3\, A_1 B_1 p_4}{m w_1}} t.$

6.5. Set III

$$\lambda = \frac{T}{-A_1{}^2 w_1 + A_0{}^2} \sqrt{1/4\, \frac{4\, J p_6 \left(-A_1{}^2 w_1 + A_0{}^2\right)^2 + 3\, A_0{}^2 A_1{}^2 p_4{}^2}{m p_6}},$$

$$u_4(x,t) = \left\{ 1/2\, \frac{2\, A_1 \mathrm{e}^{-\phi\left(x - \frac{T}{B_0} \sqrt{1/3\, \frac{3\, J B_0{}^2 + A_1{}^2 p_6}{m}} t\right)} p_6 B_1 - 2\, A_1 B_0 p_6 - 3\, B_0 B_1 p_4}{p_6 B_1 \left(B_0 + B_1 \mathrm{e}^{-\phi\left(x - \frac{T}{B_0} \sqrt{1/3\, \frac{3\, J B_0{}^2 + A_1{}^2 p_6}{m}} t\right)} \right)} \right\}^{\frac{1}{2}}. \tag{77}$$

6.7. Set V

$$\lambda =$$

$$\frac{T}{B_0 (2\, A_1 p_6 + B_1 p_4)} \sqrt{\frac{12\, J B_0{}^2 (2\, A_1 p_6 + B_1 p_4)^2 + A_1{}^2 p_6 (4\, A_1 p_6 + 3\, B_1 p_4)^2}{12 m}},$$

$$\mu = \frac{16\, p_2 p_6 - 3\, p_4{}^2}{32 p_6}, \tag{78}$$

$$A_0 = -\frac{B_0 (A_1 p_6 + B_1 p_4)}{B_1 p_6},\quad A_2 = 0,$$

where $\eta = x - \frac{T}{B_0 (2\, A_1 p_6 + B_1 p_4)} \sqrt{1/12\, \frac{12\, J B_0{}^2 (2\, A_1 p_6 + B_1 p_4)^2 + A_1{}^2 p_6 (4\, A_1 p_6 + 3\, B_1 p_4)^2}{m}} t.$

$$\mu = 1/32\, \frac{16\, p_2 p_6 - 3\, p_4{}^2}{p_6},\ A_2 = 0, \tag{74}$$

$$B_0 = -1/3\, \frac{-A_1 w_1 (4\, A_1 p_6 + 3\, B_1 p_4) + 4\, A_0{}^2 p_6}{A_0 p_4},\quad w_2 = \frac{A_1{}^2 w_1 + A_0{}^2}{A_0 A_1},$$

$$w_2^2 - 4 w_1 = \frac{\left(-A_1{}^2 w_1 + A_0{}^2\right)^2}{A_0{}^2 A_1{}^2},$$

$$u_3(x,t) =$$

$$\left\{ 3\, \frac{(A_0 + A_1 \mathrm{e}^{-\phi(\eta)}) A_0 p_4}{3\, B_1 \mathrm{e}^{-\phi(\eta)} A_0 p_4 + 4\, A_1{}^2 p_6 w_1 + 3\, A_1 B_1 p_4 w_1 - 4\, A_0{}^2 p_6} \right\}^{\frac{1}{2}}, \tag{75}$$

where

$$\eta = x - \frac{T}{-A_1{}^2 w_1 + A_0{}^2} \sqrt{1/4\, \frac{4\, J p_6 \left(-A_1{}^2 w_1 + A_0{}^2\right)^2 + 3\, A_0{}^2 A_1{}^2 p_4{}^2}{m p_6}} t.$$

6.6. Set IV

$$\lambda = \frac{T}{B_0} \sqrt{1/3\, \frac{3\, J B_0{}^2 + A_1{}^2 p_6}{m}},$$

$$\mu = 1/32\, \frac{16\, p_2 p_6 - 3\, p_4{}^2}{p_6},$$

$$A_0 = -1/2\, \frac{B_0 (2\, A_1 p_6 + 3\, B_1 p_4)}{p_6 B_1},\ A_2 = 0, \tag{76}$$

$$w_1 = 1/2\, \frac{B_0{}^2 (2\, A_1 p_6 + 3\, B_1 p_4)}{p_6 B_1{}^2 A_1},\quad w_2 =$$

$$-1/2\, \frac{(4\, A_1 p_6 + 3\, B_1 p_4) B_0}{A_1 p_6 B_1},\quad w_2^2 - 4 w_1 = 9/4\, \frac{B_0{}^2 p_4{}^2}{A_1{}^2 p_6{}^2},$$

$$w_1 = \frac{B_0{}^2 (A_1 p_6 + B_1 p_4)(4\, A_1 p_6 + B_1 p_4)}{p_6 B_1{}^2 A_1 (4\, A_1 p_6 + 3\, B_1 p_4)},$$

$$w_2 = -\frac{(8\, A_1{}^2 p_6{}^2 + 8\, A_1 B_1 p_4 p_6 + 3\, B_1{}^2 p_4{}^2) B_0}{p_6 B_1 A_1 (4\, A_1 p_6 + 3\, B_1 p_4)},\ w_2^2 - 4 w_1 =$$

$$9\, \frac{B_0{}^2 p_4{}^2 (2\, A_1 p_6 + B_1 p_4)^2}{p_6{}^2 A_1{}^2 (4\, A_1 p_6 + 3\, B_1 p_4)^2},$$

$$u_5(x,t) = \left\{ \frac{A_1 \mathrm{e}^{-\phi(\eta)} p_6 B_1 - A_1 B_0 p_6 - B_0 B_1 p_4}{p_6 B_1 \left(B_0 + B_1 \mathrm{e}^{-\phi(\eta)} \right)} \right\}^{\frac{1}{2}}, \tag{79}$$

Jalil Manafian, Walla Rahim Juadih, Amitha Manmohan Rao, Baharak Eslami, Natavan Allahverdiyeva, Parvin Mustafayeva
New Solitary Waves for Thin-Film Ferroelectric Material Equation Arising in Dielectric Materials

DOI 10.2478/ama-2024-0041

6.8. Set VI

$$\lambda = \frac{T}{B_0(8A_1p_6+3B_1p_4)}\sqrt{\frac{3JB_0{}^2(8A_1p_6+3B_1p_4)^2+4A_1{}^2p_6(4A_1p_6+3B_1p_4)^2}{3m}}, \quad \mu = \frac{16p_2p_6-3p_4{}^2}{32p_6}, \tag{80}$$

$$A_0 = -1/4 \frac{B_0(4\,A_1p_6+3\,B_1p_4)}{B_1p_6}, \quad A_2 = 0, \quad w_1 = \frac{B_0{}^2}{B_1{}^2},$$

$$\eta = x - \frac{T}{B_0(8\,A_1p_6+3\,B_1p_4)}\sqrt{1/3\,\frac{3\,JB_0{}^2(8\,A_1p_6+3\,B_1p_4)^2+4\,A_1{}^2p_6(4\,A_1p_6+3\,B_1p_4)^2}{m}}\,t.$$

$$w_2 = -1/4\,\frac{B_0(32\,A_1{}^2p_6{}^2+24\,A_1B_1p_4p_6+9\,B_1{}^2p_4{}^2)}{B_1p_6A_1(4\,A_1p_6+3\,B_1p_4)}, \quad w_2{}^2 - 4w_1 = \frac{9\,B_0{}^2p_4{}^2(8\,A_1p_6+3\,B_1p_4)^2}{16\,p_6{}^2A_1{}^2(4\,A_1p_6+3\,B_1p_4)^2},$$

$$u_6(x,t) = \left\{1/4\,\frac{4\,A_1e^{-\phi(\eta)}p_6B_1-4\,A_1B_0p_6-3\,B_0B_1p_4}{p_6B_1(B_0+B_1e^{-\phi(\eta)})}\right\}^{\frac{1}{2}}, \tag{81}$$

where

6.9. Set VII

$$\lambda = \frac{T}{B_1}\sqrt{1/3\,\frac{3\,JB_1{}^2+4\,A_2{}^2p_6}{m}}, \quad \mu = 1/32\,\frac{16\,p_2p_6-3\,p_4{}^2}{p_6},$$

$$A_0 = \frac{B_0(A_1B_1-A_2B_0)}{B_1{}^2}, \tag{82}$$

$$w_1 = 1/4\,\frac{(A_1B_1-A_2B_0)(4\,A_1B_1p_6-4\,A_2B_0p_6+3\,B_1{}^2p_4)}{A_2{}^2B_1{}^2p_6},$$

As a result (Group 1), the soliton solution is given by

$$u_7(x,t) = \left\{\frac{\left(-\frac{3}{2S}\frac{B_1{}^3p_4A_2}{(3\,B_1{}^2p_4+4\,p_6S)}\tanh\left(\frac{3B_1p_4}{8A_2p_6}(\eta+E)\right)-\frac{(3\,B_1{}^2p_4+8\,p_6S)A_2B_1}{2S(3\,B_1{}^2p_4+4\,p_6S)}\right)^{-1}A_2B_1+A_1B_1-A_2B_0}{B_1{}^2}\right\}^{\frac{1}{2}}, \tag{84}$$

where $\eta = x - \frac{T}{B_1}\sqrt{\frac{3\,JB_1{}^2+4\,A_2{}^2p_6}{3m}}\,t$ and $S = A_1B_1 - A_2B_0$.

$$w_2 = 1/4\,\frac{8\,p_6(A_1B_1-A_2B_0)+3\,B_1{}^2p_4}{A_2B_1p_6}, \quad w_2{}^2 - 4w_1 = \frac{9\,B_1{}^2p_4{}^2}{16\,A_2{}^2p_6{}^2},$$

$$u_7(x,t) = \left\{\frac{e^{-\phi\left(x-\frac{T}{B_1}\sqrt{\frac{3\,JB_1{}^2+4\,A_2{}^2p_6}{3m}}t\right)}A_2B_1+A_1B_1-A_2B_0}{B_1{}^2}\right\}^{\frac{1}{2}}. \tag{83}$$

$$A_1 = 1/4\,\frac{8\,A_2B_0p_6-3\,B_1{}^2p_4}{B_1p_6}, \quad w_1 = \frac{A_0}{A_2}, \quad w_2 = 1/4\,\frac{8\,A_2B_0p_6-3\,B_1{}^2p_4}{A_2p_6B_1},$$

6.10. Set VIII

$$\lambda = \frac{T}{B_1}\sqrt{1/3\,\frac{3\,JB_1{}^2+4\,A_2{}^2p_6}{m}}, \quad \mu = \frac{64\,A_2p_6{}^2(A_0B_1{}^2-A_2B_0{}^2)+48\,B_1{}^2p_6(A_2B_0p_4+B_1{}^2p_2)-9\,B_1{}^4p_4{}^2}{96\,p_6B_1{}^4}, \tag{85}$$

$$u_8(x,t) = \left\{\frac{4A_2e^{-2\phi\left(\frac{T}{B_1}\sqrt{\frac{3\,JB_1{}^2+4\,A_2{}^2p_6}{3m}}\right)}p_6B_1+e^{-\phi\left(\frac{T}{B_1}\sqrt{\frac{3\,JB_1{}^2+4A_2{}^2p_6}{3m}}\right)}(8\,A_2B_0p_6-3\,B_1{}^2p_4)+4\,A_0p_6B_1}{4p_6B_1\left(B_0+B_1e^{-\phi\left(\frac{T}{B_1}\sqrt{\frac{3\,JB_1{}^2+4A_2{}^2p_6}{3m}}\right)}\right)}\right\}^{\frac{1}{2}}. \tag{86}$$

As a result (Groups 1 and 2), the soliton and periodic solutions are given, respectively by

$$u_8(x,t) = \left\{\frac{4\,A_2\Phi^2p_6B_1+\Phi(8\,A_2B_0p_6-3\,B_1{}^2p_4)+4\,A_0p_6B_1}{4p_6B_1(B_0+B_1\Phi)}\right\}^{\frac{1}{2}}, \tag{87}$$

$$\Phi = \left\{-\frac{\sqrt{S}}{8B_1p_6A_0}\tanh\left(\frac{\sqrt{S}}{8A_2p_6B_1}(\eta+E)\right)-\frac{8\,A_2B_0p_6-3\,B_1{}^2p_4}{8B_1p_6A_0}\right\}^{-1},$$

$$w_2{}^2 - 4w_1 = -\frac{1}{16}\frac{64\,A_2p_6{}^2(A_0B_1{}^2-A_2B_0{}^2)+48\,A_2B_0B_1{}^2p_4p_6-9\,B_1{}^4p_4{}^2}{A_2{}^2B_1{}^2p_6{}^2},$$

$S > 0,$

$$\Phi = \left\{\frac{\sqrt{-S}}{8B_1p_6A_0}\tan\left(\frac{\sqrt{-S}}{8A_2p_6B_1}(\eta+E)\right)-\frac{8\,A_2B_0p_6-3\,B_1{}^2p_4}{8B_1p_6A_0}\right\}^{-1},$$

$S < 0,$

where $S = \left(8\,A_2B_0p_6 - 3\,B_1{}^2p_4\right)^2 - 64\,A_0A_2B_1{}^2p_6{}^2$

and $\eta = x - \frac{T}{B_1}\sqrt{\frac{3\,JB_1{}^2+4\,A_2{}^2p_6}{3m}}\,t.$

6.11. Set IX

$$\mu = \frac{16\,p_6{}^2(A_1B_1-2\,A_2B_0)^2+8\,B_1{}^2p_4p_6(A_1B_1-2\,A_2B_0)+16\,B_1{}^4p_2p_6-3\,B_1{}^4p_4{}^2}{32p_6B_1{}^4}, \tag{88}$$

$$A_0 = 1/4\,\frac{A_1{}^2}{A_2}, \quad w_1 = 1/4\,\frac{A_1{}^2B_1{}^2p_6+4\,A_1A_2B_0B_1p_6-8\,A_2{}^2B_0{}^2p_6+3\,A_2B_0B_1{}^2p_4}{A_2{}^2B_1{}^2p_6},$$

$$\lambda = \frac{T}{B_1}\sqrt{1/3\,\frac{3\,JB_1{}^2+4\,A_2{}^2p_6}{m}}, \quad w_2 = \frac{8\,A_1B_1p_6-8\,A_2B_0p_6+3\,B_1{}^2p_4}{4A_2B_1p_6},$$

$$w_2^2 - 4w_1 = 3/16\,\frac{(4\,A_1B_1p_6-8\,A_2B_0p_6+3\,B_1{}^2p_4)(4\,A_1B_1p_6-8\,A_2B_0p_6+B_1{}^2p_4)}{A_2{}^2B_1{}^2p_6{}^2},$$

$$u_9(x,t) = \left\{ 1/4\,\frac{4\,A_2{}^2 e^{-2\,\phi\left(x-\frac{T}{B_1}\sqrt{\frac{3\,JB_1{}^2+4\,A_2{}^2p_6}{3m}}t\right)}+4\,A_1 e^{-\phi\left(x-\frac{T}{B_1}\sqrt{\frac{3\,JB_1{}^2+4\,A_2{}^2p_6}{3m}}t\right)}A_2+A_1{}^2}{A_2\left(B_0+B_1 e^{-\phi\left(x-\frac{T}{B_1}\sqrt{\frac{3\,JB_1{}^2+4\,A_2{}^2p_6}{3m}}t\right)}\right)} \right\}^{\frac{1}{2}}. \tag{89}$$

As a result (Groups 1 and 2), the soliton and periodic solutions are given, respectively, by

$$u_9(x,t) = \left\{ 1/4\,\frac{4\,A_2{}^2\Phi^2+4\,A_1\Phi A_2+A_1{}^2}{A_2(B_0+B_1\Phi)} \right\}^{\frac{1}{2}}, \tag{90}$$

$$\Phi = \left\{ -1/2\,\frac{\sqrt{3}\sqrt{S}A_2B_1}{K}\tanh\left(1/8\,\frac{\sqrt{3}\sqrt{S}}{A_2B_1p_6}(\eta+E)\right) - \frac{(8\,p_6(A_1B_1-A_2B_0)+3\,B_1{}^2p_4)A_2B_1}{2K} \right\}^{-1}, \quad S>0,$$

$$\Phi = \left\{ 1/2\,\frac{\sqrt{3}\sqrt{-S}A_2B_1}{K}\tan\left(1/8\,\frac{\sqrt{3}\sqrt{-S}}{A_2B_1p_6}(\eta+E)\right) - \frac{(8\,p_6(A_1B_1-A_2B_0)+3\,B_1{}^2p_4)A_2B_1}{2K} \right\}^{-1}, \quad S<0,$$

where

$$\eta = x - \frac{T}{B_1}\sqrt{\frac{3\,JB_1{}^2+4\,A_2{}^2p_6}{3m}}t, \quad K = A_1{}^2B_1{}^2p_6+4\,A_1A_2B_0B_1p_6-8\,A_2{}^2B_0{}^2p_6+3\,A_2B_0B_1{}^2p_4,$$

$$S = (4\,A_1B_1p_6-8\,A_2B_0p_6+3\,B_1{}^2p_4)(4\,A_1B_1p_6-8\,A_2B_0p_6+B_1{}^2p_4).$$

6.12. Set X

$$\lambda = \frac{T}{B_1}\sqrt{\frac{3\,JB_1{}^2+4\,A_2{}^2p_6}{3m}}, \quad \mu = \frac{16\,p_2p_6-3\,p_4{}^2}{32p_6}, \quad A_0 = \frac{32\,A_2B_0p_6(8\,A_2B_0p_6-3\,B_1{}^2p_4)+9\,B_1{}^4p_4{}^2}{256\,B_1{}^2p_6{}^2A_2}, \tag{91}$$

$$A_1 = \frac{16\,A_2B_0p_6-3\,B_1{}^2p_4}{8B_1p_6}, \quad w_1 = \frac{256\,A_2{}^2B_0{}^2p_6{}^2-9\,B_1{}^4p_4{}^2}{256\,A_2{}^2B_1{}^2p_6{}^2}, \quad w_2 = 2\,\frac{B_0}{B_1}, \quad w_2^2-4w_1 = \frac{9\,B_1{}^2p_4{}^2}{64\,A_2{}^2p_6{}^2},$$

$$u_{10}(x,t) = \left\{ \frac{256A_2{}^2 e^{-2\phi(\eta)}B_1{}^2p_6{}^2+32e^{-\phi(\eta)}A_2B_1p_6\Upsilon_1+32A_2B_0p_6(8A_2B_0p_6-3B_1{}^2p_4)+9B_1{}^4p_4{}^2}{256B_1{}^2p_6{}^2A_2(B_0+B_1 e^{-\phi(\eta)})} \right\}^{\frac{1}{2}}, \tag{92}$$

where $\eta = x - \frac{T}{B_1}\sqrt{\frac{3\,JB_1{}^2+4\,A_2{}^2p_6}{3m}}\,t$. As a result (Group 1), the soliton solution is given by

$$u_{10}(x,t) = \left\{ \frac{256A_2{}^2\Phi^2B_1{}^2p_6{}^2+32\Phi A_2B_1p_6\Upsilon_1+32A_2B_0p_6(8A_2B_0p_6-3B_1{}^2p_4)+9B_1{}^4p_4{}^2}{256B_1{}^2p_6{}^2A_2(B_0+B_1\Phi)} \right\}^{\frac{1}{2}}, \tag{93}$$

$$\Upsilon_1 = (16A_2B_0p_6-3B_1{}^2p_4),$$

$$\Phi = \left\{ -\frac{3}{8}\frac{B_1{}^3p_4A_2p_6}{2\,A_2{}^2B_0{}^2p_6{}^2-\frac{9\,B_1{}^4p_4{}^2}{128}}\tanh\left[\frac{3}{16}\frac{B_1p_4}{A_2p_6}\left(x-\frac{T}{B_1}\sqrt{\frac{3\,JB_1{}^2+4A_2{}^2p_6}{3m}}t\right)\right] - \frac{2\,B_0B_1A_2{}^2p_6{}^2}{2\,A_2{}^2B_0{}^2p_6{}^2-\frac{9\,B_1{}^4p_4{}^2}{128}} \right\}^{-1}.$$

6.13. Set XI

$$\lambda = \frac{T}{B_1}\sqrt{\frac{3\,JB_1{}^2+4\,A_2{}^2p_6}{3m}}, \quad \mu = \frac{64\,p_2p_6-15\,p_4{}^2}{128p_6}, \quad A_0 = \frac{32\,A_2B_0p_6\left(8\,A_2B_0p_6-3\,B_1{}^2p_4\right)+9\,B_1{}^4p_4{}^2}{256\,B_1{}^2p_6{}^2A_2}, \tag{94}$$

$$A_1 = 1/8\,\frac{16\,A_2B_0p_6-3\,B_1{}^2p_4}{B_1p_6}, \quad w_1 = \frac{256\,A_2{}^2B_0{}^2p_6{}^2+9\,B_1{}^4p_4{}^2}{256\,A_2{}^2B_1{}^2p_6{}^2}, \quad w_2 = 2\,\frac{B_0}{B_1}, \quad w_2^2-4w_1 = -\frac{9\,B_1{}^2p_4{}^2}{64\,A_2{}^2p_6{}^2},$$

$$u_{11}(x,t) = \left\{\frac{256A_2{}^2e^{-2\phi(\eta)}B_1{}^2p_6{}^2+32e^{-\phi(\eta)}A_2B_1p_6\Upsilon_1+32A_2B_0p_6\left(8A_2B_0p_6-3B_1{}^2p_4\right)+9B_1{}^4p_4^2}{256B_1{}^2p_6{}^2A_2\left(B_0+B_1e^{-\phi(\eta)}\right)}\right\}^{\frac{1}{2}}, \tag{95}$$

where $\eta = x - \frac{T}{B_1}\sqrt{\frac{3\,JB_1{}^2+4\,A_2{}^2p_6}{3m}}\,t$. As a result (Group 2), the periodic solution is given by

$$u_{11}(x,t) = \left\{\frac{256A_2{}^2\Phi^2B_1{}^2p_6{}^2+32\Phi A_2B_1p_6\Upsilon_1+32A_2B_0p_6\left(8A_2B_0p_6-3B_1{}^2p_4\right)+9B_1{}^4p_4^2}{256B_1{}^2p_6{}^2A_2\left(B_0+B_1\Phi\right)}\right\}^{\frac{1}{2}}, \tag{96}$$

$$\Upsilon_1 = \left(16A_2B_0p_6-3B_1{}^2p_4\right),$$

$$\Phi = \left\{\frac{3}{8}\,\frac{B_1{}^3p_4A_2p_6}{2\,A_2{}^2B_0{}^2p_6{}^2+\frac{9\,B_1{}^4p_4^2}{128}}\tan\left[\frac{3}{16}\,\frac{B_1p_4}{A_2p_6}\left(x-\frac{T}{B_1}\sqrt{\frac{3\,JB_1{}^2+4A_2{}^2p_6}{3m}}\,t\right)\right] - \frac{2\,B_0B_1A_2{}^2p_6{}^2}{2\,A_2{}^2B_0{}^2p_6{}^2+\frac{9\,B_1{}^4p_4^2}{128}}\right\}^{-1}.$$

6.14. The graphical discussion and physical significance

By selecting the appropriate values for the parameter, we were able to generate the desired types of solutions that indicate wave discrepancy. The analytical solutions are coded in maple and the parametric and sensitivity analysis are carried out using the codes. The parametric results are presented in Figs. 1–4. The present results from the simulations show that through an inherent property of auxiliary parameters for the adjustment and control of region and rate of convergence of approximate series solutions, the MEJM and rational $\exp(-\phi(\eta))$-expansion method have proven to be very efficient and capable techniques in handling non-linear engineering problems in wider ranges of parameters. The importance of this study lies in the actuality that it can serve as a base for the experimental work that we want to undertake on the plasma physics and crystal lattice theory.

7. CONCLUSION

On the basis of the constructed auxiliary functions, the MEJM, the rational $\exp(-\phi(\eta))$-expansion method and the solitary wave solutions by utilising TFFME were inspected. The mentioned equation is non-integrable. The impact of wave motion in plasma on the physical parameters including speed and amplitudes of solitary waves has been focussed. Then, the general form rational solutions to TFFME containing soliton, kink soliton, singular soliton and periodic wave solutions were observed. We found plenty of exact solutions for two cases. The dynamical behaviour of results was investigated via graphical illustrations by using considered methods. Moreover, various important remarks about the physical meanings of solutions were presented. From these results, it may be seen that the MEJM and EEM are the power tools to solve such non-linear partial models arising in applied and engineering sciences. In the future, we can further study its soliton solutions, rogue wave solutions, solitary waves and symmetry, etc.

REFERENCES

1. Zhou X, Ke Q, Tang S, Luo J, Lu Z. Ultraviolet photodetectors based on ferroelectric depolarization field. Journal of Energy Chemistry. 2023;77:487-98.
2. Xiao X, Xu X, Lu Z, Zhao J, Liu R, Ye Y, Tang R, Liao WQ, Xiong RG, Zou, G. In-situ organic-inorganic ferroelectric layer growth for efficient perovskite solar cells with high photovoltage. Nano Energy. 2023;107:108114.
3. Umoh GV, Holguín-Momaca JT, Talamantes RP, Rojas-George G, Herrera-Pérez G, Antón RL, Aguirre-Tostado FS, Auciello O, Olive-Méndez, SF, Hurtado-Macias A. Influence of ex-situ thermal treatment on the chemical states, microstructure and ferroelectrics properties of polycrystalline BiMnO_3-δ thin films. Thin Solid Films. 2022 August 31;756:139362.
4. Tarnaoui M, Zaim N, Zaim A, Kerouad M. Investigation of the thermal and ferroelectric properties of a Spin-1 Ising thin film: Insight from path integral Monte Carlo. Materials Science and Engineering B. 2023;288:116204.
5. Yang ST, Li XY, Yu TL, Wang J, Fang H, Nie F, He B, Zhao L, Lü WM, Yan SS, Nogaret A, Liu G, Zheng LM. High-performance neuromorphic computing based on ferroelectric synapses with excellent conductance linearity and symmetry. Advanced Functional Materials. 2022;32:2202366.
6. Zhang J, Wang X, Zhou L, Liu G, Adroja DT, Silva I, Demmel F, Khalyavin D, Sannigrahi J, Nair HS, Duan L, Zhao J, Deng Z, Yu R, Shen X, Yu R, Zhao H, Zhao J, Long Y, Hu Z, Lin HJ, Chan TS, Chen CT, Wu W, Jin C. A ferrotoroidic candidate with well-separated spin chain. Advanced Materials. 2022;34:2106728.
7. Hubert MB, Justin M, Kudryashov NA, Betchewe G, Douvagai, Doka SY. Solitons in thin-film ferroelectric material. Physica Scripta. 2018;93:075201.
8. Zahran EH, Alizamini SMM, Shehata MSM, Rezazadeh H, Ahmad H. Study on abundant explicit wave solutions of the thin-film ferroelectric materials equation. Optical and Quantum Electronics. 2022;54(1):48.
9. Bekir A, Zahran EHM. Optical soliton solutions of the thin-flm ferroelectric materials equation according to the Painlevé approach. Optical and Quantum Electronics. 2021;53(2):118.
10. Gu Y, Zia SM, Isam M, Manafian J, Hajar A, Abotaleb M. Bilinear method and semi-inverse variational principle approach to the generalized (2+1)-dimensional shallow water wave equation. Results in Physics. 2023;45:106213.

11. Gu Y, Malmir S, Manafian J, Ilhan OA, Alizadeh A, Othman AJ (2022). Variety interaction between k-lump and k-kink solutions for the (3+1)-D Burger system by bilinear analysis. Results in Physics. 2022;43:106032.

12. Yao SW, Akram G, Sadaf M, Zainab I, Rezazadeh H, Inc M. Bright, dark, periodic and kink solitary wave solutions of evolutionary Zoomeron equation. Results in Physics. 2022 December;43:106117.

13. Manafian J, Lakestani M. N–lump and interaction solutions of localized waves to the (2+1)-dimensional variable-coefficient Caudrey–Dodd-Gibbon-Kotera-Sawada equation. Journal of Geometry and Physics. 2020;150:103598.

14. Manafian J, Mohammed SA, Alizadeh A, Baskonus HM, Gao W. Investigating lump and its interaction for the third–order evolution equation arising propagation of long waves over shallow water. European Journal of Mechanics-B/Fluids. 2020;84:289-301.

15. Rao X, Manafian J, Mahmoud KH, Hajar A, Mahdi AB, Zaidi M. (2022). The nonlinear vibration and dispersive wave systems with extended homoclinic breather wave solutions. Open Physics. 2022;20:795-821.

16. Alimirzaluo E, Nadjafikhah M, Manafian J. Some new exact solutions of (3+1)-dimensional Burgers system via Lie symmetry analysis. Advances in Differential Equations. 2021;2021(60):1-17.

17. Xiao Y, Fan E, Liu P. Inverse scattering transform for the coupled modified Korteweg–de Vries equation with nonzero boundary conditions. Journal of Mathematical Analysis and Applications.2021;504:125567.

18. Wu W, Manafian J, Ali KK, Karakoç SBG, Taqik AH, Mahmoud MA. Numerical and analytical results of the 1D BBM equation and 2D coupled BBM-system by finite element method. International Journal of Modern Physics B. 2022;36(28):2250201.

19. Pan Y, Manafian J, Zeynalli SM, Al-Obaidi R, Sivaraman R, Kadi A. N-Lump Solutions to a (3+1)-Dimensional Variable-Coefficient Generalized Nonlinear Wave Equation in a Liquid with Gas Bubbles. Qualitative Theory of Dynamical Systems. 2022;21:127.

20. Shen X, Manafian J, Jiang M, Ilhan OA, Shafikk SS, Zaidi M. Abundant wave solutions for generalized Hietarinta equation with Hirota's bilinear operator. Modern Physics Letters B. 2022;36(10):2250032.

21. Li R, Ilhan OA, Manafian J, Mahmoud KH, Abotaleb M, Kadi A. A Mathematical Study of the (3+1)-D Variable Coefficients Generalized Shallow Water Wave Equation with Its Application in the Interaction between the Lump and Soliton Solutions. Mathematics. 2022;10:3074.

22. Gu Y, Manafian J, Mahmoud MZ, Ghafel ST, Ilhan OA. (2021). New soliton waves and modulation instability analysis for a metamaterials model via the integration schemes. International Journal of Nonlinear Sciences and Numerical Simulation. 2021, https://doi.org/10.1515/ijnsns-2021-0443.

23. Singh S, Sakkaravarthi K, Murugesan K. Lump and soliton on certain spatially-varying backgrounds for an integrable (3+1) dimensional fifth-order nonlinear oceanic wave model. Chaos, Solitons and Fractals. 2023;167:113058.

24. Sherazi K, Sheikh N, Anjum M, Raza AG. Solar drying experimental research and mathematical modelling of wild mint and peach moisture content. Journal of Asian Scientific Research. 2023;13(2):94-107.

25. Ogunjiofor EI, Ayodele FO. Utilization of response surface methodology in optimization of locally sourced aggregates. Journal of Asian Scientific Research. 2023;13(1):54-67.

26. Rai A, Mohanty B, Agarwal S. Mathematical modeling and simulation of fluidized bed gasifier: Application to Indian coal. Iranian Journal of Chemistry and Chemical Engineering. 2023;42(1):269-85.

27. Ma H. (2021, January). Research on promotion of lower limb movement function recovery after stroke by using lower limb rehabilitation robot in combination with constant velocity muscle strength training. In 2021 7th International Symposium on Mechatronics and Industrial Informatics. 2021 January;70-73. IEEE

28. Faridi WA, Abu Bakar M, Myrzakulova Z, Myrzakulov R, Akgül A, El Din SM. The formation of solitary wave solutions and their propagation for Kuralay equation. Results in Physics. 2023;52:106774.

29. Faridi WA, Asghar U, Asjad MI, Zidan AM, ElDin SM. Explicit propagating electrostatic potential waves formation and dynamical assessment of generalized Kadomtsev-Petviashvili modified equal width-Burgers model with sensitivity and modulation instability gain spectrum visualization. Results in Physics. 2023 January;44:106167.

30. Faridi WA, Asjad MI, Jhangeer A, Yusuf A, Sulaiman TA. The weakly non-linear waves propagation for Kelvin-Helmholtz instability in the magnetohydrodynamics flow impelled by fractional theory. Optical and Quantum Electronics. 2023;55:172.

31. Faridi WA, Asjad MI, ElDin SM. Exact Fractional Solution by Nucci's Reduction Approach and New Analytical Propagating Optical Soliton Structures in Fiber-Optics. Fractal and Fractional. 2022 November 5;6(11):654.

32. Faridi WA, Abu Bakar M, Akgül A, El-Rahman MA, ElDin SM. Exact fractional soliton solutions of thin-film ferroelectric material equation by analytical approaches. Alexandria Engineering Journal. 2023 September 1;78:483-97.

33. Faridi WA, Asjad MI, Toseef M, Amjad T. Analysis of propagating wave structures of the cold bosonic atoms in a zig-zag optical lattice via comparison with two different analytical techniques. Optical and Quantum Electronics. 2022;54:773.

34. Chen D, Li Z. New optical soliton solutions of thin-film ferroelectric material equation via the complete discrimination system method. Results in Physics. 2023 July;50:106551.

35. Li Z, Peng C. Bifurcation, phase portrait and traveling wave solution of time-fractional thin-film ferroelectric material equation with beta fractional derivative. Physics Letters A. 2023 October 5;484:129080.

36. Safi Ullah M, Or-Roshid H, Ali MZ, Rezazadeh H. Kink and breather waves with and without singular solutions to the Zoomeron model. Results in Physics. 2023;49:106535.

37. Kallel W, Almusawa H, Alizamini SMM, Eslami M, Rezazadeh H, Osman MS. Optical soliton solutions for the coupled conformable Fokas-Lenells equation with spatio-temporal dispersion. Results in Physics. 2021;26:104388.

38. Vahidi J, Zabihi A, Rezazadeh H, Ansari R. New extended direct algebraic method for the resonant nonlinear SchrÖdinger equation with Kerr law nonlinearity. Optik. 2021 February ;227:169936.

39. Amir M, Haider JA, Ul Rahman J, Ashraf A. Solutions of the nonlinear evolution problems and their application. Acta Mechanica et Automatica. 2023;17(3):357-63.

40. Amir M, Haider JA, Ahmad S, Gul S, Ashraf A. Approximate solution of Painlevé equation I by natural decomposition method and Laplace decomposition method. Acta Mechanica et Automatica. 2023;17(3):417-22.

41. Manafian J. An optimal Galerkin-homotopy asymptotic method applied to the nonlinear second-order BVPs. Proceedings of the Institute of Mathematics and Mechanics. 2021;47(1):156-82.

42. Manafian J, Allahverdiyeva N. An analytical analysis to solve the fractional differential equations. Advanced Mathematical Models & Applications. 2021;6(2):128-61.

43. Li R, Bu Sinnah ZA, Shatouri ZM, Manafian J, Aghdaei MF, Kadi A. Different forms of optical soliton solutions to the Kudryashov's quintuple self-phase modulation with dual-form of generalized nonlocal nonlinearity. Results in Physics. 2023;46:106293.

44. Ali NH, Mohammed SA, Manafan J. New explicit soliton and other solutions of the Van der Waals model through the ESHGEEM and the IEEM. Journal of Modern Technology and Engineering. 2023;8(1):5-18.

45. Aldhabani MS, Nonlaopon K, Rezaei S, Bayones FS, Elagan SK, El-Marouf SAA. Abundant solitary wave solutions to a perturbed Schrodinger equation with Kerr law nonlinearity via a novel approach. Results in Physics. 2022 April;35:105385.

46. Aghazadeh MR, Asgari T, Shahi A, Farahm A. Designing strategy formulation processing model of governmental organizations based on network governance. Quarterly Journal of Public Organizations Management. 2016;4(1):2952.

New Solitary Waves for Thin-Film Ferroelectric Material Equation Arising in Dielectric Materials

47. Gholamiangonabadi D, Nakhodchi S, Jalalimanesh A, Shahi A. Customer churn prediction using a metaclassifier approach; A case study of Iranian banking industry. Proceedings of the International Conference on Industrial Engineering and Operations Management. 2019:364375.
48. Gaur M, Singh K. Lie group of transformations for time fractional Gardner equation. AIP Conf Proc. 2022;2357:090006.
49. Aneja M, Gaur M, Bose T, Gantayat PK, Bala R. Computer-Based Numerical Analysis of Bioconvective Heat and Mass Transfer Across a Nonlinear Stretching Sheet with Hybrid Nanofluids. In: Bhateja, V., Yang, XS., Ferreira, M.C., Sengar, S.S., Travieso-Gonzalez, C.M. (eds) Evolution in Computational Intelligence. FICTA 2023. Smart Innovation. Systems and Technologies. Vol 370. Springer. Singapore. https://doi.org/10.1007/978-981-99-6702-5 55.

Jalil Manafian: https://orcid.org/0000-0001-7201-6667

Walla Rahim Juadih: https://orcid.org/0009-0001-5112-7496

Amitha Manmohan Rao: https://orcid.org/0000-0003-2361-7778

Baharak Eslami: https://orcid.org/0000-0002-6723-1909

Natavan Allahverdiyeva: https://orcid.org/0000-0002-5600-2385

Parvin Mustafayeva: https://orcid.org/0000-0001-7412-3458

NORMAL TEMPERATURE MECHANICAL PROPERTIES OF 6082 ALUMINIUM ALLOY
AS A FUNCTION OF TEMPERING TEMPERATURE:
EXPERIMENTAL AND NUMERICAL APPROACH

Abdelmalek ELHADI */**, **Salah AMROUNE */****, **Amin HOUARI *****, **Madani KOUDER ******

*Mechanical Department, Faculty of Technology, University of Msila, Msila, Algeria
**Materials and Structural Mechanics Laboratory (LMMS). University of M'sila, M'sila, Algeria
***Faculty of Technology, M'hamed Bougara, University, Boumerdes 35000, Algeria
****Laboratoire de Matériaux, et Mécanique des Structures (LMMS), Université SBA. Sidi Bel Abesse, Algérie

abdelmalek.elhadi@univ-msila.dz, salah.amroune@univ-msila.dz, houari.latif2016@gmail.com, koumad10@yahoo.fr

Abstract: Our study involved a combination of practical experiments and numerical simulations using the Abaqus computational software. The main aim was to enhance our understanding of the mechanical characteristics exhibited by 6082 aluminium alloy when exposed to tensile forces. To achieve this, we produced 18 samples of standardized dimensions utilizing a parallel lathe. These samples then underwent a thermal treatment comprising a solution treatment, water quenching and various tempering procedures at different temperatures (280°C, 240°C, 200°C, 160°C and 120°C), resulting in a range of hardness levels. To obtain the experimental results, we conducted tensile tests on a specialized machine, which were subsequently supplemented with numerical analyses. By adopting this approach, we gained valuable insights into the behaviour of aluminium alloy 6082, specifically regarding its mechanical properties such as hardness, tensile strength, elongation and necking coefficient. This newfound knowledge holds potential significance in the realm of designing and optimizing aluminium structures that operate within high-temperature environments.

Key words: 6082 Aluminium alloy, quenching and tempering, hardness, tensile strength

1. INTRODUCTION

Aluminium alloys have been the subject of much research and scientific progress. Their importance in industry lies in their characteristic properties such as low densities. They are two to four times lower than that of steels. They are widely applied, especially in the aeronautical, automotive and shipbuilding industries (1). The mechanical properties of Al–Mg–Si alloys can be improved by including the additional elements or by applying appropriate heat treatments such as solution heat treatment, quenching and ageing treatment (2). 6082 Aluminium alloy (ISO name: AlSiMg07) has favourable mechanical properties, good corrosion resistance and good mechanical strength of the order of 320 MPa in the T6 state (3). The presence of silicon and copper reduces thermal expansion, while magnesium increases it (4, 5). This alloy is mainly used in the transport and structural engineering industry, such as in bridges, cranes, frameworks, transport aircraft and transport ships. The mechanical properties of this alloy such as hardness and breaking strength are improved by heat treatments (6-8).

Researchers have studied the effect of temperature on the mechanical properties of aluminium alloy in tensile tests on cylindrical specimens accompanied by numerical simulation. Terena et al (9) conducted experimental and numerical analyses on the drawing efficiency of quenching aluminium alloy 6082. Validation of the theoretical evaluation was performed by superimposing the graph of the triaxiality ratio of stress versus strain plastic and the failure envelope graph of the 60820 quenching aluminium alloy for a zero seam parameter. The numerical simulation gave the ratio of triaxiality of the stresses compared with the correlation of the plastic deformations. In order to define the constituent material model and the rupture envelope, experimental determinations and numerical simulations of plane stress and plane strain specimens were carried out (tensile, shear and compression tests for various samples). The simulation results were compared with experimental observations. Recent works have revealed important data on this matter, such as Clausen et al. (10), who targeted the dependence of fracture on strain rate, triaxiality and temperature, or Field et al. (11), who studied various materials submitted to high rate shock. Mean stress was proven to play an important role in the fracture of metals by Bao and Wierzbicki (12). Their numerical simulations with the cut-off value in fracture loci successfully captured the main features observed in tensile tests under hydrostatic pressure. Furthermore, Wierzbicki et al. (12) have shown the advantage of working with plane stress.

This work studies the influence of heat treatment on the mechanical properties of alloy 6082 before and after the treatment. To do this, the specimens are organized into 6 series of 3 samples each; therefore a total of 18 samples are tested. It should be noted that the experimental results obtained are processed and analysed by the Abaqus software in order to make a comparison between the experimental and numerical results.

Abdelmalek Elhadi, Salah Amroune, Amin Houari, Madani Kouder
Normal Temperature Mechanical Properties of 6082 Aluminium Alloy as a Function of Tempering Temperature: Experimental and Numerical Approach

DOI 10.2478/ama-2024-0042

2. MATERIALS AND METHODS

The tensile specimen material used is 6082 aluminium alloy. The results of the spectroscopic chemical analysis (wt%) of the alloy are shown in Tab. 1.

Tab. 1. The chemical composition of the material (wt%)

Mg	Si	Cu	Fe	Mn	Zn	Ni	Cr
1.20	0.85	0.71	0.78	0.56	0.40	0.17	0.07

The geometric shape and the dimensions (in mm) of the specimen chosen (according to the ISO 6892-1 standard) are represented in Fig. 1.
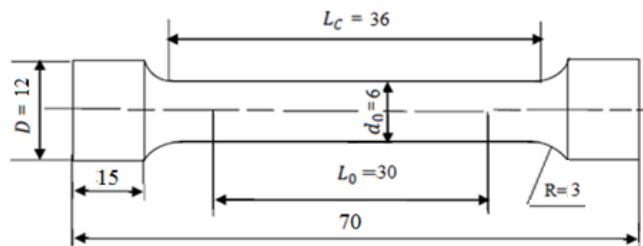

Fig. 1. Tensile specimen used

The average hardness of the material in the delivered state is 118 HV. The type of heat treatment applied to the specimens is the structural hardening treatment as shown in Fig. 2:

− Solid solution: Heating at a high temperature of 540°C for 30 min to put in the solid solution (a mixture of pure substances forming a homogeneous solid) the greatest possible number of soluble hardening elements.
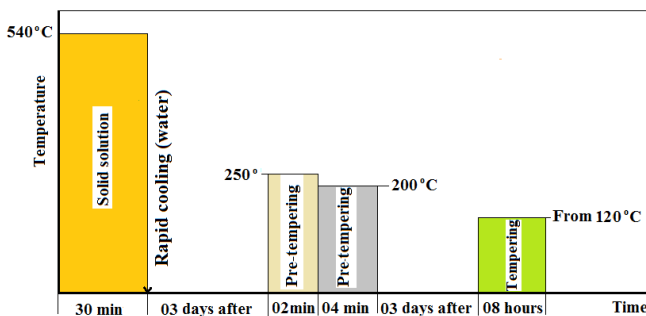− Rapid cooling: Water is often used as a quenching medium to keep the solid solution supersaturated (13).

Fig. 2. Structural hardening treatment of AL6082

Industrially, tempering cannot always be carried out immediately after quenching. Different effective remedies have been found, particularly in the case of A-SG alloys. This can be done for example as follows:

Pre-tempering in two stages: Pre-tempering for 2 min at 250°C + 4 min at 200°C (3 days after quenching). The short pre-tempering time allows for 1 week in which the time between quenching and tempering has minimal effect on the mechanical properties and prevents natural ageing.

The tempering time of AL6082 is between 8 and 10 h.

Tempering for 8 h: Applied 3 days after pre-tempering at 280°C, 240°C, 200°C, 160°C and 120°C.

Fig. 3 shows two specimens, one before heat treatment (Fig. 3a) and the other after solution treatment and quenching (Fig. 3b).
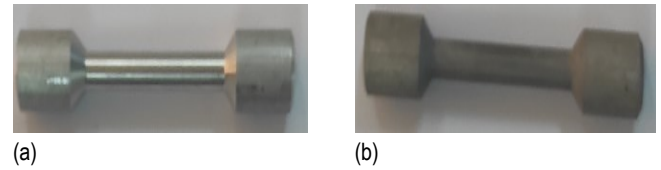

(a)   (b)

Fig. 3. Tensile specimens: (a) Before heat treatment; (b) After heat treatment

The average value of the hardness of the 6082 aluminium alloy used after quenching is 103 HV (decrease in hardness compared with the untreated material [118 HV]).

It is a hydraulically driven vertical column test apparatus (Fig. 4). The maximum pulling force can reach 50 KN in both directions. The strain rate used in tensile tests 0.2 mm/min.

Fig. 4. Tensile machine WP 310

3. CONSTITUTIVE MODEL OF THE MATERIAL AL6082-T6

The aluminium specimens were modelled as an elastoplastic material by Johnson–Cook plasticity and the damage criterion introduced in the ABAQUS calculation code was used to develop numerical simulations in order to study the structural response of the metal. This criterion provides a satisfactory description of the behaviour of metal and alloys, since it takes into account large strains, high strain rates and temperature-dependent viscoplasticity. For this purpose, the temperature factor in the aluminium alloy was taken into account. We therefore used the Johnson–Cook material model to determine the equivalent stress in the following form:

$$\sigma_{JC} = [A + B(\bar{\varepsilon}^p)^n]\left[1 + C.\ln\left(\overline{\dot{\varepsilon}^p}/\dot{\bar{\varepsilon}}_0^p\right)\right]\left[1 - \left(\frac{\theta_w - \theta_0}{\theta_m - \theta_0}\right)^m\right] \quad (1)$$

Where A, B, C, m and n are five materials constants. A is the yield strength, B is the strength coefficient, C is the strain rate coefficient, n is the strain hardening coefficient and m is the thermal softening coefficient. σ_{JC} is the flow stress, ε is the plastic strain, $\bar{\dot{\varepsilon}}^p$ is the plastic strain rate, $\dot{\bar{\varepsilon}}_0^p$ is the reference plastic strain rate, θ_0 is the temperature of the workpiece material, θ_w is the reference temperature, and θ_m is the melting temperature of

the materials. Many criteria have been developed to predict the damage of metallic materials in the case of loadings. They rely on either of the maximum strain conditions for damage initiation. Damage in the Johnson–Cook material model is predicted using the following cumulative damage law:

$$W = \sum \frac{\Delta \bar{\varepsilon}^p}{\bar{\varepsilon}_f^p} \qquad (2)$$

where $\Delta \bar{\varepsilon}^p$ is the accumulated increment of equivalent plastic strain during an integration step, W is the damage parameter for fracture initiation when it is equal to 1 $\bar{\varepsilon}_f^p$ as the deformation equivalent to rupture, deduced as follows:

$$\bar{\varepsilon}_f^p = \left[D_1 + D_2 \exp\left(D_3 \frac{P}{\sigma_{JC}} \right) \right] \left[1 + D_4 . \ln\left(\overline{\dot{\varepsilon}^p} / \dot{\varepsilon}_0^p \right) \right] \left[1 - D_5 \left(\frac{T_w - T_0}{T_m - T_0} \right)^m \right] \qquad (3)$$

$\frac{P}{\sigma_{JC}}$ is the mean stress normalized by the equivalent stress, and parameters D_1, D_2, D_3, D_4 and D_5 are constants. The evolutions of the damage are defined by the energy condition to create new free surfaces (Eq. 2). The choice of the energy approach is often governed by the size of the finite elements.

$$G_f = \int_{\bar{\varepsilon}_0^{pl}}^{\bar{\varepsilon}_f^{pl}} L \sigma_y d\bar{\varepsilon}^{pl} \qquad (4)$$

Thus, following the initiation of the damage, the variable of damage increases according to the following equation:

$$D = \frac{L d\bar{\varepsilon}^{pl}}{\bar{u}_f^{pl}} \qquad (5)$$

$\frac{P}{\sigma_{JC}}$ the plastic displacement equivalent to fracture is calculated as in the following equation:

$$\bar{u}_f^{pl} = \frac{2G_f}{\sigma_{y0}} \qquad (6)$$

where σ_{y0} is the elastic limit of the material, G_f is the fracture energy and L is the characteristic of the finite element.

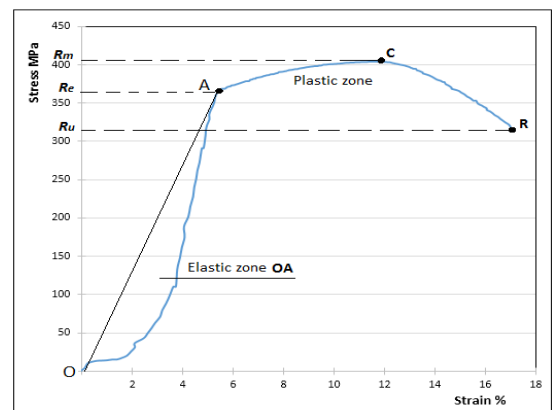**Tab. 2.** Johnson–Cook constitutive model constants for AL 6082-T6

Al6082-T6	Value
A (MPa)	360
B (MPa)	249
N	0.62
C	0.008
Tm (°C)	582
T (°C)	25
M	0.6122
D1 [17]	0.0164
D2 [17]	2.245
D3 [17]	−2.798
D4	−0.284
D5	−2.342

Tab. 2. Johnson–Cook constant and static tensile strength for AL 6082-T6 of the strain rate used in tensile tests 0.2 mm/min.
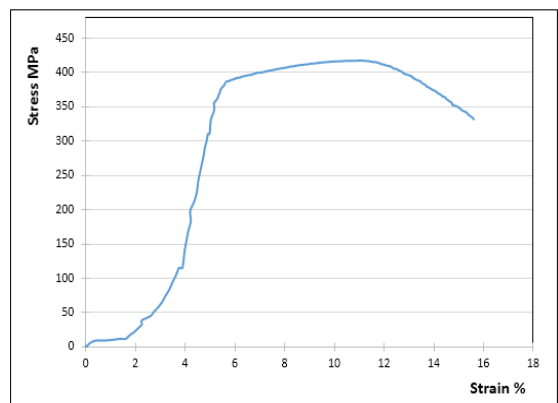
4. RESULTS
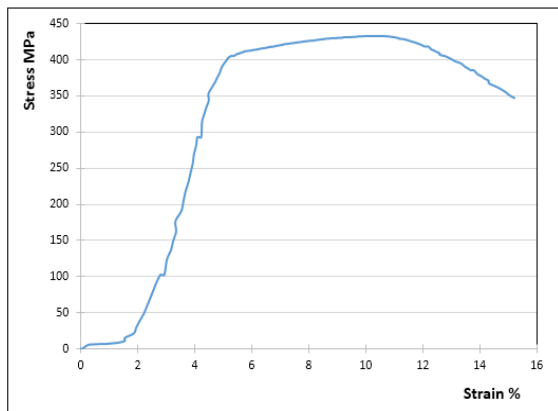
4.1. Tensile behaviour of material

The specimens used in a tensile test are prepared according to international standards. The stress–strain curve obtained from the tensile test of the material used is shown in Fig. 5. All the curves have the same appearance and reflect the mechanical behaviour of the material in tension (14, 15). We take as an example Fig. 5a, where the stress–strain curve starts with the elastic strain or the material, following Hooke's law which expresses the linearity between stress and strain. With increase in the tensile load, the material enters the plastic zone (permanent deformation) up to a maximum load which is expressed in relation to the initial section of the specimen tensile strength Rm. After that, the deformation is concentrated in one area, where there is necking (throttling) and then the rupture of the specimen.
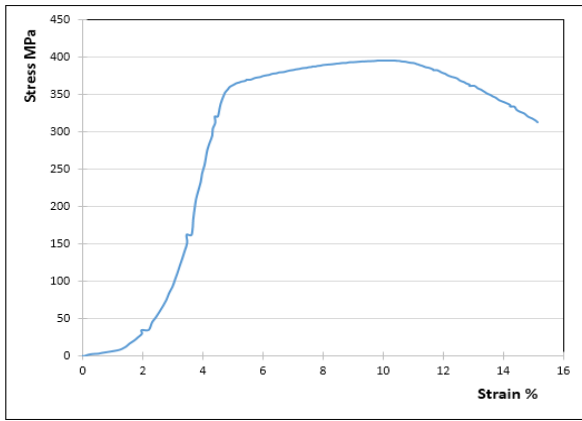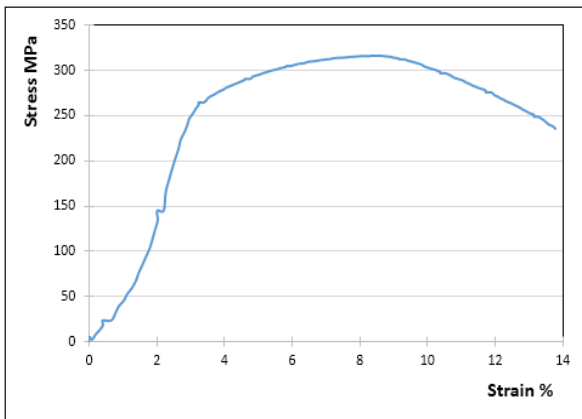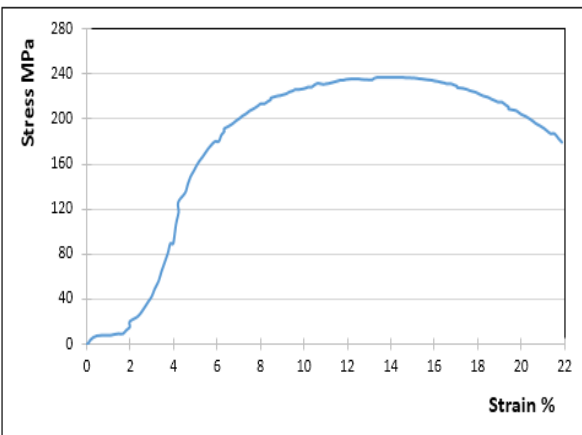

(a)

(b)

(c)

Abdelmalek Elhadi, Salah Amroune, Amin Houari, Madani Kouder
Normal Temperature Mechanical Properties of 6082 Aluminium Alloy as a Function of Tempering Temperature: Experimental and Numerical Approach

DOI 10.2478/ama-2024-0042

(d)

(e)

(f)

Fig. 5. Stress–strain curves of 6082 Aluminium alloy: (a) Untreated; Tempering temperature: (b) 120°C; (c) 160°C; (d) 200°C; (e) 240°C; (f) 280°C

4.2. Variation of hardness as a function of tempering temperature

The evolution of the average hardness HV as a function of tempering temperature is shown in Fig. 6. This figure shows a decrease in hardness values with increasing tempering temperature.
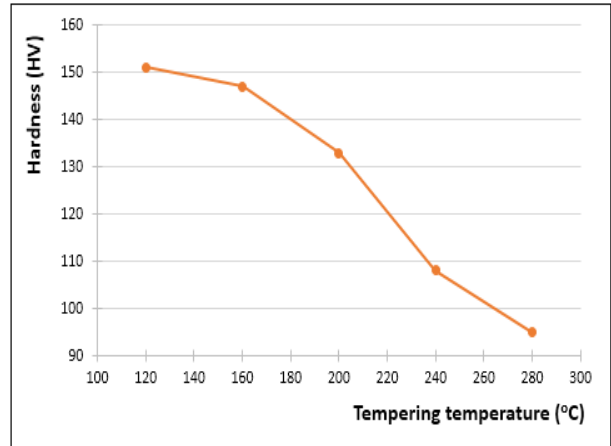

Fig. 6. Evolution of the hardness as a function of tempering temperature

4.3. Evolution of the tensile strength Rm as a function of tempering temperature

The evolution of the tensile strength as a function of the tempering temperature is shown in Fig. 7. It can be seen that the tensile strength for the material is maximum at the tempering temperature of 160°C (422 MPa) and minimum at the tempering temperature of 280°C (214 MPa).
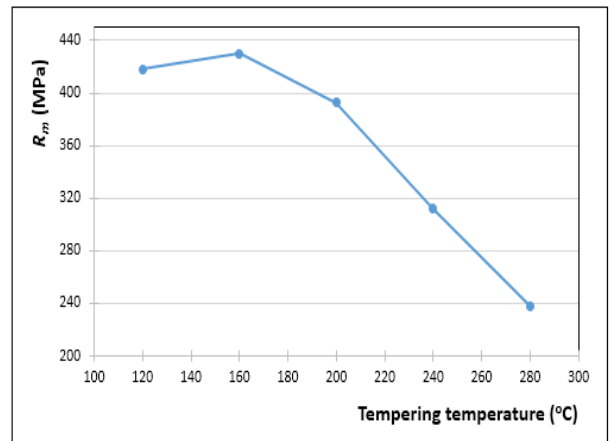
Fig. 7. Evolution in tensile strength as a function of tempering treatment

4.4. Variation in elongation A% as a function of tempering temperature

The percent elongation of the specimen (*A%*) is the ultimate elongation. It is determined by the formula:

$$A\% = 100\,\frac{Lu - Lo}{Lo} \tag{7}$$

with L_o: initial length and L_u: length after rupture.

Fig. 8 shows that the elongation values of the material are close to the tempering temperature of 120–280°C. On the contrary, the elongation is maximum at the tempering temperature of 280°C (7.92%).
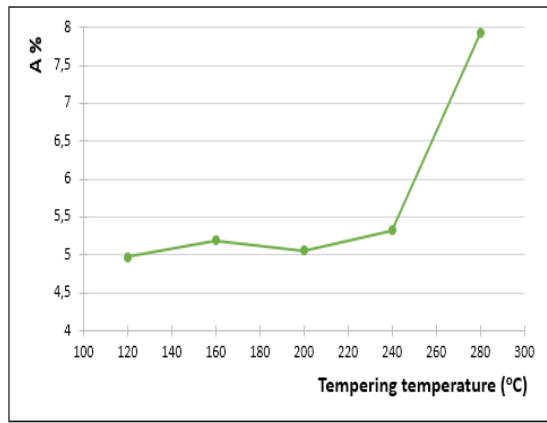
Fig. 8. Evolution of elongation as a function of tempering temperature

4.5. Variation of the coefficient of necking Z% as a function of tempering temperature

The percent necking coefficient of the specimen (Z%) is determined by the formula:

$$Z\% = 100 \frac{So - Su}{So} \tag{8}$$

with S_o: initial section and S_u: section after rupture.

Fig. 9 presents the results obtained from the coefficient of necking Z% as a function of tempering temperature. The increase in the tempering temperature leads to an increase in the coefficient of necking.
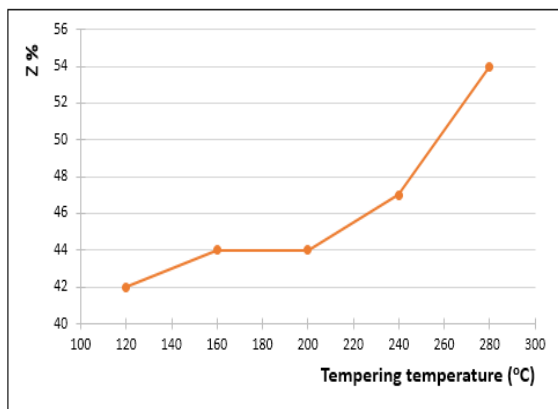

Fig. 9. Evolution of the coefficient of necking as a function of tempering temperature

5. DISCUSSION

For the alloy quenched and tempered to 280°C, the hardness decreased (95 HV), when compared with the hardness value of the alloy in the as-delivered condition (untreated: 110 HV). This is attributed to the effect of quenching, which prevents the reformation of precipitates and therefore a softening of the alloy. The tempering effect (artificial ageing) of 6082 aluminium alloy shows an increase in hardness and mechanical strength, but a slight decrease in elongation with the decrease in tempering temperature. This phenomenon is due to structural hardening by precipitation, which is the decomposition of a supersaturated solid solution (homogeneous phase having at least two constituents) into a mixture of two phases of different composition: intermetallic iron

IMF and Mg2Si (16). From the above, it can be said that the optimum tempering temperature is at 160°C, due to its best characteristics given for the alloy; in other words, a fairly large hardness, resistance to rupture important and improved ductility, which is necessary for the operating conditions of materials in general.

A comparison of the experimental and numerical curves shows that numerical results are consistent with the experimental results. They exhibit similar trends with an almost constant difference for all the tensile tests, as shown in Figs. 10–15.

Fig. 10. Experimental and numerical curves of the evolution of stress–strain as a function of the tempering temperature of the material in the delivery state T = 25°C

Fig. 11. Experimental and numerical curves of the evolution of stress-strain as a function of the tempering temperature (120°C)

Fig. 12. Experimental and numerical curves of the evolution of stress–strain as a function of the tempering temperature (160°C)

Fig. 13. Experimental and numerical curves of the evolution stress-strain as a function of the tempering temperature (200°C)

Fig. 14. Experimental and numerical curves of the evolution stress-strain as a function of the tempering temperature (240°C)
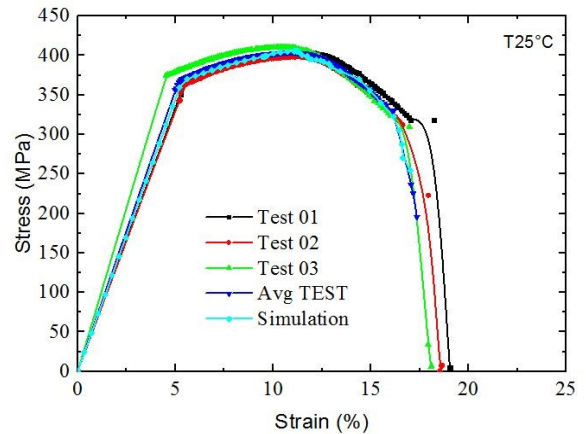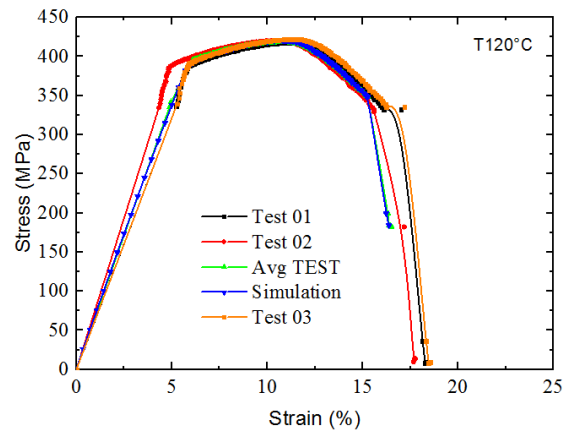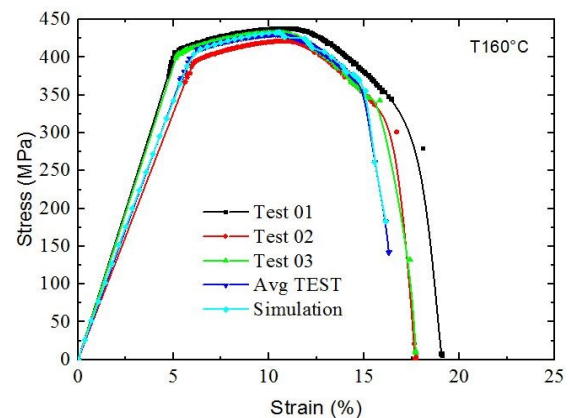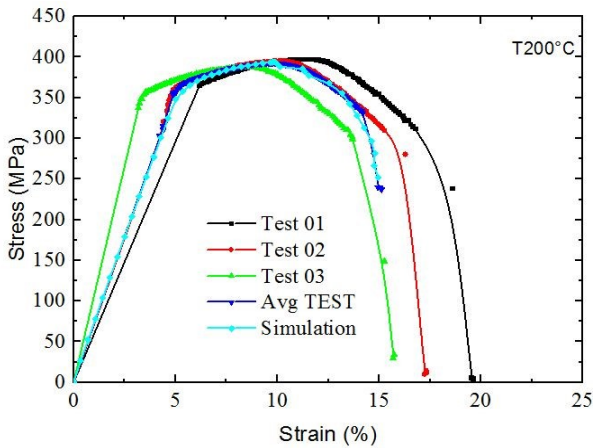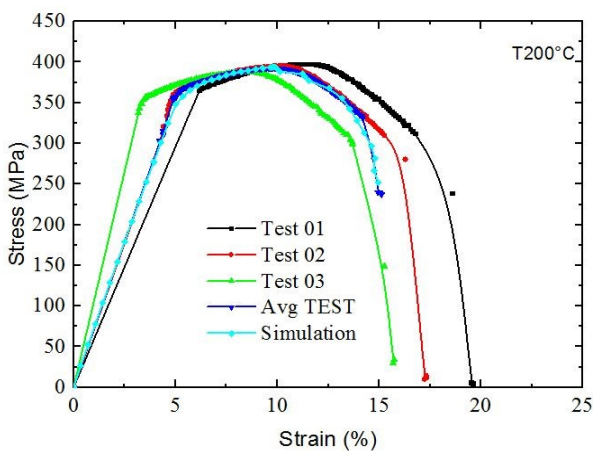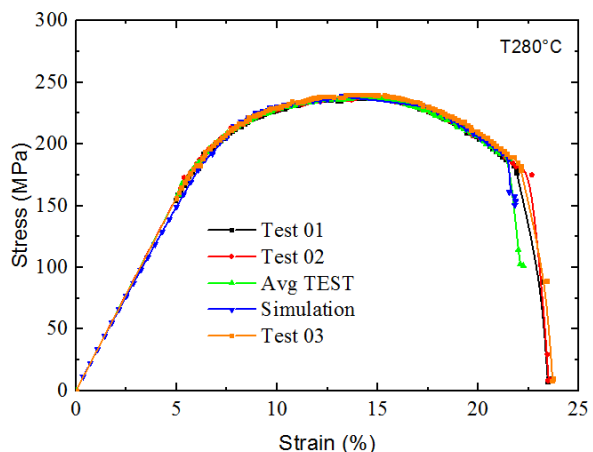

Fig. 15. Experimental and numerical curves of the evolution stress–strain as a function of the tempering temperature (280°C)

To verify the accuracy of the numerical model, tensile tests were conducted using Abaqus at temperatures ranging from 280°C, 240°C, 200°C, 160°C to 120°C. The experimental curves and the numerical load–displacement curves are presented in Figs. 10–15, where their shapes are fairly consistent, and the corresponding ultimate failure models are also reasonably well identified for the different temperatures. To illustrate, let us con-

sider the specimens treated at temperatures of 160°C and 280°C. The experimental results reveal average mechanical characteristics, with the rupture stress recorded at 430.19 MPa and 237.96 MPa, and the corresponding rupture strains at 10.58% and 13.92%, respectively. On the contrary, in the numerical simulation, the rupture stress is estimated to be around 432.89 MPa and 238.84 MPa, while the rupture strains are approximately 10.35% and 13.21%. This leads to small deviations, with errors of approximately 0.62% and 0.37% for the stress values, and 2.17% and 5.10% for the strain values. Furthermore, it is noteworthy that these errors remain within acceptable limits. The comparison between the experimental and simulated data shows a strong agreement, with discrepancies not exceeding 6% for all the cases studied. This outcome emphasises the reliability and accuracy of the simulation results. Similarly, the comparison for the remaining cases exhibits a consistent pattern. The error between the experimental and simulation results for both stress and strain measurements remains below 6%. This demonstrates a high level of agreement and confirms the fidelity of the simulation model. Overall, the analysis of the experimental and simulation data demonstrates that the numerical model accurately captures the mechanical behaviour of the specimens. The results consistently align with the experimental findings, confirming the validity and effectiveness of the simulation approach. Such close agreement between the two datasets provides confidence in the reliability of the numerical simulations and their ability to predict the mechanical response of the specimens under different conditions and temperatures.

Tab. 3 indicates the different results obtained experimentally and numerically of the tensile strength as a function of the tempering temperature.

Tab. 3. Different results experimentally and numerically of the tensile strength

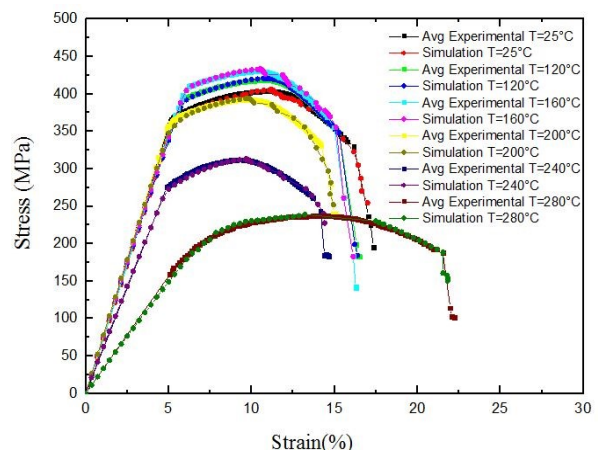| Tempering temperature	R_m (MPa) moy	Simulation R_m	Error %
25	403.94	404	0.01
120	418	420.8	0.67
160	430	432.5	0.58
200	393	392.5	0.13
240	312	313	0.32
280	237.96	238.5	0.23

Fig. 16. Experimental and numerical stress–strain curves from tensile tests

Fig. 16 groups together all the experimental and numerical stress–strain curves for all the tensile tests carried out.

Tab. 4 indicates the different results obtained experimentally, numerically and error of the stress and strain as a function of the tempering temperature. The experimental values obtained represent the average of three tests recorded for each series of speci-mens at different temperatures.

The comprehensive examination of this table unmistakably reveals that the disparity between the experimental and simulation results, be it for stress or strain, is well within the acceptable range.

Tab. 4. Different results obtained experimentally, numerically and error of the stress and strain as a function of the tempering temperature

Tempering temperature (°C)	Exp σ (MPa)moy	Sim σ (MPa)	Error E1 (%)	Exp δ (%)	Sim δ (%)	Error E2 (%)
25	403.94	405.36	0.36	10.92	11.25	3.02
120	418.76	420.87	0.51	10.75	11.04	2.69
160	430.19	432.89	0.62	10.58	10.35	2.17
200	392.532	393.98	0.37	9.78	10.01	2.35
240	412.21	313.05	0.26	9.30	9.69	4.19
280	237.96	238.84	0.37	13.92	13.21	5.10

6. CONCLUSION

The aluminium alloy 6082 plays a crucial role in the industry, and its mechanical properties can be enhanced through suitable heat treatments to enable optimal usage conditions. This study has yielded the following conclusions:

– Quenching the 6082 aluminium alloy resulted in a reduction in its hardness, with an average value of 103 HV after quenching, compared with the initial hardness of the material before quenching, which was 118 HV.
– Tempering above a certain temperature may decrease the hardness of the alloy less than the hardness as delivered, for example, tempering at 280°C reduced the average hardness to 95 HV.
– A decrease in the tempering temperature results in an increase in hardness. The tempering temperature 120°C is required to obtain high hardness 151 HV.
– Heat treatments can modify the mechanical properties of aluminium alloy 6082 in various ways.
– A decrease in the tempering temperature leads to an increase in mechanical strength and a slight decrease in elongation.
– The tensile strength for the material is maximum at the tempering temperature of 160°C (422 MPa) and minimum at tempering temperature 280°C (214 MPa).
– The experimental and simulation outcomes, whether it be for stress or strain, are comfortably within the acceptable range, with an error margin of no more than 6%. These results emphasize a significant alignment between the two sets of data and provide confirmation of the simulation's precision.

REFERENCES

1. Narayana GV, Sharma V, V Diwakar, Kumar KS, Prasad R. Fracture behaviour of aluminium alloy 2219–T87 welded plates. Science and technology of welding and joining. 2004;9(2):121-130. https://doi.org/10.1179/136217104225017035
2. Kang H, Park JY, Choi YS, Cho DH. Influence of the solution and artificial aging treatments on the microstructure and mechanical properties of die-cast Al–Si–Mg alloys. Metals. 2021;12(1):71. https://doi.org/10.3390/met12010071
3. Gabryelczyk A, S Ivanov, Bund A, Lota G. Corrosion of aluminium current collector in lithium-ion batteries: A review. Journal of Energy Storage. 2021;43:103226. https://doi.org/10.1016/j.est.2021.103226
4. Kumar CR, Malarvannan RRR, JaiGanesh V. Role of SiC on Mechanical. Tribological and Thermal Expansion Characteristics of B4C/Talc-Reinforced Al-6061 Hybrid Composite. Silicon. 2020; 12(6):1491-1500. 10.1007/s12633-019-00243-0
5. Watts SJ, Hill RG, O'Donnell MD, Law RV. Influence of magnesia on the structure and properties of bioactive glasses. Journal of Non-Crystalline Solids. 2010;356(9):517-524. https://doi.org/10.1016/j.jnoncrysol.2009.04.074
6. Prabhu Swamy NR, Ramesh CS, Chandrashekar T. Effect of heat treatment on strength and abrasive wear behaviour of Al6061-SiCp composites. Bulletin of Materials Science. 2010;33(1):49-54. 10.1007/s12034-010-0007-y
7. Aboulkhair NT, Maskery I, Tuck C, Ashcroft I, Everitt NM. The microstructure and mechanical properties of selectively laser melted AlSi10Mg: The effect of a conventional T6-like heat treatment. Materials Science and Engineering: A. 2016;667:139-146. https://doi.org/10.1016/j.msea.2016.04.092
8. Wang LF, Sun J, Yu XL, Shi Y, Zhu XG, Cheng LY, Liang HH, Yan B, Guo LJ. Enhancement in mechanical properties of selectively laser-melted AlSi10Mg aluminum alloys by T6-like heat treatment. Materials Science and Engineering: A. 2018;734:299-310. https://doi.org/10.1016/j.msea.2018.07.103
9. Trană E, Rotariu AN, Lixandru P, Matache LC, Enache C, Zecheru T. Experimental and numerical investigation on 6082 0 temper aluminium alloy cartridge tubes drawing. Journal of Materials Processing Technology. 2015;216:59-70. https://doi.org/10.1016/j.jmatprotec.2014.08.032
10. Clausen AH, Børvik T, Hopperstad OS, Benallal A. Flow and fracture characteristics of aluminium alloy AA5083–H116 as function of strain rate, temperature and triaxiality. Materials Science and Engineering: A. 2004;364(1):260-272. https://doi.org/10.1016/j.msea.2003.08.027
11. Field J E, Walley SM, Proud WG, Goldrein HT, Siviour CR. Review of experimental techniques for high rate deformation and shock studies. International Journal of Impact Engineering. 2004;30(7):725-775. https://doi.org/10.1016/j.ijimpeng.2004.03.005
12. Wierzbicki T, Bao Y, Lee YW, Bai Y. Calibration and evaluation of seven fracture models. International Journal of Mechanical Sciences. 2005;47(4):719-743. https://doi.org/10.1016/j.ijmecsci.2005.03.003
13. Mohamed A, Samuel FH. A review on the heat treatment of Al-Si-Cu/Mg casting alloys. Heat Treatment-Conventional and Novel Applications. 2012;1:55-72. http://dx.doi.org/10.5772/79832

14. Rao D, Huber K, Heerens J, dos Santos JF, Huber N. Asymmetric mechanical properties and tensile behaviour prediction of aluminium alloy 5083 friction stir welding joints. Materials Science and Engineering: A. 2013;565:44-50. https://doi.org/10.1016/j.msea.2012.12.014
15. Kang J, Rao H, Zhang R, Avery K, Su X. Tensile and fatigue behaviour of self-piercing rivets of CFRP to aluminium for automotive application. in IOP Conference Series: Materials Science and Engineering. 2016. IOP Publishing.
16. Hillert M. On theories of growth during discontinuous precipitation. Metallurgical and Materials Transactions B. 1972;3(11):2729-2741. 10.1007/bf02652840

Abdelmalek Elhadi: https://orcid.org/0000-0002-9825-9488

Salah Amroune: https://orcid.org/0000-0002-9565-1935

Amin Houari: https://orcid.org/0009-0004-2617-2182

Kouider Madani: https://orcid.org/0000-0003-3277-1187

TRIBOLOGICAL BEHAVIOUR OF CROSS-SHAPED DIMPLES ON SLIDING SURFACES UNDER HYDRODYNAMIC LUBRICATION

Hardik GANGADIA* , Saurin SHETH**

*Research Scholar, Gujarat Technological University, Ahmedabad, Gujarat, India
**Mechatronic Engineering Department, G. H. Patel College of Engineering & Technology, V. V. Nagar, Gujarat, India

hgangadia@gmail.com, saurinsheth@gcet.ac.in

Abstract: This study reports on the tribological behaviour of sliding surfaces having cross-shaped micro-dimples on a surface. One wall is smooth and moving at a constant speed against the other fixed wall with dimples. The laser machine helps to create the desired dimples on the surface of the fixed wall. For the purpose of generating hydrodynamic pressure and tribological behaviour, the effects of cross-shaped dimples and oriented cross-shaped dimples have been compared with circular-shaped dimples. Additionally, the impact of sliding speed, dimple area density and depth on tribological behaviour was examined. The findings show that compared with a circular-shaped dimple, an unconventional cross-shaped and orientated cross-shaped dimple generates a higher net hydrodynamic pressure in the fluid domain and offers superior stability between the sliding surfaces. It has been demonstrated that geometrical factors like dimple depth and area density as well as operational factors like sliding speed have a substantial impact on the hydrodynamic average pressure and tribological behaviour of sliding surfaces. The experimental findings indicate that, for the same geometric and operating parameters, cross- and orientated cross-shaped dimples have a 20%–25% lower friction coefficient between the sliding surfaces than circular dimples. The results of the experiment support those of the analysis and CFD.

Key words: surface texturing, tribology, hydrodynamic lubrication, sliding surfaces, CFD, LST

1. INTRODUCTION

Hydrodynamic lubrication between sliding surfaces in components such as journal bearings and mechanical seals plays a significant role in improving the sliding surface tribological behaviour. Many studies have been conducted to improve the hydrostatic and hydrodynamic load capacity of sliding surfaces. Researchers have discovered that coning, misalignment, waviness and hydropads produce positive pressure and increase the load capacity of sliding surfaces [1–6].

Laser texturing has grown as a proven method in mechanical components over the past two decades. When one of the sliding surfaces is texturised, there is a rise in hydrodynamic pressure and load-carrying capacity, as well as a reduction in generated friction and wear. Many locales are now using surface texturing to get tribological benefits. Surface texturing techniques such as shot blasting [7], laser texturing [8, 9], reactive ion etching [10] and micro-electrolytic etching [11] are now widely available. When compared with other technologies, laser texturing provides unparalleled control of the surface microstructure while having a low environmental impact.

When the upper surface slides, the fluid is pushed forward by the relative motion of the sliding surfaces. When the fluid enters the dimple as it flows forward, the pressure reduces owing to diverging action, and as it moves out of the dimple, the pressure rises due to converging action. This difference in pressure drop and pressure rise results in net positive pressure creation, which enhances the load-carrying capacity between the sliding surfaces.

Etsion and Burstein [12] introduced a mechanical seal model with a hemispherical regular micro-surface on one of the mating seal faces, resulting in improved seal performance in terms of lower friction. According to Etsion et al. [13, 14], optimal pore depth over pore diameter maximises fluid film stiffness in laser textured surfaces. So far, the investigation has primarily concentrated on circular micro dimples [12–16] because of their ease of fabrication. Some researchers concentrated on conventionally shaped micro dimples such as square, triangle and ellipse [17–20], while others [18, 19] discovered that an elliptical dimple placed perpendicular to the direction of sliding produced the best load-carrying capacity results. Qui et al. [21] optimised the texture geometry and density of six different dimple shapes for maximum load-carrying capacity in the case of gas-lubricated parallel sliding bearings, concluding that the ellipsoidal dimple shape results in the highest load-carrying capacity and discovering that optimal geometry and density are almost independent of operating conditions. Following that, Qui et al. [22] adjusted the texture geometry and density of six alternative dimple shapes in terms of least friction coefficient and maximum bearing stiffness, reporting that the ellipsoidal shape yielded the lowest friction coefficient and the highest bearing stiffness. According to Raeymaekers et al. [23] and Yan et al. [11], dimple area density is also an important characteristic that influences the tribological behaviour of sliding surfaces.

Researchers [24–27] are now using commercial CFD programmes like FLUENT to compute the pressure and velocity distribution characteristics of lubricant flow between textured sliding surfaces. Recently, Liu et al. [26] used numerical simulation to study the geometry and operating parameters of spherical micro dimples on the tribological performance of textured surfaces

sciendo

Hardik Gangadia, Saurin Sheth
Tribological Behaviour of Cross-Shaped Dimples on Sliding Surfaces Under Hydrody-Namic Lubrication

DOI 10.2478/ama-2024-0043

and recommended a dimple area density of 25%–35% in practical applications. Wei et al. [27] conducted a numerical analysis to determine the effect of geometric features on pressure build-up for various dimple shapes. This article also presented a new parameter called dimple surface angle, and it was discovered that the most significant changes occur when the dimple surface angle is <30°. Much study has been done to date to investigate the load-carrying capacity of one standard dimple shape (i.e. round, square, triangle, spherical, etc.) rather than the tribological characteristics of unconventional-shaped texturing. Non-traditional dimple shapes, on the contrary, have more potential for convergence and thus better tribological benefits than standard dimple shapes, which cannot be ignored. Gangadia and Sheth [28, 29] studied the effect of bowtie and star dimple shapes on circular dimple shapes to generate hydrodynamic pressure between two parallel sliding surfaces and found that when the sliding speed, area density and dimple depth are all the same, the oriented bowtie-shaped dimple produces the most hydrodynamic pressure.

In this study, an analytical model is developed to estimate the hydrodynamic pressure generation for the application of laser textured mechanical seals, followed by CFD analysis, to explore the effect of straight and orientated cross-shaped dimples over circular-shaped dimples. The experimental results matched well with the analytical method and CFD results. Coefficient of friction was examined for the parametric analysis to study the effect of different geometric characteristics such as dimple depth and dimple area density on hydrodynamic pressure. Operating parameters such as sliding speed are also covered.

2. ANALYTICAL MODEL

The geometrical model of laser textured mechanical seal is presented in Fig. 1. Inner and outer radii of the seal ring are ri and ro, respectively, and the values for the same has been taken as 0.0181 m and 0.0235 m, respectively. Each dimple is modelled by radius rd and depth HD. Laser textured dimples are distributed uniformly on one of the faces of seal. Each dimple is modelled at the centre of the imaginary square cell with length of 2r1 as shown in Fig. 1(c). Length of the imaginary square cell with relation to circular dimple area density Sp is as follows:
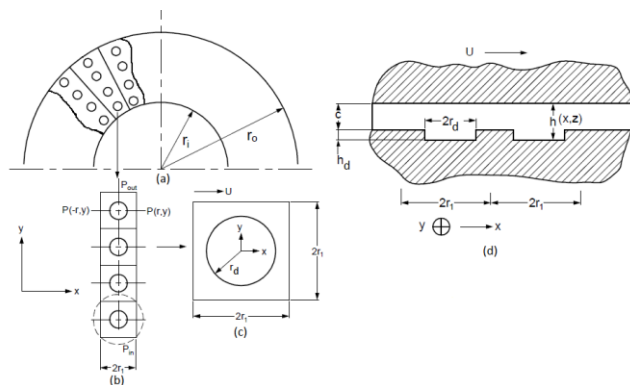
$$2r_1 = \sqrt{\frac{\pi}{S_d} \times r_d} \tag{1}$$

Fig. 1. (a) Geometric model of laser textured mechanical seal (b) Dimple column (c) Dimple cell (d) Cross-section of the textured seal

Fig. 1(d) shows the cross-section of the textured seal. One of the two surfaces of the seal is provided with a micro-texture and the other surface is flat. The flat surface moves with relative velocity U to the textured surface. The minimum distance between the bearing surfaces is indicated by c and the local distance is indicated by h(x, z), where x and z are Cartesian coordinates as shown in the Fig 1. It is assumed that seal faces are separated by incompressible viscous fluid.

The generalised Reynolds's equation for mechanical seal is expressed as:

$$\frac{\partial}{\partial x}\left(h^3 \frac{\partial p}{\partial x}\right) + \frac{\partial}{\partial z}\left(h^3 \frac{\partial p}{\partial z}\right) = 6\mu U \frac{\partial h}{\partial x} \tag{2}$$

Modified Reynold's equation in order to deal with cavitation is:

$$\frac{\partial}{\partial x}\left(K\beta h^3 \frac{\partial \xi}{\partial x}\right) + \frac{\partial}{\partial z}\left(K\beta h^3 \frac{\partial \xi}{\partial z}\right) = 6\mu U \frac{\partial(\xi h)}{\partial x} \tag{3}$$

Dimensionless Reynold's equation is as follows:

$$\frac{\partial}{\partial X}\left(KH^3 \frac{\partial \xi}{\partial X}\right) + \frac{\partial}{\partial Z}\left(KH^3 \frac{\partial \xi}{\partial Z}\right) = \Lambda \frac{\partial \xi H}{\partial X} \tag{4}$$

Only one radial column of dimples is taken for consideration as it is assumed that micro-dimples are evenly distributed. Boundary condition for one radial column is as shown in Fig. 1(b):

$$p(x, z = r_i) = p_{in}$$

$$p(x, z = r_o) = p_{out} \tag{5}$$

In this analysis the pressure distribution is assumed to be periodic in the direction of the circumference with the same period as the length of the imaginary square cell. Therefore, in the direction of the circumference, the periodic condition of the pressure is applied as:

$$p(x = -r_1, z) = p(x = r_1, z) \tag{6}$$
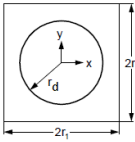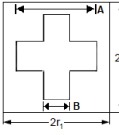
Boundary conditions in dimensionless form are given as follows:

$$P\left(X, Z = \frac{r_i}{r_d}\right) = \frac{p_{in}}{p_a}$$

$$P\left(X, Z = \frac{r_o}{r_d}\right) = \frac{p_{out}}{p_a}$$

$$P\left(X = -\frac{r_1}{r_d}, Y\right) = P\left(X = \frac{r_1}{r_d}, Y\right) \tag{7}$$

Circular, cross and oriented cross-shaped dimples are considered in this study, which can be fabricated using LST. All dimples have flat bottoms and flat boundary walls and are symmetric to the XZ plane. Dimple-shaped geometry, the equivalent dimension of dimple with respect to area of circular dimple, cell size and dimensionless local spacing H(X,Y) between the textured and flat surface within one unit cell for cross shape dimple are described in Tab. 1.

Equation 3 can be solved for the pressure distribution in the seal clearance by a finite difference method with a relaxation factor of 1.4. Uniform Cartesian grid of 100 × 100 nodes is selected based on convergence and accuracy. The opening force to prevent contact between seal faces can be obtained by integrating the pressure over the seal area. SAE 30 oil was used in analysis and experimental work, which has dynamic viscosity of 0.29 kg/m/s at 20°C and density of 875 kg/m³.

Tab. 1. Details of all dimple-shaped geometries

Shape	Equivalent dimensions	Cell size	Conditions for film thickness H(X,Y)
 Circular	$r_d = \sqrt{\dfrac{A_{circle}}{S_d}}$	$2r_1 = \sqrt{\dfrac{\pi}{S_d} \times r_d}$	$\begin{cases} 1+\dfrac{\varepsilon}{\delta} & if\ X^2+Y^2 \leq 1 \\ 1 & if\ X^2+Y^2 > 1 \end{cases}$
 Cross	$B = \sqrt{\dfrac{\pi}{5}} \times r_d$ And $A = 3B$	$2r_1 = \sqrt{\dfrac{5B^2}{S_d}}$	$\begin{cases} 1+\dfrac{\varepsilon}{\delta} & if\ \|X\| \leq \dfrac{1}{\sqrt{10}}\ AND\ \|Y\| \leq \dfrac{3}{\sqrt{10}} \\ & or\ \|Y\| \leq \dfrac{1}{\sqrt{10}}\ AND\ \|X\| \leq \dfrac{3}{\sqrt{10}} \\ 1 & else \end{cases}$

3. CFD MODEL

CFD analysis has been done to reconfirm the results obtained from the analytical model prior to the experimental work. Fig. 1 (d) is reconstructed and shown as Fig. 2, which represents the geometric parameters of the CFD model. The length of the domain is denoted by $2r_1$ and it will remain constant for all the analyses. The quantity c is the fluid film thickness which is set in advance and quantity HD is the depth of the dimple.
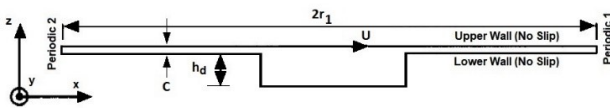
Fig. 2. Geometric parameters and boundary conditions of model

The fluid domain's boundary conditions are likewise depicted in the same Fig. 1(d). At both the upper and lower walls, fluid has no slip condition. A periodic boundary condition confines the edges in the *x*-direction. The bottom wall is stationary, while the upper wall with an untextured surface slides in the *x*-direction with constant velocity *U*. Fig. 3 depicts the design and modelling of circular, cross and oriented cross surface texturing shapes for hydrodynamic analysis.
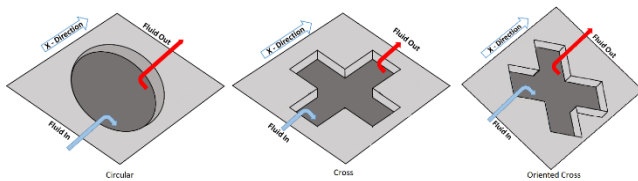
Fig. 3. Designs of dimple shapes

The lubricant is considered to be an incompressible Newtonian fluid with constant viscosity and density and no body force when filled between sliding pairs. The flow is described as laminar and isothermal. It is also believed that the effect of fluid pressure on the formation of a solid structure is insignificant. The flow of lubricant is governed by Navier–Stokes (momentum) equations and the continuity equation, which can be expressed as:

$$\rho(v.\nabla)v = -\nabla p + \nabla.(\mu \nabla v) \tag{8}$$

$$\nabla.v = 0 \tag{9}$$

The pressure distribution of lubricant between sliding pairs was estimated using the commercial CFD software FLUENT. The process for improving performance in creating hydrodynamic pressure in textured sliding surfaces is assumed to be connected to cavitation generation inside the dimples. In this case, the lubricant pressure may fall below the gas saturation pressure at the diverging zone, resulting in cavitation. The Rayleigh–Plesset multiphase cavitation model [30], which has been successfully evaluated in references [31, 32], is employed. This is a multi-phase model in which lubricant vapour is created when the pressure drops below saturation. The results are obtained in dimensional form but given in nondimensional form to allow for easy comparison of different outcomes.

4. EXPERIMENTAL SET-UP

All of the experiments were carried out on a pin on disc wear testing machine, as indicated schematically in Fig. 4.
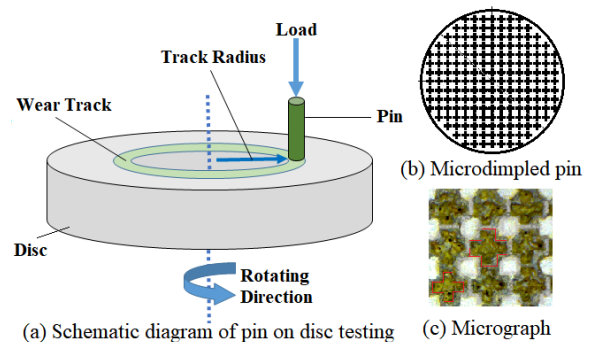
Fig. 4. Experimental set-up and specimen preparation (a) Schematic diagram of pin on disc testing (b) Microdimpled pin (c) Micrograph

The lower disc is fixed and can be rotated at any speed between 100 rpm and 2,000 rpm using a motor. The EN31 disc measures 165 mm in diameter and 8 mm in thickness. The textured pin is attached to a metallic holder that allows for vertical loading and self-alignment. SAE 30 oil is used as lubricating fluid. The mechanical assembly, base plate, guide base, AC motor, AC drive, spindle assembly, loading arrangement with pulley and 1:1

leverage ratio, loading pan with dead weights, wear disc holder, scale to set the track radius, LVDT for displacement and load cell for frictional force are the main components of the pin-on-disc wear testing apparatus. Sliding takes place between a rotating disc and a fixed pin. To fit the test conditions, the normal load and rotating speed can be changed. With the use of electronic sensors, frictional force and wear (displacement) are measured and recorded into a PC using Mag-view 2015 software. Software on the computer directly displays the value of the friction coefficient, which was calculated based on the magnitude of the frictional force. The sensor is affixed to the textured pin, which can measure the coefficient of friction. Data from sensors are acquired on a computer using an acquisition system.

4.1. Specimen preparation

The cylindrical textured pin measures 10 mm $\times$ 30 mm and 78.5 mm^2 in contact area with the disc. Cylindrical textured pins are made of EN31 steel with hardness of 45–50 HRC. Each pin was then polished and optically flattened before being examined for flatness with a monochromatic check light. A laser marking machine is used to texture the pins. A dimple drawing that meets the necessary dimensions is created and imported into the laser marking machine. The loop count is adjusted in accordance with the required depth, speed is set to 150 mm/s, power is set to 100% and frequency is set to 20 KHz. As illustrated in Fig. 4(b), micro-dimples are distributed evenly across the pin. Each trial began with the pins being cleaned with a cleaner and the weight of each pin being measured with a digital electronic weighing machine.

4.2. Test procedure

For the test first the textured pin was installed with a metallic holder and the flow of lubricant was started. Track radius is kept at 40 mm. After reaching a certain level of lubricant, the motor is started. The force sensor measures the friction force between the two samples by means of a piezoelectric (load cell) sensor as an acquisition mechanism. By dividing the normal force by the friction force produced in direct contact between the two surfaces as determined by a data collecting device, one can get the friction coefficient. Then the values of coefficient of friction were recorded for different speeds and loading conditions. When the test is over, the textured pin is replaced by a new one.

5. RESULTS AND DISCUSSION

More hydrodynamic pressure between the sliding surfaces results in a larger separation being maintained between them, which reduces friction. Greater tribological advantage and less friction are associated with increased hydrodynamic pressure. While the results of the experiments will be converted directly into a friction coefficient, the simulations will yield pressure as their output.

5.1. Effect of dimple shape

Fig. 5 illustrates the pressure distributions for a circular dimple, a cross dimple and a cross dimple positioned at 45°. All vari-

ables, such as sliding speed (U) = 6 m/s, dimple depth (h_p) = 15 m, dimple area density = 30% and fluid film thickness (c) = 2 µm, are held constant during the simulation for each shape to enable simple comparison. Fig. 5 shows that when the fluid flows along the x-axis and enters the dimple, the pressure drops owing to divergence initially. As the fluid travels forward, the pressure builds up due to the dimple's convergence, and this generated pressure is greater than the pressure decrease during divergence, resulting in a net positive pressure. Because all other geometric and operational parameters remain constant, the variation in net pressure is related solely to differences in dimple geometric shapes. However, geometric parameters such as dimple depth and area density, as well as operational parameters such as sliding speed, have a considerable role in the formation of hydrodynamic pressure.
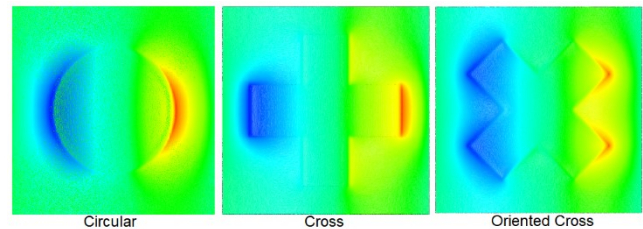
Fig. 5. Pressure distribution in the fluid domain for different shapes
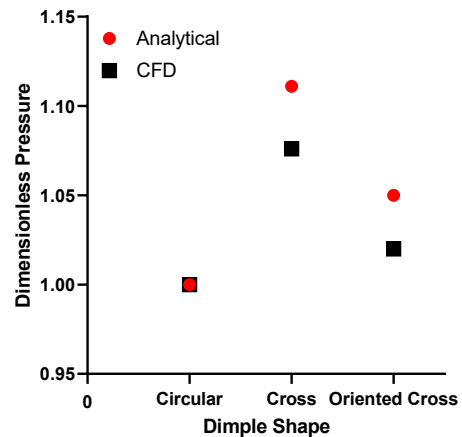
Fig. 6. Comparison of dimensionless average pressure

Fig. 6 compares dimensionless average pressures for the dimple shapes discussed in this paper. The graphic clearly shows a significant correlation between the results of the analytical investigation and the CFD study. Because of the significant convergence around the dimple's periphery, the cross-shaped dimple generates approximately 10% more hydrodynamic pressure than the circular-shaped dimple for the values of geometric and operating parameters mentioned above. In the case of an orientated cross-shaped dimple, pressure build-up is also high.

5.2. Effect of dimple depth

Fig. 7 depicts pressure distributions for minimum and maximum depth for cross dimples and oriented cross dimples. To conveniently capture the hydrodynamic effect of depth for cross

and oriented cross dimples, all other variables are maintained identical. Fig. 7 clearly shows that as dimple depth increases, the convergence becomes more concentrated in the case of both cross and orientated cross-shaped dimples.
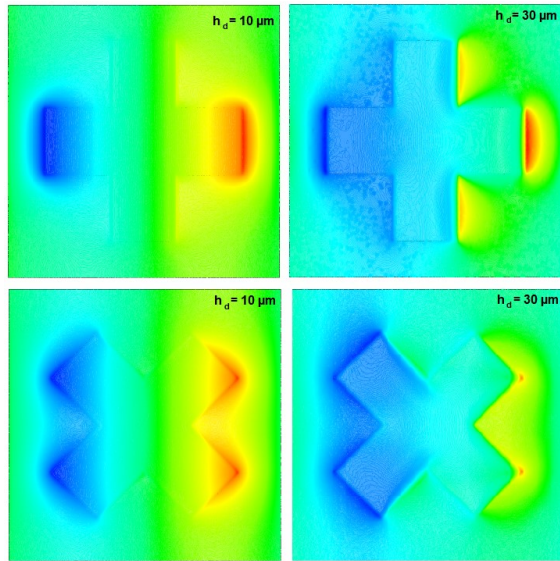
Fig. 7. Pressure distribution in the fluid domain for different depths
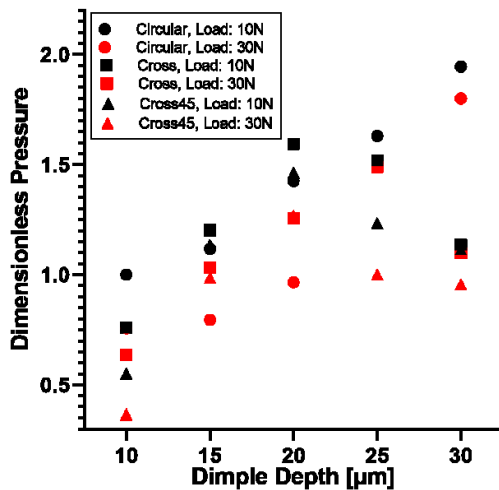
Fig. 8. Dimensionless pressure as a function of dimple depth

Fig. 8 shows the dimensionless pressure as a function of dimple depth for two different loading conditions and dimple forms. In the case of circular dimples, the dimensionless pressure increases as the dimple depth grows. One way to think about it is that when the dimple depth increases, more lubricant gets trapped, which lubricates the sliding surfaces even more. Conversely, when the dimple depth increases, the lubricant undergoes a greater shear force, leading to the formation of microvortices. By narrowing the gap between the sliding surfaces and raising friction, these vortices may cause the lubricant to be sucked inward. In the case of cross and oriented cross dimples, dimensionless pressure increases as dimple depth grows from 10 µm to 20 µm and then drops as dimple depth climbs further. It is also discovered that cross dimples generate more pressure than circular dimples in the 15–20 µm dimple depth range. It is clear that dimple depth has an effect on the generation of hydrodynam-

ic pressure.

To understand the effects of dimple depth on tribological performance, the pressure distribution of the lubricant is analysed. Figs. 9–11 depict the pressure distribution of various dimple shapes along the *x*-direction at various depths, while Figs. 12–14 depict the pressure distribution of various dimple shapes along the *y*-direction at various depths. As the liquid enters the dimple, the pressure decreases and rises owing to convergence going in the *x*-direction.
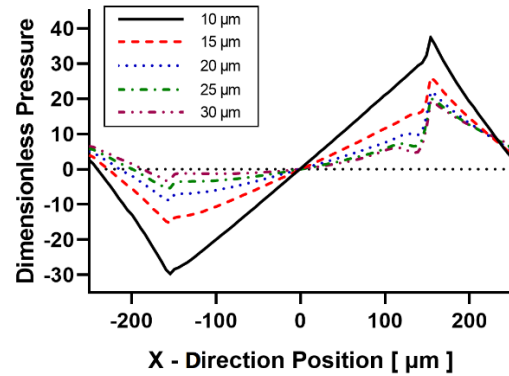
Fig. 9. Effect of dimple depth on pressure distribution along x-direction
(for circular dimple)

The maximum magnitude of positive pressure is probably larger than the maximum magnitude of negative pressure. As a result of the net pressure rises, load-carrying capacity increases and friction reduces. The sliding surface's stability is dependent on how evenly the pressure is distributed. The stability of the sliding surface and the homogeneity of the pressure distribution increase with decreasing negative and positive pressure differential. Figs. 9–11 show that as the dimple depth increases, the difference between negative and positive pressure in the *x*-direction reduces and stability increases. It can also be seen that the pressure distribution in the *x*-direction is smoother in the case of oriented cross-shaped dimples, resulting in improved stability between the sliding surfaces. When discussing pressure distribution in the *Y*-direction, it can be seen in Figs. 12–14 that for each depth considered, approximate straight line of pressure distribution is observed in the case of cross-shaped dimples and two peaks of same level are seen in pressure distribution lines in the case of oriented cross-shaped dimples, and thus cross-shaped dimples and oriented cross-shaped dimples offer better stability than circular-shaped dimples between the sliding surfaces.
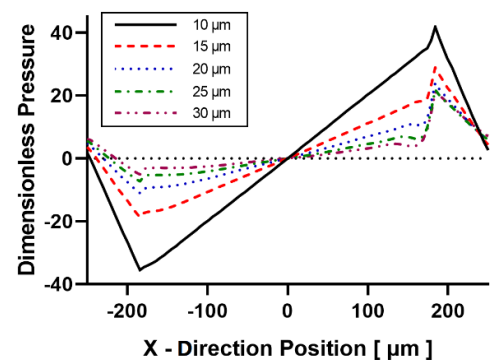
Fig. 10. Effect of dimple depth on pressure distribution along x-direction
(for cross dimple)

Hardik Gangadia, Saurin Sheth
Tribological Behaviour of Cross-Shaped Dimples on Sliding Surfaces Under Hydrody-Namic Lubrication

DOI 10.2478/ama-2024-0043
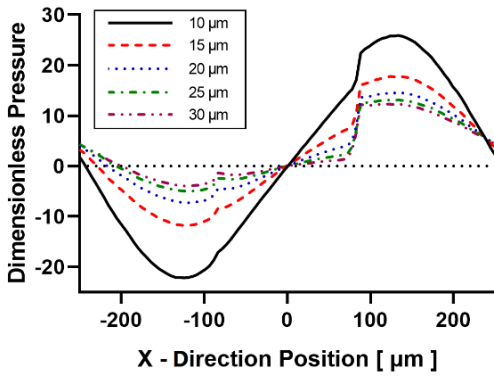
Fig. 11. Effect of dimple depth on pressure distribution along x-direction (for oriented cross dimple)
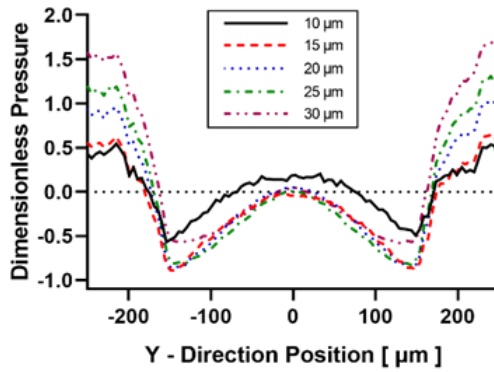
Fig. 12. Effect of dimple depth on pressure distribution along y-direction (for circular dimple)
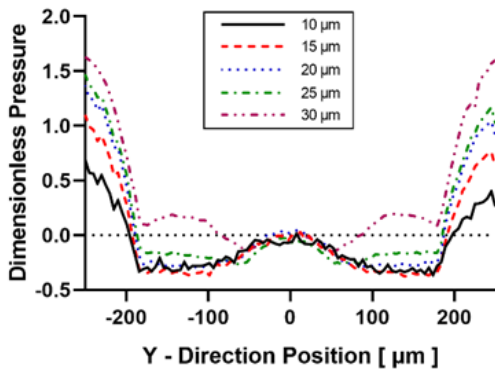
Fig. 13. Effect of dimple depth on pressure distribution along y-direction (for cross dimple)
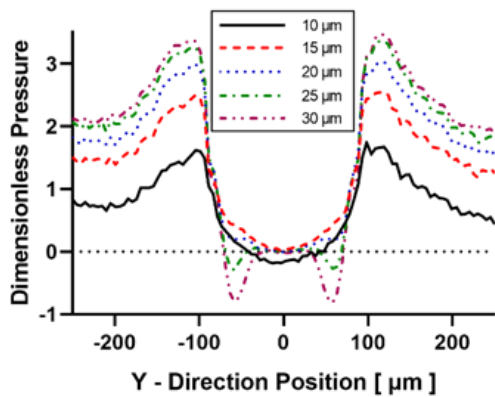
Fig. 14. Effect of dimple depth on pressure distribution along y-direction (for oriented cross dimple)

5.3. Effect of dimple area density
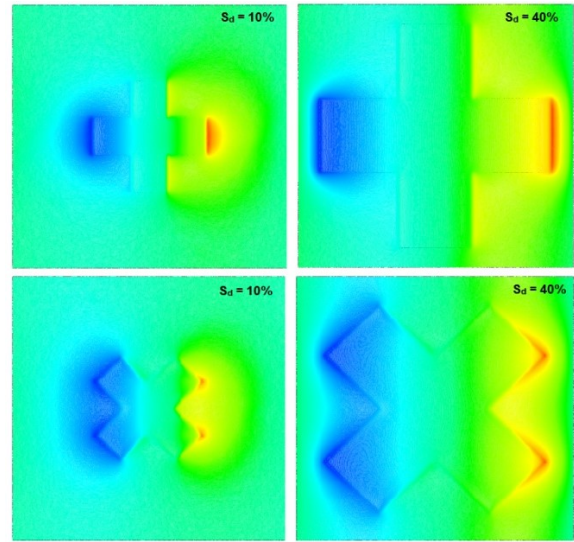
Fig. 15. Pressure distribution in the fluid domain for different dimple area density

Fig. 15 shows pressure distributions for minimum and maximum dimple area density for cross dimple and oriented cross dimple. To conveniently capture the hydrodynamic effect of dimple area density for cross and oriented cross dimples, all other variables are held constant. The graphic clearly shows that when dimple area density increases, the amount of generating pressure decreases due to weaker convergence.
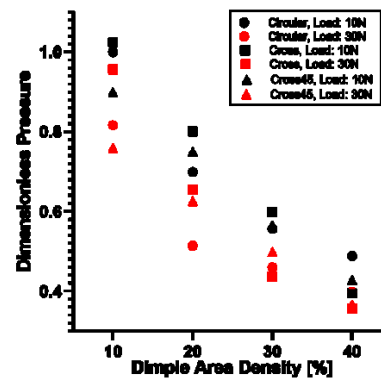
Fig. 16. Dimensionless pressure as a function of dimple area density

Fig. 16 depicts dimensionless pressure as a function of dimple area density for several dimple shapes under two different loading conditions. When the dimple area density is increased from 10% to 40%, the pressure declines steadily. Increasing the dimple area density could have a negative effect on the generation of hydrodynamic pressure. The shape of the dimple resembles a plane plate where micro hydrodynamic pressure is less likely to be generated when the dimple size is raised while the cell size stays constant. For a dimple area density of 10%, the pressure generated by a cross dimple is greater than the pressure generated by a circular dimple. The pressure generated by 20% and 30% dimple area density is greater than the pressure generated by circular dimple in both cross-shaped dimple and oriented cross-shaped dimple. When the dimple area density is increased

to 40%, circular dimples provide maximum pressure for both 10N and 30N loading conditions. It is obvious from this that modifying the dimple area density has a direct effect on the creation of hydrodynamic pressure.
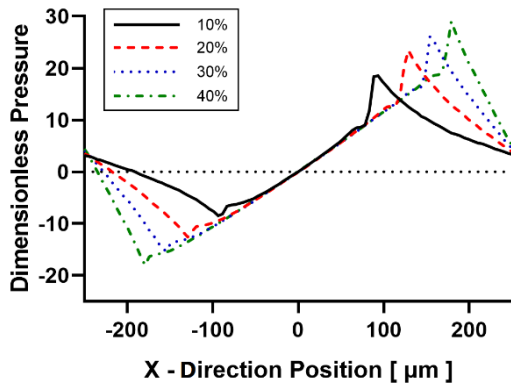
Fig. 17. Effect of dimple area density on pressure distribution along x-direction (for circular dimple)
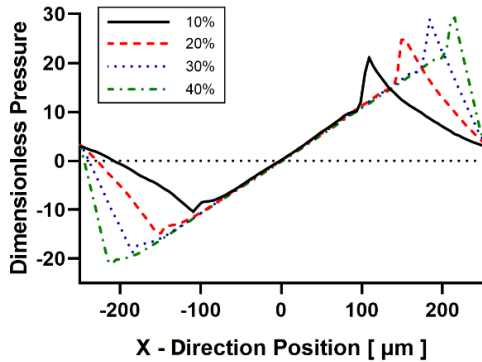
Fig. 18. Effect of dimple area density on pressure distribution along x-direction (for cross dimple)
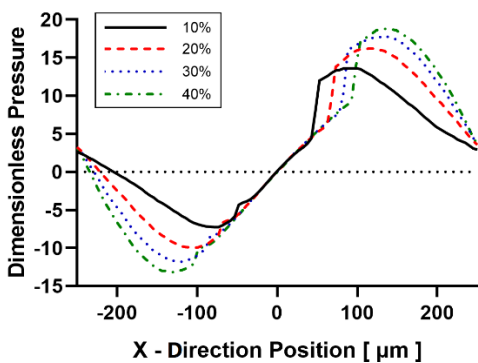
Fig. 19. Effect of dimple area density on pressure distribution along x-direction (for oriented cross dimple)

The lubricant's pressure distribution is studied in order to investigate the effects of dimple area density on tribological performance. Figs. 17–19 depict the pressure distribution of various dimple forms along the *x*-direction at various dimple area densities, while Figs. 20–22 depict the pressure distribution of various dimple shapes along the *y*-direction at various dimple area densities. As previously stated, as the liquid enters the dimple, the

pressure decreases and the pressure rises owing to convergence going in the *x*-direction.
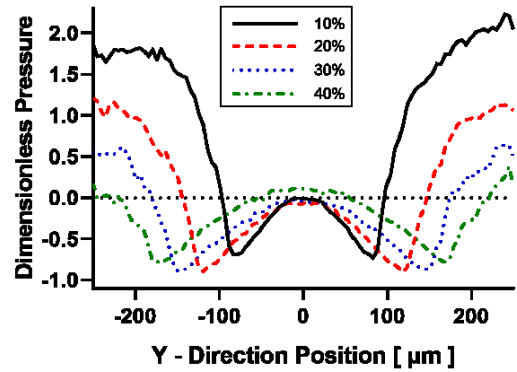
Fig. 20. Effect of dimple area density on pressure distribution along y-direction (for circular dimple)
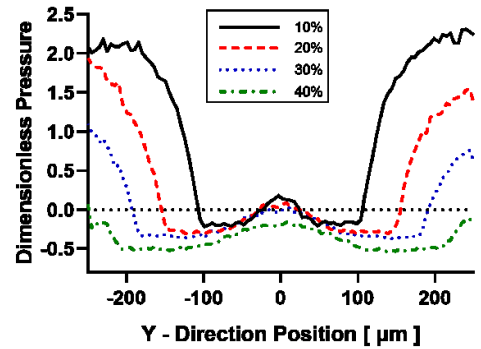
Fig. 21. Effect of dimple area density on pressure distribution along y-direction (for cross dimple)
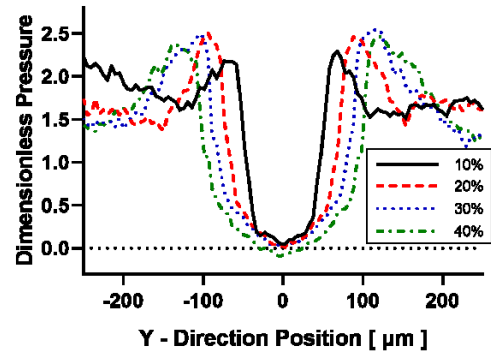
Fig. 22. Effect of dimple area density on pressure distribution along y-direction (for oriented cross dimple)

The increase in net pressure can also be noticed here since the maximum magnitude of positive pressure is greater than the maximum magnitude of negative pressure. As a result, there is a tribological advantage between the sliding surfaces. Figs. 17–19 show that as dimple area density grows, so does the difference between negative and positive pressure in the *x*-direction. There is uniformity in the behaviour of circular dimples and cross-shaped dimples. However, when we talk about oriented cross dimples, the pressure distribution is more smooth and stable. When we look at the pressure distribution in the *Y*-direction, we

sciendo

Hardik Gangadia, Saurin Sheth   DOI 10.2478/ama-2024-0043
Tribological Behaviour of Cross-Shaped Dimples on Sliding Surfaces Under Hydrody-Namic Lubrication

can see that as the dimple area density grows from 10% to 40%, there is a little decrease in pressure for each form of dimple. The pressure distribution is more stable in the case of cross-shaped dimples than in the case of circular-shaped dimples and oriented cross dimples. Comparing Figs. 20 and 21, it is evident that the cross-shaped dimple improves the stability between the sliding surfaces by providing a more uniform pressure distribution at the centre than the circular dimple does. Over all the other scenarios evaluated, a cross-shaped dimple with 40% density provides the most stability.

5.4. Effect of speed

It has been demonstrated numerous times that the hydrodynamic pressure between sliding surfaces increases with increasing speed. However, if one of the sliding surfaces is textured, the hydrodynamic effects of speed on dimples of various shapes may differ. Fig. 23 illustrates the dimensionless pressure as a function of sliding speed for three different dimple shapes. Fig. 23 clearly shows that under both loading situations, there is little variation in the generated hydrodynamic pressure for all dimple forms analysed at a speed of 3 m/s, but this disparity grows as the speed increases.
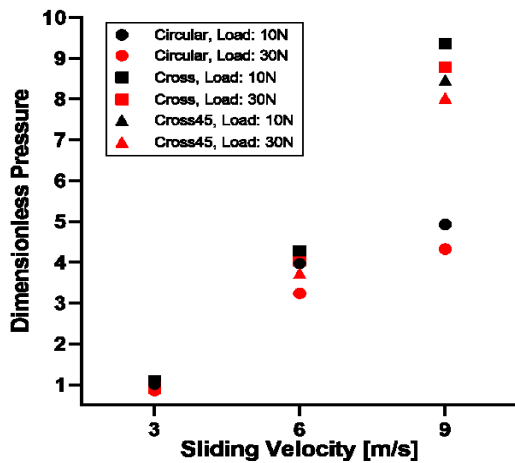
Fig. 23. Dimensionless pressure as a function of sliding speeds

Cross dimples and oriented cross dimples generate higher hydrodynamic pressure than circular dimples at 6 m/s, but the difference grows significantly larger at 9 m/s. The pressure created in the cross dimple is approximately 100% larger than the pressure generated in the circular dimple at a sliding speed of 9 m/s. Thus, using cross dimples as texturing rather than traditional circular dimples leads to a significant boost in hydrodynamic pressure generation and thus tribological benefits.

6. EXPERIMENTAL RESULTS

Fig. 24 represents the variation of the friction coefficient throughout several testing for circular, cross and orientated cross-shaped dimples. For each dimple shape, the dimple depth is set to 30 μm, the dimple area density is set to 40% and the RPM of the disc is set to 6 m/s sliding speed. The graph shows that the friction

coefficient for cross-shaped dimples and oriented cross-shaped dimples is substantially lower than the friction coefficient for circular-shaped dimples. As a result, cross-shaped and oriented cross-shaped dimples give better tribological advantages in sliding textured surfaces than traditional circular-shaped dimples.

Figs. 25–27 show the influence of dimple depth and density on average friction coefficient under uniform loading conditions for various dimple forms. Dimple depth is assumed to be 10 μm and 30 μm, and dimple area density is assumed to be 10% and 40%. As a result, four findings for three distinct speeds for each dimple shape are compared. As it can be seen, when rotational speed increases, the average friction coefficient drops. The lubricant film is created as a result of the rough surface, and film formation is easier at greater rotational speeds. The effect of dimple depth and dimple area density on tribological performance, on the other hand, appears to be significant.
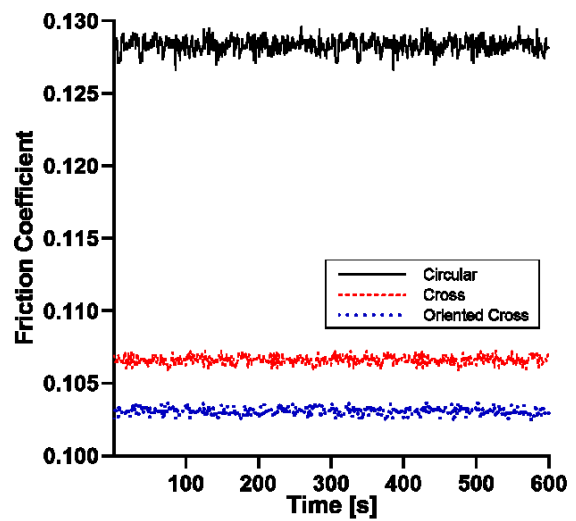
Fig. 24. Variation of friction coefficient during different tests for different dimple shapes
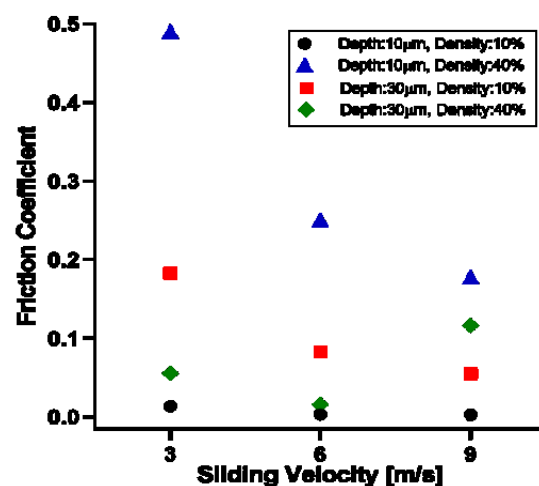
Fig. 25. Friction coefficient as a function of sliding velocity for circular shaped dimple

Maximum friction coefficient is found for a combination of 10 μm dimple depth and 40% dimple area density at each speed and for each shaped dimple, and minimum friction coefficient is found

for a combination of 10 µm dimple depth and 10% dimple area density. Cross-shaped dimples have the lowest friction coefficient of all the shapes studied, with a velocity of 9 m/s for a combination of 10 µm dimple depth and 10% dimple area density. These experimental findings indicate a strong connection with the CFD and numerical model results.
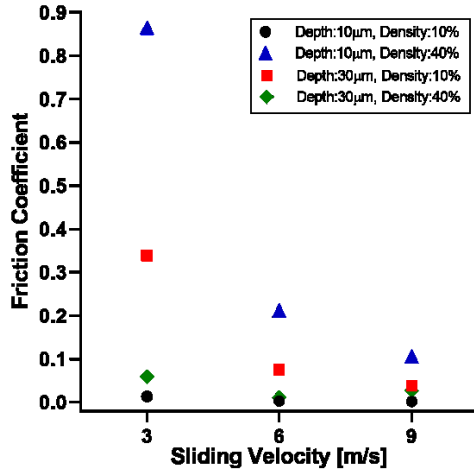
Fig. 26. Friction coefficient as a function of sliding velocity for cross-shaped dimple
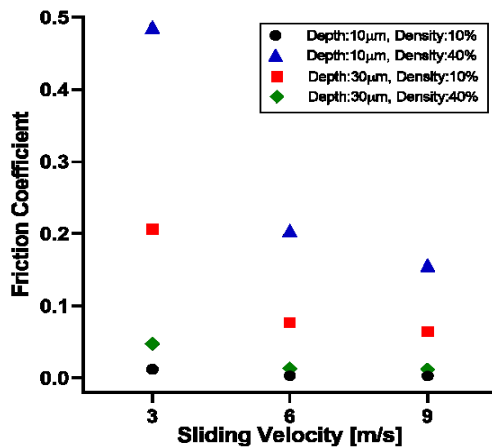
Fig. 27. Friction coefficient as a function of sliding velocity for oriented cross-shaped dimple

7. CONCLUSIONS

Using CFD and numerical models, it was found how cross and oriented cross dimple shapes affected the generation of hydrodynamic pressure between two parallel sliding surfaces. The tribological performance of textured samples with various types of dimples was then evaluated using pin on disc testing. The outcomes are listed below.

1. The cross-shaped dimple provides the most hydrodynamic pressure of the three dimple types when the dimple depth, dimple area density and sliding speed are all identical.
2. Dimple depth and area density have an effect on tribological behaviour and hydrodynamic pressure generation. Cross-shaped dimples and orientated cross-shaped dimples produce hydrodynamic pressure more effectively than circular dimples

for each dimple depth and dimple area density considered in the analysis.

3. The generated hydrodynamic pressure for all of the analysed dimple forms is about the same for all loading conditions at a speed of 3 m/s, but the discrepancies become more obvious as the speed increases. Cross-shaped dimples and oriented cross-shaped dimples produce greater hydrodynamic pressure than circular dimples at 6 m/s, but the difference becomes substantial at 9 m/s.
4. Cross-shaped dimples and oriented cross-shaped dimples provide more stability as pressure is distributed more evenly both in the x-direction and y-direction in the case of cross-shaped dimples and oriented cross-shaped dimples than it is in the case of circular-shaped dimples.
5. The shape of the dimple also has an effect on the friction coefficient. The test results accord well with those from the numerical and CFD models. The experimental findings indicate that, for the same geometric and operating parameters, cross- and orientated cross-shaped dimples have a 20%–25% lower friction coefficient between the sliding surfaces than circular dimples. The density and depth of dimples appear to have a substantial impact on tribological performance.
6. When used to obtain a tribological advantage on sliding surfaces under a variety of test load, dimple depth, dimple area density and rotational speed conditions, unconventional (cross or orientated cross)-shaped texturing outperforms traditional circular texturing.

It should be mentioned that there can be slight variations in the outcomes if the surface-to-surface contact is significant. To achieve the best response, surface textures' shape, orientation and geometric parameters can be optimised for particular applications and mechanical components.

REFERENCES

1. Haardt R., Godet M. Axial vibration of a misaligned radial face seal under a constant closure force. ASLE Trans, 1975;18: 55-61. https://doi.org/10.1080/05698197508982747
2. Etsion I. Squeeze effects in radial face seals. J. Lubric. Technol., 1980; 102: 145-152. https://doi.org/10.1115/1.3251452
3. Sharoni A., Etsion I. Performance of end-face seals with diametral tilt and coning-hydrodynamic effects. ASLE Trans, 1978; 24: 61-70. https://doi.org/10.1080/05698198108982998
4. Lebeck A. O., Teale J. L., Pierce R. E. Hydrodynamic lubrication and wear in wavy contacting face seals. J.Lubric. Technol, 1978; 100: 81-91. https://doi.org/10.1115/1.3453120
5. Ruddy A. V., Dowson D., Taylor C. M. The prediction of film thickness in a mechanical face seal with circumferential waviness on both the face and the seat. Journal of Mechanical Engineering Science, 1982; 24(1): 37-43. https://doi.org/10.1243/JMES_JOUR_1982_024_008_02
6. Key W. E., Salant R. F., Payvar P., Gopalakrishnan S., Vaghasia G. Analysis of a Mechanical Seal with Deep Hydropads. Tribology Transactions, 1989; 32(4): 481–489. https://doi.org/10.1080/10402008908981916
7. Nakano M, Korenaga A, Korenaga A, et al. Applying micro-texture to cast iron surfaces to reduce the friction coefficient under lubricated conditions. Tribol. Lett. 2007; 28: 131–137. https://doi.org/10.1007/s11249-007-9257-2.
8. Etsion I. State of the art in laser surface texturing. J. Tribol. 2005;127:248–253. https://doi.org/10.1115/1.1828070.
9. Voevodin A.A, Zabinski J.S. Laser surface texturing for adaptive solid

lubrication. Wear. 2006;261(11–12):1285–1292.
https://doi.org/10.1016/j.wear.2006.03.013.

10. Wang X, Adachi K, Otsuka K, Kato K. Optimization of the surface texture for silicon carbide sliding in water. Appl. Surf. Sci. 2006;253(3):1282–1286.
https://doi.org/10.1016/j.apsusc.2006.01.076.

11. Yan D, Qu N, Li H, Wang X. Significance of dimple parameters on the friction of sliding surfaces investigated by orthogonal experiments. Tribol. Trans. 2010;53:703–712.
https://doi.org/10.1080/10402001003728889.

12. Etsion I, Burstein L. A model for mechanical seals with regular micro-surface structure. Tribol. Trans. 1996;39:677-683.
https://doi.org/10.1080/10402009608983582.

13. Etsion I, Halperin G, Greenberg Y. Increasing mechanical seal life with laser-textured seal faces. In Proc. of 15th international conference on fluid sealing BHR group, Maastricht. 1997;3-11.

14. Etsion I, Kligerman Y, Halperin G. Analytical and experimental investigation of laser-textured mechanical seal faces. Tribol. Trans. 1999;42:511-516. https://doi.org/10.1080/10402009908982248.

15. Kligerman Y, Etsion I. Analysis of the hydrodynamic effects in a surface textured circumferential gas seal. Tribol. Trans. 2001;44(3):472-478. https://doi.org/10.1080/10402000108982483.

16. Etsion I, Halperin G. A laser surface textured hydrostatic mechanical seal. Tribol. Trans. 2002;45(3):430-434.
https://doi.org/10.1080/10402000208982570.

17. Wang X.L, Hsu S.M. Integrated surface technology for friction control: A new paradigm effects of geometric shapes on friction. The 4th china international symposium on tribology. 2004;12-20.

18. Yu H, Wang X, Zhou F. Geometric shape effects of surface texture on the generation of hydrodynamic pressure between conformal contacting surfaces. Tribol. Lett. 2010;37:123-130.
https://doi.org/10.1007/s11249-009-9497-4

19. Bai S, Peng X, Li Y, Sheng S. A hydrodynamic laser surface-textured gas mechanical face seal. Tribol. Lett. 2010;38(2):187-194.
https://doi.org/10.1007/s11249-010-9589-1.

20. Yu H, Deng H, Haung W, Wang X. The effect of dimple shapes on friction of parallel surfaces. Proceedings of the institution of mechanical engineers, Part J: Journal of engineering tribology. 2011;225(8):693-703. https://doi.org/10.1177/1350650111406045.

21. Qiu M, Delic A, Raeymaekers B. The effect of texture Shape on the load-carrying capacity of gas-lubricated parallel slider bearings. Tribol. Lett. 2012;48:315–327.
https://doi.org/10.1007/s11249-012-0027-4.

22. Qiu M, Minson B, Raeymaekers B. The effect of texture shape on the friction coefficient and stiffness of gas-lubricated parallel slider bearings. Tribology International. 2013;67:278-288.
https://doi.org/10.1016/j.triboint.2013.08.004.

23. Raeymaekers B, Etsion I, Talke F.E. A model for the magnetic tape/guide interface with laser surface texturing. In proceedings of the ASME/STLE international joint tribology conference. 2008;669–671. https://doi.org/10.1115/IJTC2007-44173.

24. Han J, Fang L, Sun J, Ge S. Hydrodynamic lubrication of microdimple textured surface using three-dimensional CFD. Tribology Transactions. 2010;53:6:860-870.
https://doi.org/10.1080/10402004.2010.496070.

25. Han J, Fang L, Sun J, Wang Y, Ge S, Zhu H. Hydrodynamic lubrication of surfaces with asymmetric microdimple. Tribology Transactions. 2011;54:4:607-615.
https://doi.org/10.1080/10402004.2011.584364.

26. Liu W, Ni H, Chen H, Wang P. Numerical simulation and experimental investigation on tribological performance of micro-dimples textured surface under hydrodynamic lubrication. International Journal of Mechanical Sciences. 2019;163:105095.
https://doi.org/10.1016/j.ijmecsci.2019.105095.

27. Wei Y, Tomkowski R, Archenti A. Numerical study of the influence of geometric features of dimple texture on hydrodynamic pressure generation. Metals, 2020;10:361. https://doi.org/10.3390/met10030361.

28. Gangadia H, Sheth S. Influence of the Bowtie Shaped Dimples on the Performance of Sliding Surfaces under Hydrodynamic Lubrication. Tribologia - Finnish Journal of Tribology. 2023; 40(1−2): 46−58.
https://doi.org/10.30678/fjt.126885

29. Gangadia H, Sheth S. The Effect of Star Shaped Dimples on Sliding Surfaces under Hydrodynamic Lubrication, Tribology Online. 2023; 18(7): 457-468. https://doi.org/10.2474/trol.18.457

30. ANSYS Fluent Theory Guide, 2013.

31. Bakir F, Rey R, Gerber A. G, Belamri T, Hutchinson B. Numerical and experimental investigations of the cavitating behaviour of an inducer. Int. J. Totat. Mech. 2004; 10(1): 15-25.
https://doi.org/ 10.1080/10236210490258034

32. Cupillard S, Glavatskih S, Cervantes M. J. Computational fluid dynamics analysis of a journal bearing with surface texturing. Proceedings of the Institution of Mechanical Engineers. Part J: Journal of Engineering Tribology. 2008; 222(2): 97-107.
https://doi.org/10.1243/13506501JET319

Hardik Gangadia: https://orcid.org/0000-0002-7734-3402

Saurin Sheth: https://orcid.org/0000-0002-1792-7051

MODELLING OF AN INFLUENCE OF LIQUID VELOCITY ABOVE THE NEEDLE
ON THE BUBBLE DEPARTURES PROCESS

Paweł DZIENIS*

*Faculty of Mechanical Engineering, Białystok University of Technology, ul. Wiejska 45C, 15-351 Białystok, Poland

p.dzienis@pb.edu.pl

Abstract: In the present paper, the influence of liquid flow above the needle on a periodic or chaotic nature of the bubble departures process was numerically investigated. During the numerical simulations bubbles departing from the needle was considered. The perturbations of liquid flow were simulated based on the results of experimental investigations described in the paper [1]. The numerical model contains a bubble growth process and a liquid penetration into a needle process. In order to identify the influence of liquid flow above the needle on a periodic or chaotic nature of bubble departures process, the methods data analysis: wavelet decomposition and FFT were used. It can be inferred that the bubble departure process can be regulated by altering the hydrodynamic conditions above the needle, as variations in the liquid velocity in this area affect the gas supply system's conditions. Moreover, the results of numerical investigations were compared with the results of experimental investigation which are described in the paper [2]. It can be considered that, described in this paper, the numerical model can be used to study the interaction between the bubbles and the needle system for supplying gas during the bubble departures from two needles, because the interaction between the bubbles is related to disturbances in the liquid flow above the needle.

Keywords: bubble, liquid penetration into the needle, bubbles interaction

1. INTRODUCTION

The gas bubble flow and gas bubble formation process in liquids are found in: oceans, chemical processes, pharmaceutical industries or industrial equipment [3]. In medicine, ultrafine gas bubbles are used to transport medications [4]. The studies on bubble departure and interactions between bubbles serve as a preliminary exploration into the greenhouse effect caused by methane bubbles escaping in the oceans [5,6]. The gas bubble flow is investigated during the exploitation of the hydrocarbon deposits, too [7]. Moreover, the knowledge of bubble flow and bubble departure process is very important during the aeration or saturation process which helps e.g. in purifying the surface water or municipal sewage [8]. Regulating the bubble departure process can enhance mass transfer in the bubble column. [9,10,11]. The investigations of interaction between bubbles is treated as an introduction to investigations of the bubble formation process during boiling [12,13,14].

There are papers, in which researchers describe their results of investigations of single, double or multi bubble departures. In these investigations bubbles are generated from single [15-20] and twin or more needles or orifices [23-30].

The bubble departure time can be split into bubble growth time and bubble waiting time [15-19]. At relatively low gas flow rates, the liquid penetrates the needle or orifice during the bubble waiting time. The liquid penetration into the needle or orifice is connected with a decrease in gas pressure occurrences in the system for supplying gas [15-19] and it is modified by liquid pressure changes above the needle caused by perturbations in liquid flow in the needle neighbourhood [2]. In other papers [19-21], the chaotic nature of bubble departures process, phenomena which

influence chaotic bubble departures or bubbles trajectories were investigated. In the paper [20] it was demonstrated that the chaotic nature of bubble trajectories is due to the shape of the departing bubbles and the liquid flow induced by the moving bubbles within the bubble column. In papers we can find that two groups of phenomena are responsible for the chaotic nature of bubble behaviours: the first group is connected with the bubble interface oscillations, liquid flow around the needle or orifice and the second group is connected with the processes which appear in the gas supply system of the needle [16,21,22].

The process of bubble departures from due or more needles or orifices was investigated in papers [23-31]. In papers [23,24] it was shown that interactions between bubbles, departed from twin needles can lead to synchronous or alternative bubble departures, bubble coalescence or bubble bouncing. The bubble coalescence or bubble bouncing depend on the bubble Reynolds numbers. The impact of the distance between needles on bubble interactions in selected kinds of liquids (with different physical and chemical properties) was investigated in paper [25]. In the paper [2] the interaction between bubbles or bubbles and the system for supplying gas was investigated based on the process of liquid penetration into the needle. It has been demonstrated that hydrodynamic interactions can result in periodic, chaotic or multi-period changes of liquid movement inside the needle gas supply system. Those changes influence the nature of the bubble departures process [2]. In the paper [25] the regimes of alternative bubble departures (ABD coefficient) in different kinds of liquids are proposed. The coefficient ABD comprises the distance between needles, air volume flow rate, liquid properties, and the frequency of bubble departures. The regimes of synchronous bubble departures from twin orifices are investigated in the paper [26]. Fur-

thermore, the ABD coefficient was examined for orifices in this study [26]. Additionally, research on bubble interactions was conducted in papers [26-29]. In another study [20], bubbles were generated in water and an aqueous glycerine solution, revealing that such interactions alter bubble trajectories. The study concluded that bubble interactions and coalescence depend on gas flow rates, tube spacing, and liquid properties. The interaction of a bubble pair in viscoelastic shear-thinning fluids was experimentally investigated in paper [33]. It was suggested the elasticity and deformability are responsible for interaction between the bubble pair like in non-Newtonian fluids.

The numerical models of bubble departures (for example, proposed in the papers [31,32]) are very sensitive to changes in boundary conditions and modifications to systems of differential equations. Consequently, modelling the interaction between bubbles departed from twin needles or the influence of bubble departure from one needle on the process of bubbles departing from neighbouring needles is very difficult or impossible. Based on results of experimental investigation of liquid changes caused by departed bubbles (presented in the paper [1]), in present paper the modelling of the perturbations of bubble departures process caused by liquid flow above the needle is proposed. Proposed model can be used to investigate bubble and gas supply system interaction. The results of numerical investigation are consistent with the results of experimental investigation described in paper [2], where the perturbations of liquid flow above the needle are caused by growing and moving bubbles.

In the present paper the influence of liquid flow above the needle (according to the modifications caused by departed and moving bubbles) on the bubble departure process was numerically investigated. For modelling the interaction the model of the bubble growth and liquid movement inside the needle, which has been proposed in paper [32] was used. The model was modified so that it was possible to change the liquid inflow frequency and speed of liquid flow above the needle from which the bubbles are generated. In order to identify the influence of liquid flow above the needle on the bubble departure process, the wavelet decomposition and FFT methods were used. To investigate how liquid flow interactions affect disturbances in the gas supply system, manifesting as fluctuations in the depth of liquid penetration into the needle, 3D attractors were reconstructed. The study demonstrated that regulating the bubble departure process can be achieved by adjusting the hydrodynamic conditions above the needle. This is because variations in liquid velocity above the needle directly impact the conditions within the gas supply system. Described in this paper, the numerical model can be used to study the interaction between the bubbles and the needle gas supply systems during the bubble departures from two needles, because the interaction between the bubbles is related to disturbances in the liquid flow above the needle.

The structure of the paper is as follows. The numerical model is described in Chapter 2. Results of the numerical data analysis are shown in Chapter 3 -"Results of numerical investigations". A summary of the obtained results is shown in the "Conclusion" section.

2. DESCRIPTION OF NUMERICAL MODEL

For modelling the influence of liquid flow perturbations above the needle on the nature of the bubble departure process, the model of the bubble growth and liquid movement inside the nee-

dle was used. The model was proposed in papers [29,30], but it was modified so that it was possible to change the liquid inflow frequency and speed of liquid flow above the needle from which the bubbles are generated. These liquid flow perturbations correspond to the hydrodynamic interaction between the bubbles and the needle's gas supply systems. In the considered numerical model, attempts were made to maintain the same hydrodynamic conditions as in the experimental studies presented in paper [2]. The model was prepared in the SciLab environment, in which the equations described below were solved using the function ODE. The schema of bubble growth and liquid movement into the needle model is shown in Fig.1.
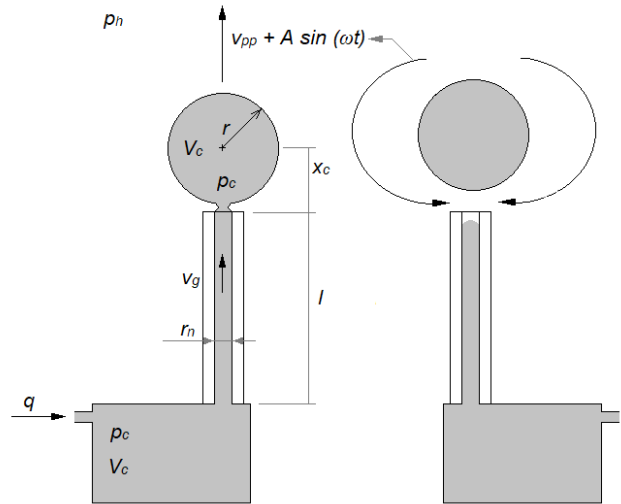
Fig. 1. The schema of bubble growth and liquid movement into the needle model

During the bubble growing stage an isothermal process was considered while the bubble growth was described by the Rayleigh–Plesset equation [30]:

$$r_b \left(\frac{d^2 r_b}{dt^2} \right) - \frac{3}{2} \left(\frac{dr_b}{dt} \right)^2 = \frac{1}{\rho_l} \left(p_b - p_h - \frac{2\sigma}{r_n} - \frac{4\mu_l dr_b}{dt} \right) \quad (1)$$

where: r_b is the radius of the bubble (m), ρ_l is the liquid density (kg/m³), p_b is the air pressure in bubble (Pa), p_h is the hydrostatic pressure (Pa), σ - the surface tension (N/m), r_n is the inner radius of the needle (m), μ_l is the dynamic viscosity of the liquid (kg/ms).

The air volume flow rate supplied to the bubble through the needle was determined by the Hagen–Poiseuille equation:

$$\frac{dp_b}{dt} = \left(\frac{p_b}{V_b} \right) \left[\left(\frac{\pi}{8\mu_g} \right) \left(\frac{r_n^4}{l} \right) (p_c - p_b) - \frac{dV_b}{dt} \right] \quad (2)$$

where: V_b is the volume of bubble (m³), μ_g is the dynamic viscosity of the gas (kg/ms), l is the needle length (m), p_c is the air pressure in the gas supply system (Pa).

Pressure changes in the air supply system are described by the following equation:

$$\frac{dp_c}{dt} = \frac{k_c p_c}{V_c} \left[q - \frac{\pi r_n^4}{8\mu_g l} (p_c - p_b) \right] \quad (3)$$

where: V_c is the gas supply system volume (m³), q is the gas volume flow rate supply to the needle (m³/s).

Moreover, the model of the bubble growth stage accounts the forces acting on a growing bubble [30]:

– drag force:

$$F_d = 0.5 C_d \rho_l \pi r^2 \left(\frac{dx_c}{dt} - v_{pp}\right) \left|\frac{dx_c}{dt} - v_{pp}\right| \qquad (4)$$

– buoyancy force:

$$F_B = g(\rho_l - \rho_g) \cdot V_b \qquad (5)$$

– maximum value of the surface tension force:

$$F_\sigma = 2\pi \, r_n \sigma \qquad (6)$$

– added mass force:

$$F_{AM} = -\rho_l \frac{d}{dt}\left[C_M V_b \left(\frac{dx_c}{dt} - v_{pp}\right)\right] \qquad (7)$$

– gas momentum:

$$F_M = \rho_g \frac{q_b^2}{\pi r_o^2} \; where \; q_b = \left(\frac{\pi}{8\mu_g}\right)\left(\frac{r_n^4}{l}\right)(p - p_b) \qquad (8)$$

where: g is the gravitational acceleration (m/s^2), r_o is the bubble diameter (m), q_b is the air flow rate supplied to the bubble (m^3/s), Cd is the drag force coefficient, CM = 0.5 is the added-mass coefficient for a sphere, v_{pp} is the velocity of the liquid around the growing bubble (m/s) and x_c is the position of the bubble centre (m).

The liquid movement inside the needle is described by the equation of motion of the liquid mass centre [32]:

$$\frac{d}{dt}\left\{\left[0.5\rho_l \pi r_n^2 x_l + \rho_l \frac{4}{3}\pi(2r_n)^3\right]\frac{dx_l}{dt}\right\} = F_1 - F_2 \qquad (9)$$

The force $F1$ is related to the pressure difference that occurs in the system:

$$F_1 = -s\Delta p = -\pi r_n^2 \left[p_g - \left(p_h + \rho_l g(2x_l) + A \cdot 2\frac{\sigma}{r_n} - \frac{\rho v_{pp}|v_{pp}|}{2}\right)\right] \qquad (10)$$

The force $F2$ is related to the resistance of the movement of the liquid in the needle:

$$F_2 = 2 * 8\pi\mu_l x_l \frac{dx_l}{dt} \qquad (11)$$

The corresponding pressure changes in the air supply system are described by the following equation [32]:

$$\frac{dp_c}{dt} = \frac{p_c}{V_c}\left(q + \pi r_n^2 \frac{dx_l}{dt}\right) \qquad (12)$$

where: x_l is the height of the liquid penetration into the needle (m) and s is the cross-sectional area of the needle (m^2).

Criterion for the end of the liquid movement is correlated with the depth of liquid penetration – $x_l < 0$.

Based on the above equations, it can be assumed that, the perturbations in liquid flow velocity above the needle affect the bubble waiting time (v_{pp} – Eq.10) or bubble growing time (Eq.4 and Eq.7) and consequently the perturbations in liquid flow velocity can modify the nature of periodic or chaotic bubble departure process. The character of changes of liquid flow above the needle is the same as liquid pressure changes above the needle. Therefore, based on the results of experimental investigations of liquid pressure changes modification above the needle cause by bubble departure process (being a periodic function), presented in paper [1] (Fig.2), the variations in liquid flow were approximated using a sinusoidal function, and based on equations (Eq.9, Eq.11, Eq.14), the velocity of the liquid above the needle was computed as:

$$v_{pp} = v'_{pp} + A sin(\omega t) \qquad (13)$$

V_{pp}' is the liquid flow above the needle induced by bubbles depart-

ing from it, focusing on the study of liquid movement in this process and $Asin(\omega t)$ is disturbances in the liquid velocity.

The changes of pressure above the needle, during the bubble departures (Fig.2).
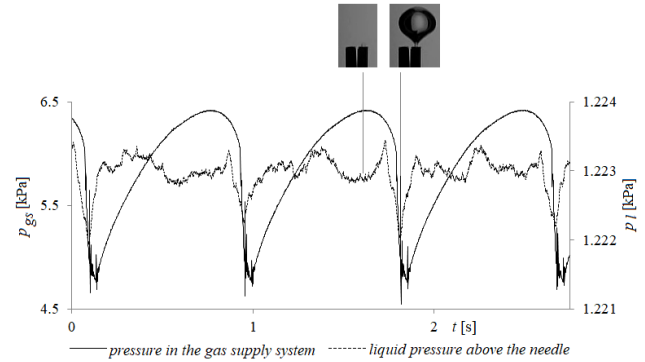
Fig. 2 The changes of liquid pressure above the needle, during the bubble departures [1]

In the paper [1] it was shown that subsequent stages of bubble growth change the hydrodynamic conditions around the needle. The measured fluctuations in the liquid pressure above the needle, caused by bubble departures, are periodic. The greatest pressure changes are observed during the bubble growing time (Fig.2). During the bubble waiting time only slight fluctuations in liquid pressure above the needle are observed.

3. RESULTS OF NUMERICAL INVESTIGATIONS

In numerical simulations, the time series of liquid penetration into the needle during successive cycles of bubble departure, and the time series of liquid flow velocity above the needle, were analyzed. The simulations aimed to capture periodic and chaotic behaviors in the time series, depending on the frequency of variations in liquid flow velocity above the needle. Fig. 3 illustrates the time series of liquid penetration into the needle and disturbances in liquid velocity.

The model exhibits high sensitivity to boundary conditions; therefore the initial minimum value of pressure in the system for supplying gas was set at 5.4 kPa. The air volume flow rate was set as 0.006 l/min. The amplitude (A) was set as 0.1 and it simulates the velocity of liquid flow perturbations. Frequency of liquid perturbations was modified by changes in ωt and frequencies were changed during the simulations. For presented considerations the frequency was equal to 10.02 Hz, 11.52Hz and 11.9 Hz.

In order to investigate the repeatability of time series of liquid penetration into the needle, the 3D attractors were used. The 3D attractor is reconstructed using the stroboscope coordination. In this method, The coordinates of attractor points are determined by computing the positions based on sampled data points where the distance between them equals the time delay (τ. If the subsequent trajectories on the attractor are closely spaced, the signal is considered quasi-periodic. However, if the trajectories in the attractor reconstruction begin to diverge from each other, it indicates that the analyzed signal exhibits chaotic behavior. The reconstruction of the 3D attractor, the time delay (τ) was calculated for all time series separately. To determine τ, the mutual information method was used [32-34]. The first minimum of the function is treated as

the proper value of *τ*:

$$I(x_i, x_{i+\tau}) = \sum_{x_{i+\tau}} \sum_{x_i} p[x_i, x_{i+\tau}] \log_2 \left\{ \frac{p[x_i, x_{i+\tau}]}{p[x_i], p[x_{i+\tau}]} \right\} \quad (14)$$

where: $p[x_i, x_{i+\tau}]$ is the joint probability function of $\{x_i\}$ and $\{x_{i+\tau}\}$, $p[x_i]$ and which are the marginal probability distribution functions of $\{x_i\}$ and $\{x_{i+\tau}\}$.
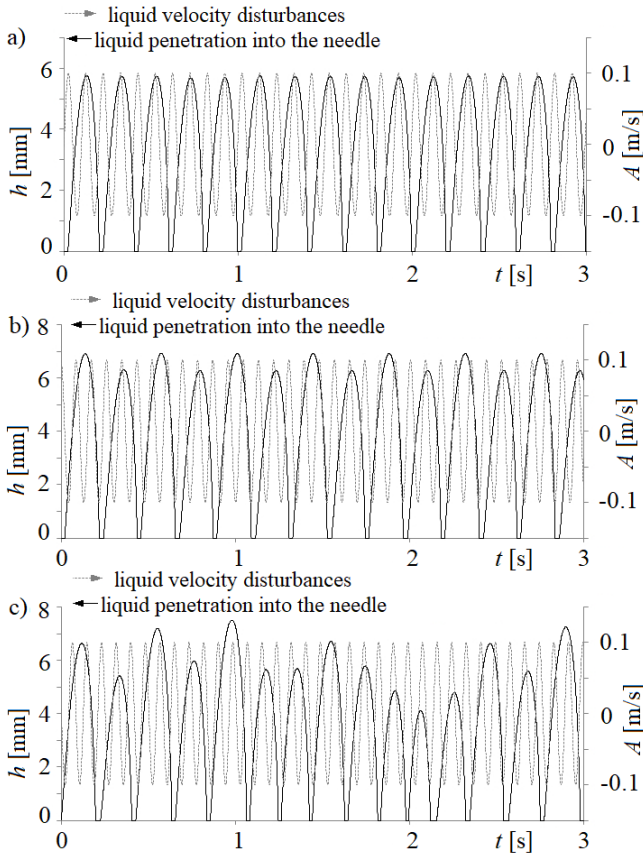
Fig. 3. The time series of liquid penetration into the needle obtained from numerical investigations (continuous lines) and liquid velocity perturbations (dotted lines) for selected frequencies of liquid velocity disturbances (f_{sin}). a) f_{sin} = 10.02 Hz, b) f_{sin} = 11.52 Hz, c) f_{sin} = 11.9 Hz

In Fig. 4 are shown the reconstructions of 3D attractors for time series of liquid penetration into the needle.
The changes of frequency of occurrence of perturbations in liquid flow above the needle modify the depths of liquid flooding into the needle in successive cycles of bubble departures (Fig.3) and the periodic or chaotic nature of bubble departures (Fig.4). In Fig.3 a the depths of liquid flooding are similar in successive cycles of bubble departures. In this case the trajectories in the 3D attractor reconstruction are close to each other for liquid flooding in subsequent stages of bubble departures (Fig.4 a). In Fig.3 b it can be observed that the changes of the depth of liquid flooding into the needle for successive cycles of bubble departures occurs with two periods. In the first cycle the depth is close to 6 mm and for the second cycle it is close to 7 mm and this process is repeatable in all analysed time series. Two distinct paths overlap in the 3D reconstruction of the attractor (Fig. 4b). As shown in Fig. 3c, the penetration of liquid into the right needle varied unpredictably. The trajectories that compose the 3D attractors are non-repetitive (Fig. 4c).

To verify the timing of bubble departures and occurrence of perturbations of liquid flow above the needle, the frequencies of these phenomena occurrences were estimated. The frequencies were estimated using the FFT method [35,36]. The frequencies of the liquid penetration into the needle and the frequencies of liquid velocity perturbations are shown in Table 2.
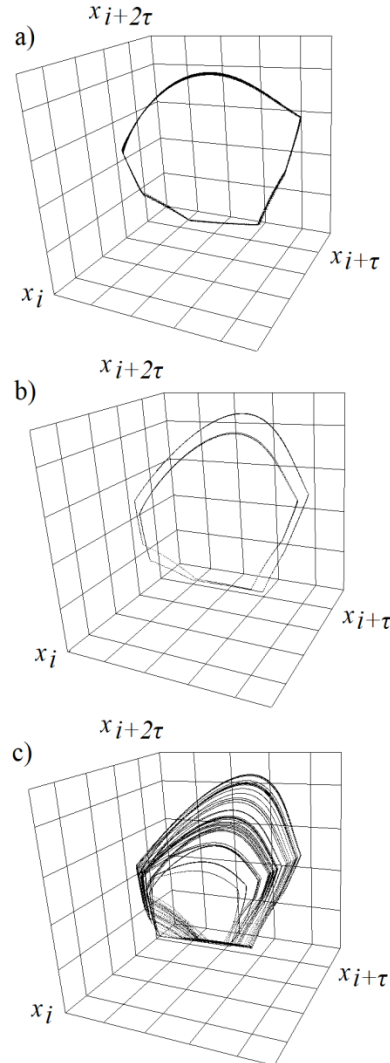
Fig. 4. 3D attractors obtained for time series of liquid flow into the needle for selected frequencies of liquid velocity perturbations (f_{sin}). a) f_{sin} = 10.02 Hz, b) f_{sin} = 11.52 Hz, c) f_{sin} = 11.9 Hz

Tab. 2. The frequencies of the liquid penetration into the needle and the frequencies of liquid velocity perturbations

f_{lp} [Hz]	5.01	4.58	4.64
f_{sin} [Hz]	10.02	11.52	11.90
f_{lp}/f_{sin} [-]	2.00	2.51	2.56

In the FFT method, the dominant frequency of liquid movement into the needle can be treated as a frequency of bubble departures, but only for periodic changes of depth of liquid penetration into the needle. In the case of chaotic changes of depth of liquid penetration into the needle the dominant frequency is treated as the frequency of the majority of departing bubbles in the analysed time series. Even though this frequency is not fully a

frequency of bubble departure, analysis of the dominant frequency can be used for investigations of liquid flow above the needle on bubble departure nature.

To explore how changes in the frequency of liquid velocity above the needle relate to the time series of liquid penetration into the needle, wavelet decomposition analysis and the FFT method were employed. This analysis was performed in Matlab with the Wavelet Toolbox. The Daubechies (db2) method in the Orthogonal Wavelet Family, and I performed five levels of frequency decomposition because the analysed time series have the non-linear character. The signal of details obtained from the 5th level of decomposition was analysed using the Fast Fourier Transform. The power spectra are shown in Fig. 5.

The disturbance frequency of the liquid flow above the needle was marked using fs in Fig.5. In Fig.5 it was shown that liquid flow disturbances modify the nature of the frequency distribution in the analysed time series. If bubble detachment is periodic, then only multiples of the frequency components are observed (Fig.5 a and Fig.5 b). The introduction of non-frequency-synchronized disturbances results in the appearance of multiple frequency components, which determines the chaotic nature of the bubble departure process.
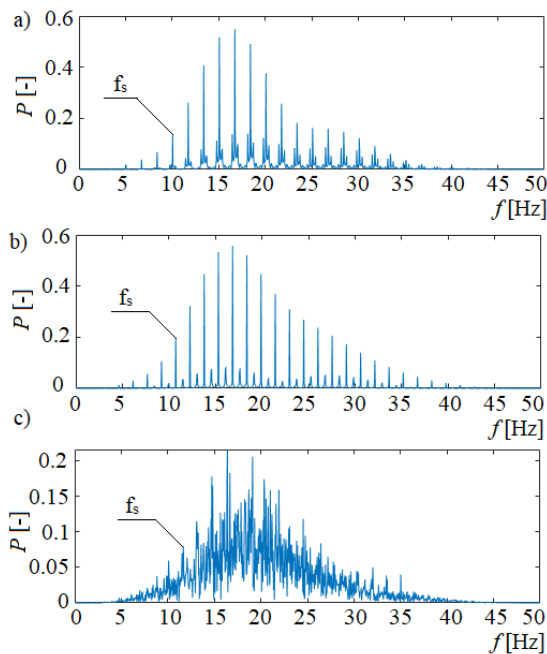
Fig. 5. Frequency components obtained from 5th level wavelet decomposition for time series of liquid flow into the needle for selected frequencies of liquid velocity perturbations (f_{sin}). a) f_{sin} = 10.02 Hz, b) f_{sin} = 11.52 Hz, c) f_{sin} = 11.9 Hz.

For presented considerations the frequencies of disturbances of liquid velocity above the needle (f_{sin}) were chosen so that the ratio of those frequencies and the frequencies of bubbles departures (f_{lp}) were: was different from the integer, and the close to half-integer (equal to 2.5) integer (equal to 2). It was shown that synchronisation in the perturbations in liquid flow above the needle and the bubble departures can modify the nature of bubble departures. In the case, when the ratio of f_{sin}/f_{lp} is integer then the depth of liquid flow into the needle changes slightly for successive cycles of bubble departures (Fig.3 a). The trajectories in 3D attractors are close to each other (Fig.4 a). In the case that the ratio of f_{sin}/f_{lp} is integer and half (2.5) then the needle is flooded by the

liquid with two consecutive levels of depths of liquid penetration in successive cycles of bubble departures (Fig 3 b). When the ratio of f_{sin}/f_{lp} is other than an integer, then the depth of liquid flooding is varying in the successive cycles of bubble departures.

It can be concluded that the disturbances in liquid flow above the needle modify the conditions in the system supplying gas. Disturbances occurring in liquid flow lead to changes in the depth of liquid penetration and the frequency of bubble detachment. This confirms that adjusting the hydrodynamic conditions above the needle can regulate the process of bubble departure. The results obtained from the numerical model are consistent with the results of the experiment presented in paper [2]. The proposed numerical model containing a component allowing for modification of the velocity and frequency of disturbances in the liquid flow above the needle can be used to numerically study the interaction between the bubbles and the needle the system supplying gas during the bubble departures from two needles.

4. CONCLUSIONS

This paper numerically investigates the influence of liquid flow above the needle (modified by departed bubbles) on the bubble departure process. For modelling the interaction the models of the bubble growth and liquid penetration into the needle were used. The numerical model containing the component allows for modification of the velocity and frequency of disturbances in the liquid flow above the needle. This component was developed based on the experimental results presented in this paper [1]. It was shown that the numerical model can be used to study the interaction between the bubbles and the needle gas supply systems during the bubble departures from two needles, because the interaction between the bubbles is related to disturbances in the liquid flow above the needle.

Moreover it was shown that the bubble departure process can be control by adjusting the hydrodynamic conditions above the needle. This is because alterations in the liquid flow and its velocity above the needle affect the boundary conditions in the system supplying gas. The changes in the conditions in the gas supply system correspond with the changes of the periodic or chaotic nature of bubbles departures (visible in time series of liquid penetration into the needle). Consequently, the depth of liquid into the needle and the frequency of bubble departures vary, depending on the occurring disturbances.

REFERENCES

1. Dzienis P, Mosdorf R. Liquid pressure fluctuations around a needle during bubble departures. Meccanica. 2023;58:1307-1313.
2. Dzienis P, Golak K, Konopka M, Mosdorf R, Baziene K, Gargasas J. A hydrodynamic interaction between bubbles and gas supply system during gas bubble departures in liquids: an experimental study. Scientific Reports. 2023;13:17979.
3. He Y, Zhang T, Lv L, Tang W, Wang Y, Zhou J, Tang S. Application of microbubbles in chemistry, wastewater treatment, medicine, cosmetics, and agriculture: a review. Environmental Chemical Letters.2023;21:3245–3271.
4. Le TH, Phan AHT, Le KCM, Phan TDU, Nguyen KT.Utilizing polymer-conjugate albumin-based ultrafine gas bubbles in combination with ultra-high frequency radiations in drug transportation and delivery. RSC Advances.2021;11(55):34440–34448.
5. Leifer I, Tang D. The acoustic signature of marine seep bubbles. The Journal of the Acoustical Society of America.2007;121:35–40.

6. Boufadel MC, Socolofsky S, Katz J, Yang D, Daskiran C, Dewar W. A review on multiphase underwater jets and plumes: Droplets, hydrodynamics, and chemistry, Reviews of Geophysics. 2023; 58:1–40.
7. Abdulmouti H. Bubbly two-phase flow: III- applications, American Journal of Fluid Dynamics.2022;12(1):16–119.
8. Gevod VS, Borisov IA. Influence of air bubble flow structure on the rate of water purification by the bubble-film extraction method, Water Supply 2019;19(8):2298–2308.
9. Budzyński P, Gwiazda A, Dziubiński M. Intensification of mass transfer in a pulsed bubble column, Chemical Engineering Processing & Process Intensification. 2017;112:18-30.
10. Suwartha N, Syamzida D, Priadi CR, Moersidik SS, Ali F. Effect of size variation on microbubble mass transfer coefficient in flotation and aeration processes. Heliyon.2020;6(4):e03748.
11. Herrmann-Heber R, Reinecke SF, Hampel U. Dynamic aeration for improved oxygen mass transfer in the wastewater treatment process. Chemical Engineering Journal.2020;386:122068.
12. Pan F, Mu L, He Y, Wang C, Zhou S. A thermal-hydrodynamic coupling method for simulating the interplay between bubble departure and wall temperature variation in nucleate boiling, Journal of Hydrodynamics. 2021; 33: 243–258.
13. Yuan J, Ye X, Shan Y. Modeling of the bubble dynamics and heat flux variations during lateral coalescence of bubbles in nucleate pool boiling. International Journal of Multiphase Flow 2021;142:103701.
14. Pan F, Mu L, He Y. Wang C. Numerical study on the activation of nucleation sites and bubble interactions in twin-bubble nucleate boiling, Proceedings of the Institution of Mechanical Engineers. Part C: Journal of Mechanical Engineering Science. 2022;236(12). doi.org/10.1177/09544062211065987
15. Stanovsky P, Ruzicka M, Martins A, Teixeira J. Meniscus dynamics in bubble formation: A parametric study. Chemical Engineering Science. 2011;66:3258-3267.
16. Dzienis P, Zaborowska I, Mosdorf R. JRP analysis of synchronization loss between signals recording during bubble departures. Nonlinear Dynamics. 2022;108:433–444.
17. Mohseni E, Reinecke SF, Hampel U. Controlled bubble formation from an orifice through harmonic gas pressure modulation. Chemical Engineering Journal 2023. doi.org/10.1016/j.cej.2023.143953
18. Cano-Lozano JC, Bolaños-Jiménez R, Gutiérrez-Montes C. Martínez-Bazán C. On the bubble formation under mixed injection conditions from a vertical needle, International Journal of Multiphase Flows. 2017; 97:23–32.
19. Dzienis P, Mosdorf R, Wyszkowski T. The dynamics of liquid movement inside the nozzle during the bubble departures for low air volume flow rate. Acta Mechanica et Automatica.2012;6(3):31–36.
20. Augustyniak J, Perkowski DM, Mosdorf R. Measurement of multifractal character of bubble paths using image analysis, International Communication in Heat and Mass Transfer 2020;117: 104701.
21. Zhang L, Shoji M. Aperiodic bubble formation from a submerged orifice, Chemical Engineering Science. 2001; 56: 5371–5381.
22. Snabre P, Magnifotcham F. Formation and rise of a bubble stream in viscous liquid The European Physical Journal B. 1997; 4: 369–377.
23. Feng X, Kunugi T, Qin S, Wu D. Flowrate effects on the lateral coalescence of two growing bubbles. The Canadian Journal of Chemical Engineering. 2023. doi.org/10.1002/cjce.24976
24. Yuan J, Ye X, Shan Y. Modelling of double bubbles coalescence behaviour on different wettability walls using LBM method, International Journal of Thermal Sciences. 2021; 168: 107037.
25. Dzienis P, Mosdorf R, Wyszkowski T. A hydrodynamic criterion of alternative bubble departures, Thermal Science. 2021; 25(1B): 553-565.
26. Capponi A, Llewellin EW. Experimental observations of bubbling regimes at in-line multi-orifice bubblers. International Journal of Multiphase Flow. 2019; 114: 66–81.
27. Padash A, Chen B, Boyce CM. Characterizing alternating bubbles emerging from two interacting vertical gas jets in a liquid. Chemical Engineering Science.2022; 248 Part B: 117199.
28. Kazakis NA, Mouza AA, Paras SV. Coalescence during bubble formation at two neighbouring pores: an experimental study in microscopic scale, Chemical Engineering Science. 2008;63(21): 5160–5178.
29. Legendre D, Magnaudet J, Mougin G. Hydrodynamic interactions between two spherical bubbles rising side by side in a viscous liquid, Journal of Fluid Mechanics. 2003; 497: 133–166.
30. Sanada T, Sato A, Shirota MT, Watanabe M. Motion and coalescence of a pair of bubbles rising side by side. Chemical Engineering Science. 2009; 64: 2659–2671.
31. Ruzicka MC, Bunganic R, Drahoš J. Meniscus dynamics in bubble formation. Part II: Model. Chemical Engineering Research and Design. 2009; 87: 1357-1365.
32. Dzienis P, Mosdorf R. Stability of periodic bubble departures at a low frequency. Chemical Engineering Science. 2014; 109: 171-182.
33. Ravisankar M, Garcidueñas Correa A, Su Y, Zenit R. Hydrodynamic interaction of a bubble pair in viscoelastic shear-thinning fluids. Journal of Non-Newtonian Fluid Mechanics. 2022;309:104912.
34. Marwan N, Romano MC, Thiel M, Kurths J. Recurrence plots for the analysis of complex systems. Physics Reports. 2007;438: 237–329.
35. Grassberger P, Procaccia I. Measuring the strangeness of strange attractors. Physica – D. 1983;9:189–208.
36. Sen AK, Litak G, Taccani R, Radu R. Chaotic vibrations in a regenerative cutting process, Chaos Solitons Fractals. 2008;38:886–893.
37. Schuster HG. Deterministic Chaos. An Introduction. PWN. Warszawa 1993 (in Polish).
38. Liebert W, Schuster HG. Proper choice of the time delay for the analysis of chaotic time series. Physics Letters A. 1989;142:107–111

Paweł Dzienis: https://orcid.org/0000-0001-9200-8760

MODELLING AND EXPERIMENTAL STUDY
OF A PASSIVE FREQUENCY-DEPENDENT VEHICLE SUSPENSION DAMPER

Bartłomiej FRANCZYK*⬤, JANUSZ GOŁDASZ**⬤

*Faculty of Mechanical Engineering, Cracow University of Technology, al. Jana Pawła II 37, 31-864 Kraków, Poland
**Faculty of Electrical and Computer Engineering, Cracow University of Technology, ul. Warszawska 24, 31-155 Kraków, Poland

bartlomiej.franczyk@doktorant.pk.edu.pl, jgoldasz@pk.edu.pl

Abstract: The recent trends in the automotive industry have enforced chassis solutions beyond the reach of conventional systems. Thus, extending the functionality of passive hydraulic dampers is vital in improving their effectiveness while maintaining low production and operating costs. This paper presents a general structure of a passive shock absorber with so-called frequency-dependent (FD) damping characteristics and points to constitutive elements of the valves used in this type of an adaptive damper. A mathematical description of FD damper is provided together with a model developed in the Siemens AMESim environment. The performance of the model was verified against the data from tests with a real, commercially available FD shock absorber. Furthermore, in order to emphasise its efficiency, the authors have carried out a study involving quarter car models (QCM) with and without the FD damper, respectively. The results have clearly shown major advantages of utilising FD dampers in a suspension.

Key words: vehicle dynamics, ride, road holding, adaptive damping, quarter car, frequency dependent damping

1. INTRODUCTION

In general, motor vehicles are subjected to constant vibrations of a random nature, the source of which are road irregularities and driving manoeuvres. Vehicle dampers that are used in suspensions systems are aimed at providing sufficient amount of damping, and thus satisfying level of driving comfort as well as good handling and roadholding of a vehicle. However, it is known that these criteria are most often contradictory, and the level of damping in the suspension is the result of a compromise between smooth ride and adequate motion control [1, 2]. The reason for this conflict can be well understood by analysing exemplary transmissibility function of a classic one degree of freedom (1-DOF) harmonic oscillator. As one can see in Fig. 1, an increase in damping coefficient reduces or even eliminates the resonant peak corresponding to the sprung mass natural frequency. On the other hand, higher damping degrades isolation from road inputs at the higher frequency excitations. Thus, high damping is required in the suspension system for frequencies below the invariant point of $\omega_0\sqrt{2}$, while above this frequency low damping would be most beneficial [3]. This conflict is amplified by the fact that the vehicle suspension consists of sprung mass and several unsprung masses, which differ by an order of magnitude. Due to such a significant difference, each of these masses needs different damping forces to be tuned optimally. Wheel vibrations, usually of high frequency but small amplitude, require less damping forces than vibrations of a vehicle body, which are characterised by low frequency and high amplitude [4, 5]. It becomes obvious that for the overall improvement of vehicle dynamics, damping dependent on current operating conditions of the suspension should be used with particular emphasis on operation frequencies.
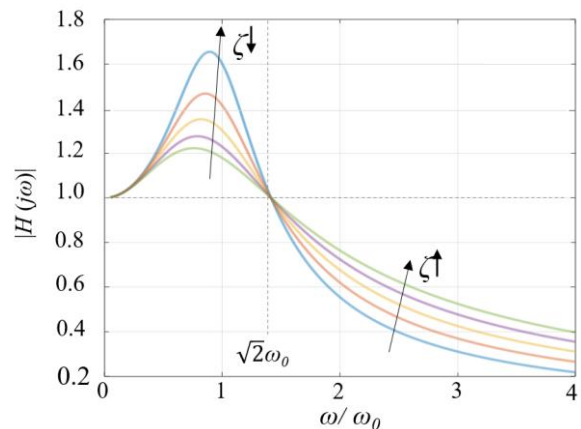
Fig. 1. Displacement transmissibility function of a 1-DOF harmonic oscillator: H(jω) – the ratio of input and output amplitudes, ω – excitation frequency, ω0 – natural frequency, ζ – damping coefficient. 1-DOF, one degree of freedom

One common way of making the damper operation dependent on the excitation frequency (or the excitation change rate in general) is the use of active or semi-active systems utilising electronic controls. Some benefits of such systems have been described in the works of Nguyen et al. [6], Slaski [7] and Pletschen et al. [8]. However, such solutions are relatively expensive and thus unavailable to ordinary car owners. Obviously, far more affordable technologies would rely on passive valving systems configured for the desired functionality. This category includes both displacement-sensitive dampers (DSD) and amplitude-sensitive dampers (ASD) [9, 10]. Notable DSD examples include, for instance, the works of Lee and Moon [11], Hazaveh et al. [12] and Ilbeigi et al. [13]. Such dampers allow for the position dependency of the

damping force in order to increase ride comfort at low amplitude inputs while keeping sufficient body control at high amplitude inputs. It should be noted, however, that such valving systems operate well only when the damper's relative displacement amplitude oscillates around their design positions where the DSD feature are usually located. If the initial position of the suspension changes significantly as in the case of a loaded vehicle, it may exhibit unexpected and disadvantageous behaviours. This shortcoming seems to be avoided when ASD valving systems are used. With these systems, the damping level depends on the amplitude of the suspension movement, regardless of the initial position; low damping is provided for low-amplitude inputs but whenever damper travel exceeds a certain amount, damping increases significantly. The ASD examples are described by Łuczko and Ferdek [14], Łuczko et al. [15], Gołdasz [16] and Zhang et al. [9] . ASD systems, however, are not free of disadvantages, namely, a sudden change in the level of forces related to the full use of the working range of elastic elements involved, increased noise or a large number of additional components [9]. Moreover, due to the time delay needed to reach high damping forces, they can negatively affect the vehicle's steering system response. Some authors, for example, Nie et al. [17] carried out attempts to achieve even more sophisticated damper characteristics and thus bringing the passive damper's performance closer to that of the active one. Xu et al. [18] presented a damper with a variable moment of inertia to improve vehicle chassis dynamic characteristics. Finally, Sikora [19] proposed a tuned mass damper in order to achieve improved dynamic characteristics with no apparent negative impact on the steady-state performance. It must be noted, however, that real implementations of such concepts usually require difficult-to-realise technological solutions.

Based on the above considerations, it can be concluded that the optimal solution to achieve both high levels of comfort (or isolation from road inputs) and driving performance is a suspension damper possessing variable damping characteristics, independent of the suspension position, providing reduced damping forces at high-frequency and low-amplitude inputs, but at the same time delivering 'normal' levels of damping forces at low frequency and high amplitude excitations. These requirements seem to be met by frequency-dependent (FD) dampers [20, 21]. With these dampers, they behave as typical, passive dampers at low frequencies and gradually decrease the magnitude of their output at the higher ones. An in-depth review of available and patented solutions in this area was carried out by Franczyk et al. [10]. The authors have shown that most FD-like valves on the market have been developed as add-on system in parallel to the main valves usually operating in the rebound portion of the damper's working cycle. In rare cases, the FD valve is integrated into base (foot) valve assemblies of twin-tube dampers [21].

It is acknowledged by the authors that the term 'frequency-dependent valve' may be misleading; the reader should refer to Dixon [22] for further discussion on the topic. In the essence, FD type valves are pressure-rate dependent valves. In the presented study, the authors simply follow the nomenclature used by automotive OEMS and other parties involved in the development of the FD technology in passenger vehicles.

To summarise, the purpose of this work is to provide a complete, functional math model of the FD damper, which can become the basis for further component level parametric studies. The authors claim that their proposed valve architecture is sufficient to copy the operating principles of most FD damping sys-

tems. The model is verified experimentally to ensure that it is capable of reproducing the characteristics of a real FD shock absorber. To the best knowledge of the authors, no such model of the valve has been presented so far and so it is novel. The second goal is to examine the impact of the FD damper on the comfort and safety of a car.

2. MATERIALS AND METHODS

This subsection presents both the method of modelling the FD shock absorber as well as the performance of a quarter model of a car equipped with this type of shock absorber.

2.1. Passive twin-tube FD damper

In this paragraph a model of a twin-tube, passive adaptive vehicle damper is presented and developed. The damper's structure is shown simply in Fig. 2. As shown, the piston and rod assembly (A) divides the inner volume into the compression chamber (B) and the rebound chamber (C). The reserve chamber (D) is formed by the volume between the inner and the outer tubes. The base valve (E) further separates the oil volume in the reservoir from the compression chamber. The volume of the rebound (upper) chamber is Vr and the pressure in the chamber is pr. Similarly, the volume of the compression chamber (below the piston) is Vc and the compression pressure is referred to as pc. Moreover, the gas pressure is pg, and the volume Vg. While in compression, the oil flows from the compression chamber to the rebound chamber and, in a volume equal to the volume of the piston rod being forced into the damper, to the reserve chamber. In the rebound (upward) stroke the flows are reversed. The references volumetric flow rates are marked in Fig. 2 by means of the following abbreviations: Q_{PVc} – flow rate between the chambers (B) and (C) through the compression side shim stack of the piston valve, Q_{PVr} – flow rate between the chambers (C) and (B) through the rebound side shim stack of the piston valve, Q_{PVO} – flow rate through the constant area orifice in the piston valve assembly linking directly the chambers (B) and (C), Q_{BVc} – flow rate between the chambers (B) and (D) through the compression side shim stack of the base valve assembly, Q_{BVr}– flow rate between the chambers (D) and (B) through the rebound side shim stack of the base valve assembly, Q_{BVO} – flow rate through the constant area orifice located in the base valve. The model assumes the presence of the FD valve assembled within the piston valve and arranged to operate specifically in the damper's rebound stroke, in parallel to the main rebound valve (Q_{FD}). The FD valve, in general, utilises the basic operation principle of conventional damper valves which means it can be considered as a shim or blow-off valve whose opening depends on the pressure difference across it. The main difference, however, is that it is additionally controlled by pilot pressure, which causes its hardening or softening depending on the pressure rate in the main rebound chamber of the shock absorber. The presented solution corresponds to most commercial architectures of the FD valves and incorporates the following three characteristic elements: the accumulator chamber of variable volume, the constant area control orifice (CO) located at its inlet and a pressure-controlled relief valve operated as a bypass in parallel to the main rebound valve [10]. The hydraulic scheme of the described piston valve assembly is shown in Fig. 3.
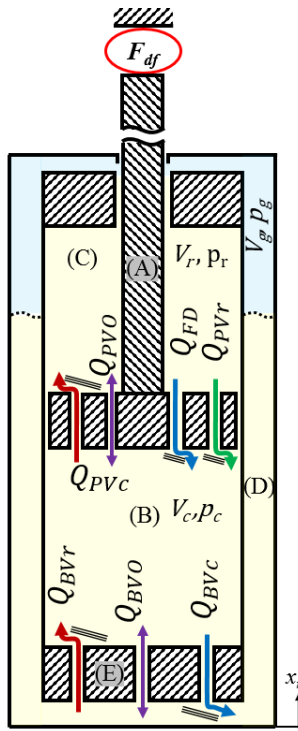
Fig. 2. Diagram of the analysed FD damper
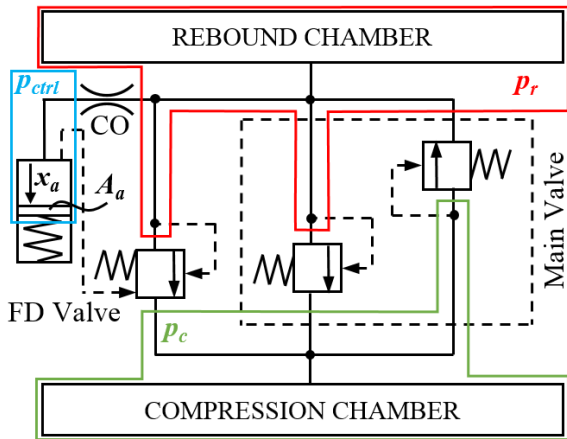
Fig. 3. Hydraulic circuit of the piston valve assembly

$$Q_{PV} = \begin{cases} Q_{PVO} + Q_{PVr} + Q_{FD} & \text{if} & p_r - p_c > 0 \\ 0 & \text{if} & p_r - p_c = 0 \\ Q_{PVO} + Q_{PVc} & \text{if} & p_r - p_c < 0 \end{cases} \quad (5)$$

$$Q_{BV} = \begin{cases} Q_{BVO} + Q_{BVr} & \text{if} & p_c - p_g > 0 \\ 0 & \text{if} & p_c - p_g = 0 \\ Q_{BVO} + Q_{BVc} & \text{if} & p_c - p_g < 0 \end{cases} \quad (6)$$

where p_g is the gas pressure.

The oil flow rates through the two orifice flow rates Q_{PVO} and Q_{BVO} are modelled by means of the modified Bernoulli equation as follows [23, 24]:

$$Q = c_o A \sqrt{\frac{2\Delta p}{\rho}} \quad (7)$$

where Q is the respective flow rate, co refers to the dynamic discharge coefficient; piston valve – *cPVO*, base valve – *cBVO*. The pressure difference Δp is calculated as $\Delta p = p_r - p_c$ (piston valve) or $\Delta p = p_c - p_g$ (base valve), and ρ is the oil density.
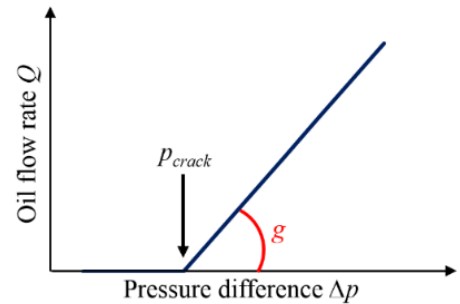
Fig. 4. Generic flow characteristic of the pressure-relief valve
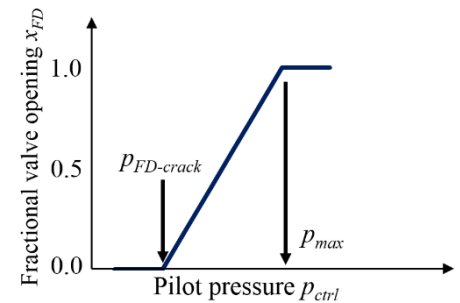
Fig. 5. Opening characteristic of the pilot-operated pressure-relief valve

To start with, taking into account the compressibility of the oil, both rebound and compression pressures (pr, pc) can be calculated using the following expressions [23]:

$$\begin{cases} \dot{p}_r = \beta \frac{(A_p - A_{pr})\dot{x}_r - Q_{PV}}{V_r} \\ \dot{p}_c = \beta \frac{-A_p \dot{x}_r + Q_{PV} + Q_{BV}}{V_c} \\ V_r = V_{r0} + (A_p - A_{pr})x_r \\ V_c = V_{c0} - A_p x_r \end{cases} \quad (1\text{–}4)$$

where β refers to the oil bulk modulus, A_p and A_{pr} are the cross-section areas of the piston and piston rod, respectively, V_r and V_c are the volumes and V_{r0} and V_{c0} are the initial volumes of the rebound and compression chambers, respectively, x_r is the displacement of the damper body with respect to the piston rod and Q_{PV} and Q_{BV} denote the total oil flow rates through the piston valve and the base valve. Based on the diagram in Fig. 1, the respective flow rates are defined as:

For the sake of clarity of the presented analysis, it is assumed that at high damper velocities the flows are realised by means of hydraulic pressure-relief valves, best characterised by the nonlinear relationship Q(Δp). Modelling the behaviour of shim valves is, therefore, outside the scope of this study. That has been well described, for example, by Skačkauska et al. [24], Xu et al. [25], Czop et al. [26] and Farjoud and Ahmadian [27]. In this study the authors further employ a more simplified functional approach. In this manner oil flow rates through each of the four relief-valves can be described using the following relationships [28]:

$$Q_{PVr} =$$
$$\begin{cases} 0 & \text{if} & p_r - p_c \leq p_{PVr-crack} \\ g_{PVr}(p_r - p_c)|p_r - p_c| & \text{if} & p_r - p_c > p_{PVr-crack} \end{cases} \quad (8)$$

Bartłomiej Franczyk, Janusz Gołdasz
Modeling and Experimental Study of a Passive Frequency-Dependent Vehicle Suspension Damper

DOI 10.2478/ama-2024-0045

$$Q_{PVc} = \begin{cases} 0 & \text{if} \quad p_c - p_r \leq p_{PVc-crack} \\ g_{PVc}(p_c - p_r)|p_c - p_r| & \text{if} \quad p_c - p_r > p_{PVc-crack} \end{cases} \quad (9)$$

$$Q_{BVr} = \begin{cases} 0 & \text{if} \quad p_g - p_c \leq p_{BVr-crack} \\ g_{BVr}(p_g - p_c)|p_g - p_c| & \text{if} \quad p_g - p_c > p_{BVr-crack} \end{cases} \quad (10)$$

$$Q_{BVc} = \begin{cases} 0 & \text{if} \quad p_c - p_g \leq p_{BVc-crack} \\ g_{BVc}(p_c - p_g)|p_c - p_g| & \text{if} \quad p_c - p_g \leq p_{BVc-crack} \end{cases} \quad (11)$$

where g_{PVr}, g_{PVc}, g_{BVr}, and g_{BVc} are constant coefficients for defining slopes of the $Q(\Delta p)$ curve as shown in Fig. 4. The pressures $p_{PVr-crack}$, $p_{PVc-crack}$, $p_{BVr-crack}$, and $p_{BVc-crack}$ are cracking pressures of the pressure-relief valves in the piston assembly and the base valve one. Based on this approach, the flow through each valve is initiated only upon exceeding the specific cracking pressure.

As shown in Fig. 3, the FD portion of the rebound valve assembly includes a more complex valve which is modelled here with a pilot-operated pressure-relief valve. The opening of the valve x_{FD} is governed by the pilot (control) pressure p_{ctrl}, the pressure due to the spring pre-tension $p_{FD-crack}$ and the pilot differential pressure p_{max} at maximum opening as revealed in Fig. 5. Then, the flow rate through the FD valve can be simply described as follows [17]:

$$Q_{FD} = \begin{cases} 0 & \text{if} \quad x_{FD} = 0 \\ c_{FD}A_{FD}(x_{FD})\sqrt{\frac{2(p_r - p_c)}{\rho}} & \text{if} \quad 0 < x_{FD} < 1 \\ c_{FD}A_{FDmax}\sqrt{\frac{2(p_r - p_c)}{\rho}} & \text{if} \quad x_{FD} = 1 \end{cases} \quad (12)$$

where c_{FD} is the dynamic discharge coefficient of the FD valve, A_{FDmax} is the maximum valve cross-sectional area and $A_{FD} = x_{FD}A_{FDmax}$ is the valve cross-sectional area due to its fractional (dimensionless) valve opening x_{FD} computed as:

$$x_{FD} = \frac{p_{ctrl} + p_{crack} - p_r}{p_{max}} \quad (13)$$

The pilot pressure p_{ctrl} is calculated considering the dynamics of the accumulator chamber and the CO. Assuming that $p_{ctrl} > F_{0a}/A_a$, where A_a is the accumulator's cross-sectional area and F_{0a} denotes the preload force in the accumulator (due to the spring preload or a precharge pressure), the flow rate into the accumulator's chamber Q_a is:

$$Q_a = A_a \dot{x}_a \quad (14)$$

where x_a is the displacement of the spring-loaded piston in the accumulator. By balancing the forces acting on the piston the following equation is obtained:

$$p_{ctrl}A_a = k_a x_a \quad (15)$$

where k_a is the stiffness ratio of the accumulator's spring. By differentiating Eq. (15), the following equation is obtained

$$\dot{p}_{ctrl}A_a = k_a \dot{x}_a \quad (16)$$

and combining Eqs (14) and (16) yields:

$$\dot{p}_{ctrl} = Q_a \frac{k_a}{A_a^2} \quad (17)$$

It must be noted that the flow rate Q_a is equal to the flow rate through the orifice CO located at the inlet into the accumulator chamber. Applying Eq. (7) here, one obtains:

$$Q_a = c_{CO}A_{CO}\sqrt{\frac{2(p_r - p_{ctrl})}{\rho}} \quad (18)$$

and finally:

$$\dot{p}_{ctrl} = \frac{k_a}{A_a^2} c_{CO}A_{CO}\sqrt{\frac{2(p_r - p_{ctrl})}{\rho}} \quad (19)$$

The variation of the gas pressure p_g can be obtained by using the adiabatic relationship [24]:

$$p_g = p_{g0}\left(\frac{V_{g0}}{V_g}\right)^n = p_{g0}\left(\frac{V_{g0}}{V_g + \int_0^t Q_{BV}}\right)^n \quad (20)$$

where p_{g0} is the initial gas pressure, V_g is the gas volume, V_{g0} is the initial gas volume and n is an adiabatic index.

Finally, the force output F_{df} can be expressed considering the pressures acting on the piston:

$$F_{df} = (A_p - A_{pr})p_r - A_p p_c + p_g A_{pr} \quad (21)$$

Given the set of Eqs (1)–(21), a mathematical model is developed in Siemens AMESim environment ver. 2022 – see Fig. 6. The model parameters were altered to fit the characteristics of a commercially available FD type damper for which experimental data were extracted using the test rig revealed in Section 2.3. The adopted model parameters are listed in Tab. 1.
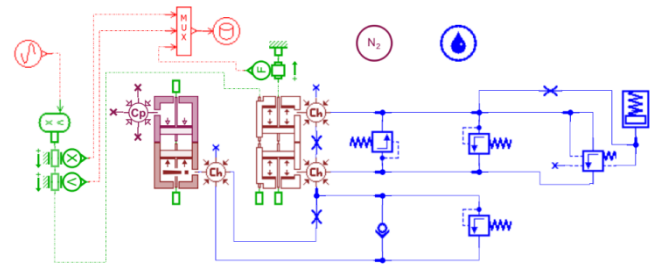
Fig. 6. Structure of the AMESim simulation model of the analysed FD damper

Tab. 1. Set of AMESIM model parameters

Damper and oil parameters								
A_p [mm²]	A_{pr} [mm²]	p_{g0} [MPa]	V_{g0} [m³]	n [ ]	β [Pa]	ρ [kg/m³]	V_{r0} [dm³]	V_{c0} [dm³]
1017.4	254.3	0.6	0.02	1.4	17,000	850	0.06	0.08
Piston valve parameters								
A_{PVO} [mm²]	c_{PVO} [–]		g_{PVr} [m³/s/MPa]		g_{PVc} [m³/s/MPa]		$p_{PVr-crack}$ [MPa]	$p_{PVc-crack}$ [MPa]
2.2	0.99		$3.8 \cdot 10^{-4}$		$1.3 \cdot 10^{-3}$		1.52	0.12

Base valve parameters					
A_{BVO} [mm^2]	c_{BVO} [¯]	g_{BVr} [m^3/s/MPa]	g_{BVc} [m^3/s/MPa]	$p_{BVr\text{-}crack}$ [MPa]	$p_{BVc\text{-}crack}$ [MPa]
0.5	0.93	$1.7 \cdot 10^{-3}$	$2.7 \cdot 10^{-4}$	0.0	1.0

FD valve parameters								
A_{CO} [mm^3]	F_{0a} [N]	A_a [mm^2]	k_a [N/mm]	c_{CO} [¯]	c_{FD} [¯]	A_{FDmax} [mm^2]	$p_{FD\text{-}crack}$ [MPa]	p_{max} [MPa]
0.16	150	397	950	0.7	0.7	5.6	0.7	0.25

2.2. 2-DOF quarter car model

A classic two degree of freedom (2-DOF) quarter car model (QCM) is used in order to study the effectiveness of using FD dampers in passenger vehicles. Such model is a common tool suitable for ride comfort and vehicle safety evaluation studies, respectively [5, 6, 11, 29]. As it is shown in Fig. 7, it consists of the sprung (body) and unsprung (wheel) masses connected by spring and damping elements. There is a spring element placed between the unsprung mass and the ground (road input) which represents the elasticity of a tire. The QCM equations of motion are then as follows:

$$M_w \ddot{z}_2 = k_t(z_2 - x_0) - k_s(z_1 - z_2) - c_s(\dot{z}_1 - \dot{z}_2) \qquad (22)$$

$$M_b \ddot{z}_1 = k_s(z_1 - z_2) - c_s(\dot{z}_1 - \dot{z}_2) \qquad (23)$$

where M_w is the mass of the wheel (kg), M_b refers to the mass of the vehicle body (kg), k_t is the tire stiffness ratio (N/m), k_s is the suspension spring stiffness ratio (N/m), c_s denotes the damping coefficient (Ns/m), z_1 and z_2 are the displacements of the vehicle body and wheel, respectively, and x_0 is the road input.
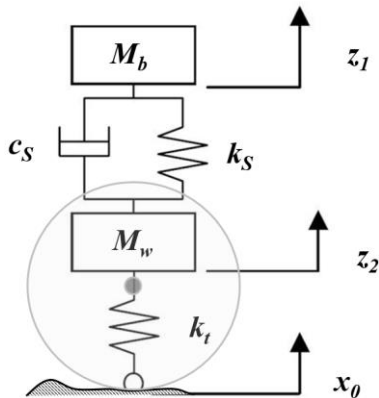
Fig. 7. A 2-DOF QCM

In this paper, two QCMs of the parameters presented in Tab. 2 are studied in terms of ride comfort and safety. The values of M_w, M_b, k_t, and k_s are adopted to suit a front quarter of a vehicle of total weight of 2,400 kg distributed between front and rear axles with the ratio of 60:40 and having the body natural frequency of 0.7 Hz. In this first scenario the authors study the performance of a QCM with a standard damper, whereas the second scenario involves the FD damper with the low frequency performance identical to the standard damper and FD valve parameters as presented in Tab. 1. The QCM models are presented in Fig. 8.

As highlighted in Fig. 9, the reduction in the output forces is ≈400 N at the frequency of 12 Hz, which in relative terms results in 40% degradation achieved at the velocity of v = 0.15 m/s and 17% at the velocity of v = 1.0 m/s.

Tab. 2. QCM parameters

M_w [kg]	M_b [kg]	k_t [kN/m]	k_s [kN/m]
30	720	200	15

QCM, quarter car model.
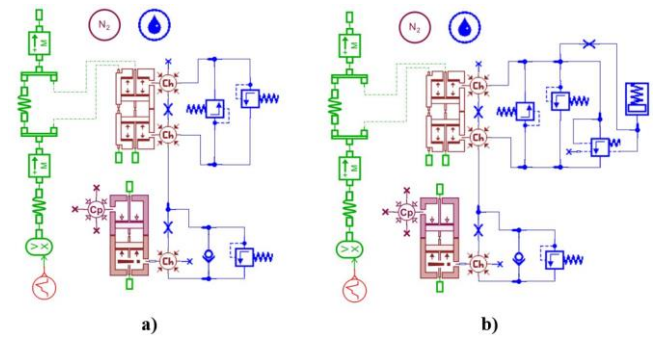
Fig. 8. QC model with (a) a conventional damper, (b) a FD damper
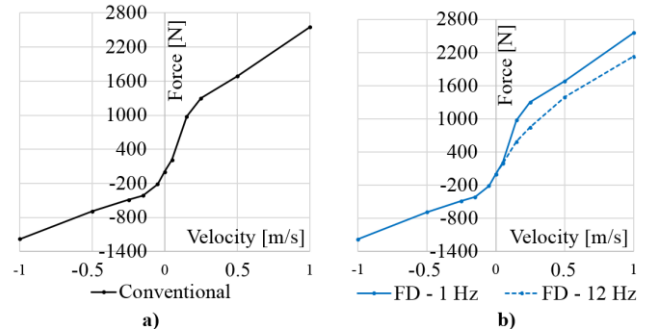
Fig. 9. Measured performance characteristics of (a) a conventional (non-FD) damper, (b) a FD damper

The QCMs were excited with stochastic ISO 8086 type road displacement signals. The ISO 8086 standard provides means for classifying road surface types depending on their quality (roughness) and describes various approaches to be used in order to simulate specific road surface profiles. In this paper, artificial road profiles of the ISO classes A–B (very good), C–D (average) and E–F (very poor), as shown in Fig. 10, are generated using the method similar to the one used in Agostinacchio et al. [30]. For each of the generated road profiles the following three cases are studied varying in the vehicle horizontal velocity – 36 km/h, 72 km/h and 108 km/h (10 m/s, 20 m/s and 30 m/s, respectively). In each case the driving distance is equal to 300 m. Testing the dynamic properties of a vehicle by means of QCM and ISO standardised profiles is an effective and proven method [14, 29, 30]. The selected range of road profiles and driving speeds takes into account the range of operating conditions to which passenger cars are commonly subjected [29, 30].
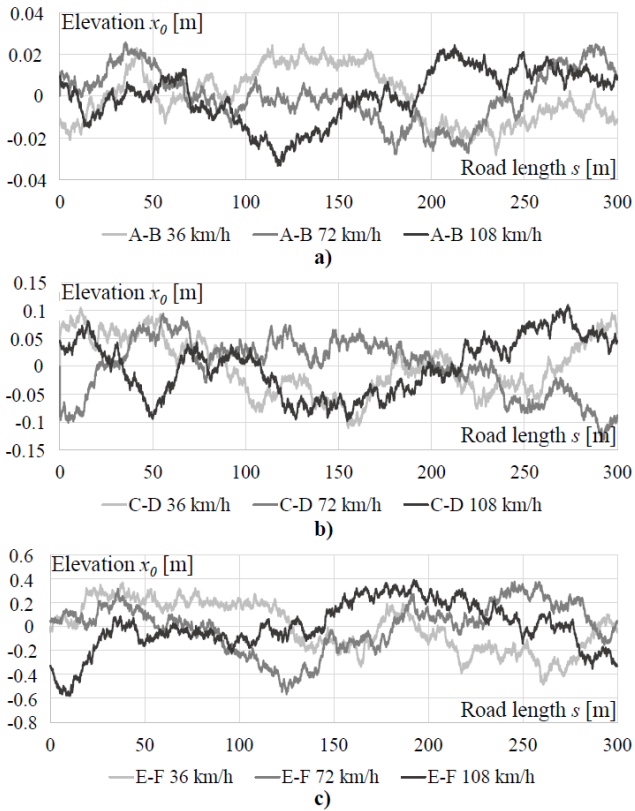
Fig. 10. Generated road profiles for three distinct vehicle's velocities, (a) A–B, (b) C–D and (c) D–E

2.3. Damper test setup, inputs

The damping force measurements of a commercially available passive FD twin-tube damper were measured using the hydraulic test rig shown in Fig. 11.

The MTS stroker was equipped with a MTS 661.21A-01 force transducer with a capacity of 25 kN and a MTS Silentflo 505.11 hydraulic pump capable of flow rates up to 41.5 l/min and the

output pressures up to 207 bar. The damper was mounted in the rig as shown in the photograph with the bottom rigidly attached to the moving base, whereas the piston rod was held stationary. All signals (displacement, velocity, force) were acquired simultaneously at the sampling rate of 1 kHz. The prescribed input displacement was a sinewave resulting in peak velocities up to 0.5 m/s and the frequencies up to 15 Hz. Such testing is the industry standard for most shock absorbers, although specific test speeds depend on manufacturer requirements and suspension design. The test frequencies were selected to cover the range of resonant frequencies of sprung masses and unsprung masses of a typical passenger car, which are usually within the range from $\approx$1.3 Hz to 15 Hz, respectively [1]. The experiment plan of both the laboratory experiments and the numerical simulations covered 75 subsequent test runs according to Tab. 3.
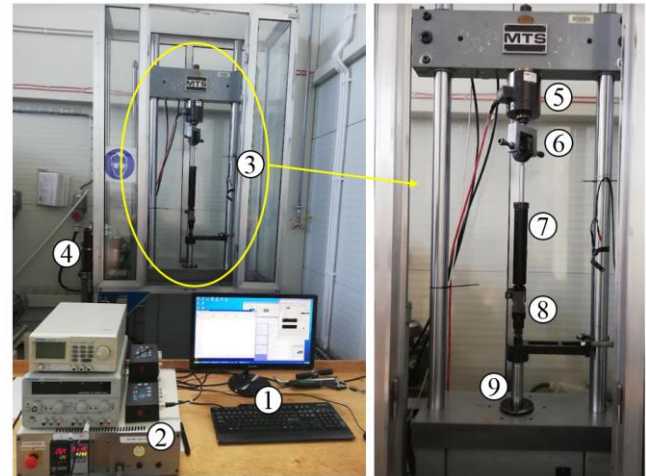
Fig. 11. Test rig for damper performance measurements; 1 – PC with dedicated software for data acquisition, 2 – control unit, 3 – stroker frame, 4 – hydraulic power supply, 5 – force transducer, 6 – upper mount, 7 – test sample, 8 – lower mount, 9 – stroker rod (exciter).

Tab. 3. Test plan for both simulations and rig tests; the input displacement amplitude *a* (mm) varies with the peak velocity *v* and the frequency *f*

f [Hz] *v* [m/s]	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
0.05	8.0	4.0	2.7	2.0	1.6	1.3	1.1	1.0	0.9	0.8	0.7	0.7	0.6	0.6	0.5
0.15	23.9	11.9	8.0	6.0	4.8	4.0	3.4	3.0	2.7	2.4	2.2	2.0	1.8	1.7	1.6
0.25	39.8	19.9	13.3	9.9	8.0	6.6	5.7	5.0	4.4	4.0	3.6	3.3	3.1	2.8	2.7
0.35	55.7	27.9	18.6	13.9	11.1	9.3	8.0	7.0	6.2	5.6	5.1	4.6	4.3	4.0	3.7
0.50	79.6	39.8	26.5	19.9	15.9	13.3	11.4	9.9	8.8	8.0	7.2	6.6	6.1	5.7	5.3

3. RESULTS AND DISCUSSION

In this section the authors reveal the results of two modelling studies. First, the mathematical model of the FD damper is verified experimentally. Second, the QCM based analysis is carried out to access the influence of the FD-damper on comfort and safety metrics.

3.1. FD damper model verification

The accuracy of the model is verified by comparing the damping forces from obtained during the experiments and those in the simulations. The comparison is carried out in the form of static force–velocity plots as well as force–velocity and force–displacement diagrams, respectively. This is a popular method of evaluating damper models in the literature [11, 24, 25].

The obtained results are presented in Figs 12–14. In more detail, Figs 12 and 13 illustrate the output of the model in the form of force–velocity and force–displacement plots, respectively, at three various excitation frequencies f = {1, 6, 12} Hz and the two peak velocities v = {0.25, 0.50} m/s. Finally, Fig. 14 shows the static force-velocity and force-frequency curves at various frequencies and velocities.
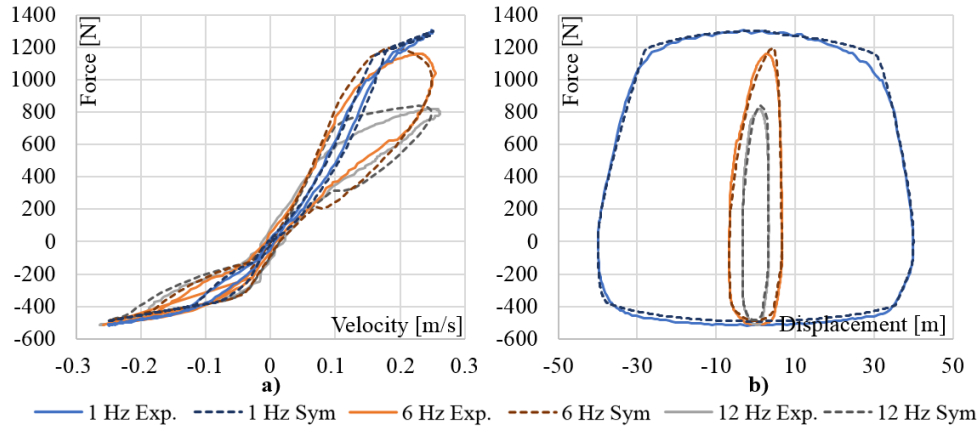
Fig. 12. Model vs experiment, v = 0.25 m/s – (a) force vs velocity, (b) force vs displacement
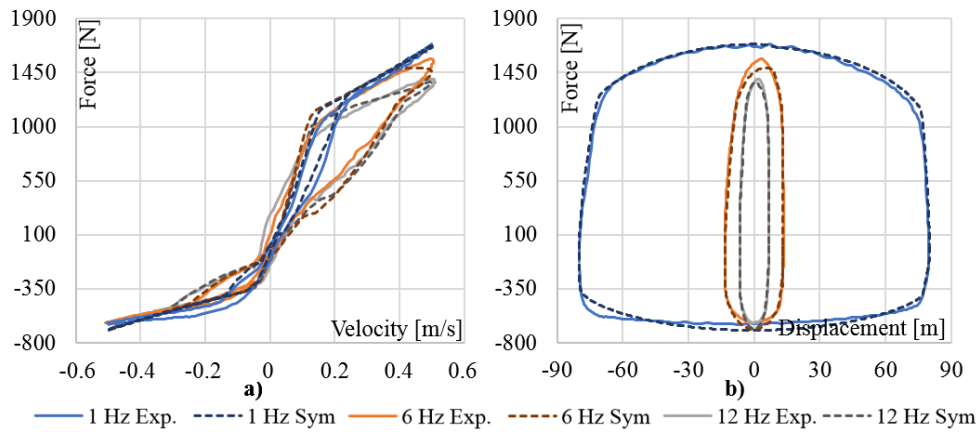
Fig. 13. Model vs experiment, v = 0.50 m/s – (a) force vs velocity, (b) force vs displacement
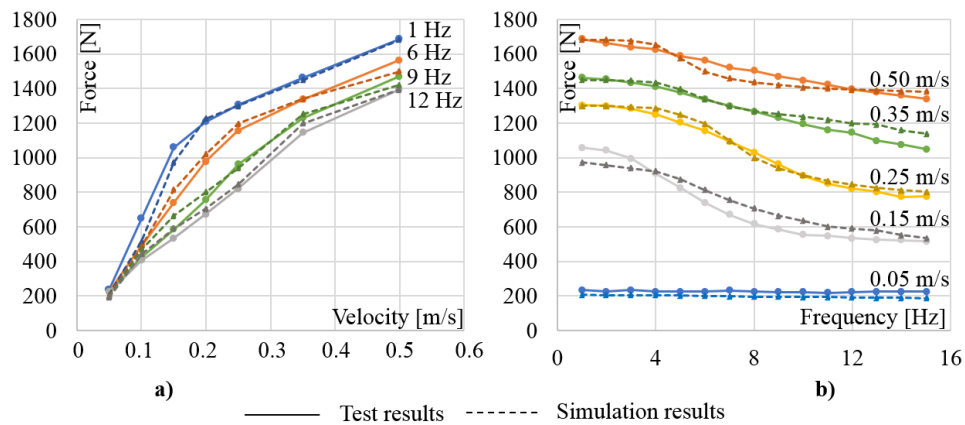
Fig. 14. Model vs experiment; variation of the damping force with frequency and velocity (a) force vs velocity plot, (b) force falloff (force vs frequency) diagram

In order to assess the model's accuracy in predicting the output of the analysed FD damper, relative error metrics were calculated according to the equation below:

$$\delta = \frac{|F_{exp} - F_{sym}|}{F_{exp}} \cdot 100\% \tag{24}$$

where F_{exp} and F_{sym} are damping force values obtained from experimental and simulation tests, respectively. The results are presented in Tab. 4. As shown, the mismatch varies from 0% to nearly 17%. It should be noted, however, that the largest error value is obtained at low damper velocities, where the forces are

small (~200 N). The average error value over the entire frequency range for the velocities from 0.15 m/s to 0.50 m/s is 4%.

Our analysis of the obtained data shows that at the smallest input velocity of 0.05 m/s the force variation against the frequency is almost negligible. However, as the frequency of the input increases the output forces degrade gradually by ≈400 N. Considering the conte nts of Tab. 4, the best agreement at the velocities is obtained at the input velocities >0.05 m/s. It is also evident from the data that the model maintains the typical performance of a passive damper at the lowest frequency. This effect is the most distinctive feature of the FD technology, which sets them apart from the rest of the adaptive dampers. Due to this feature, FD

damper application in a suspension may ensure good support for the vehicle body at low frequencies. On the other hand, at high frequencies, the damping force is built up much more slowly over time, so that the energy of small but frequent vibrations is not transferred to the body, which in turn ensures better comfort. The above comparison of the simulation results and the experimental measurements indicates a good match between the output of the real damper and its mathematical model. Thus, it can be concluded that the presented model accurately reflects the functionality of the FD type adaptive damper and can be successfully used in further research.

Tab. 4. Relative error metrics δ (%): model vs experiment

f [Hz] v [m]	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	Avg
0.05	11	9	12	9	11	12	15	13	12	13	12	13	15	17	17	13
0.15	8	8	6	1	6	10	12	14	13	14	10	10	10	5	4	9
0.20	2	2	4	4	6	4	6	6	6	6	6	5	4	4	4	5
0.25	0	0	1	3	4	3	1	3	2	1	2	3	3	5	3	2
0.35	1	0	1	1	1	0	0	0	2	4	5	5	9	8	9	3
0.50	0	1	2	2	1	4	4	5	3	3	2	0	1	2	3	2
Avg	3	3	4	3	4	5	6	6	6	6	5	5	6	6	6	x

3.2. Analysis of QCM output

The definition of vehicle ride comfort refers to the unpleasant feelings experienced by vehicle passengers and caused by vibrations. It is often evaluated by means of the weighted root-mean-square (RMS) acceleration of the sprung mass according to the following equation [29, 31, 32]:

$$RMS(\ddot{z}_1) = \sqrt{\frac{\int_0^T \ddot{z}_1^2(t)dt}{T}} \qquad (25)$$

where T is the simulation time. In order to compare damping effectiveness of FD damper in relation to a conventional one in terms of ride comfort, a reduction factor is calculated as follows:

$$\delta_{RMS(\ddot{z}_1)} = \left(1 - \frac{RMS(\ddot{z}_{1-FD})}{RMS(\ddot{z}_{1-C})}\right) \times 100\% \qquad (26)$$

where $RMS(\ddot{z}_{1-FD})$ and $RMS(\ddot{z}_{1-C})$ refer to the calculated RMS values involving FD and conventional damper scenarios, respectively.

Next, vehicle's road holding ability directly depends on the tire's ability to stay in contact with the road and to transfer both longitudinal and lateral forces. One commonly used indicator of this performance is the value of the tire normal force T_i, which influences the friction between the tire and the road surface. In this context, the aim is to obtain the highest possible force with a stable course. On the other hand, too much force, especially when dynamic overloads occur, can damage the suspension components. In this paper, road holding indicators are calculated and compared in a manner similar to the ride comfort metric presented above. For this purpose, the variable $\ddot{z}_1$ is replaced by F_t in the Eqs (25) and (26) and so the following road-holding metrics are obtained:

$$RMS(F_t) = \sqrt{\frac{\int_0^T F_t^2(t)dt}{T}} \qquad (27)$$

$$\delta_{RMS(F_t)} = \left(1 - \frac{RMS(F_{t-FD})}{RMS(F_{t-C})}\right) \times 100\% \qquad (28)$$

It should be noted that more elaborate evaluation methods exist for both ride comfort and road-holding metrics for passenger vehicles [30, 33–35]. However, for the purpose of this study, which is a preliminary research preceding more elaborated ones, the authors have decided to use the fundamental metrics.

Fig. 15 highlights exemplary sections of the acceleration $\ddot{z}_1$ and the force F_t in time domain, obtained for the ISO C–D road profile and the vehicle's horizontal velocity of 72 km/h. Fig. 16 presents a summary of the examined RMS metrics.

On analysing the results, it becomes evident that an FD damper's QCM is most effective when driven on middle-class quality roads. Note, however, that the study is limited to one specific FD damper tuning only. In the examined case up to 12% reduction in $RMS(\ddot{z}_1)$ can be observed with the average impact of 2.7% on road holding performance. When driven on good quality roads, the comfort level is increased up to 7% with a negligibly small impact on safety. The biggest yet still small impact of the FD damper on tire loads can be observed when driven on poor quality road of the E–F ISO class. It should be noted, however, that at high driving speeds, this effect is highly beneficial, as it protects the tire from overloading, as in this case the dynamic tire loads reach up to 150% of the tire static load. In particular, examining the contents of Fig. 15, it can be easily observed that the time history of acceleration is much smoother in the case of the FD-based QCM. As seen, most of the sharp peaks are removed, which may be beneficial not only for the passenger's comfort but also for improvement in noise/harshness; acceleration peaks are reduced by up to 40% (see Fig. 15 at t = 3.02 s).
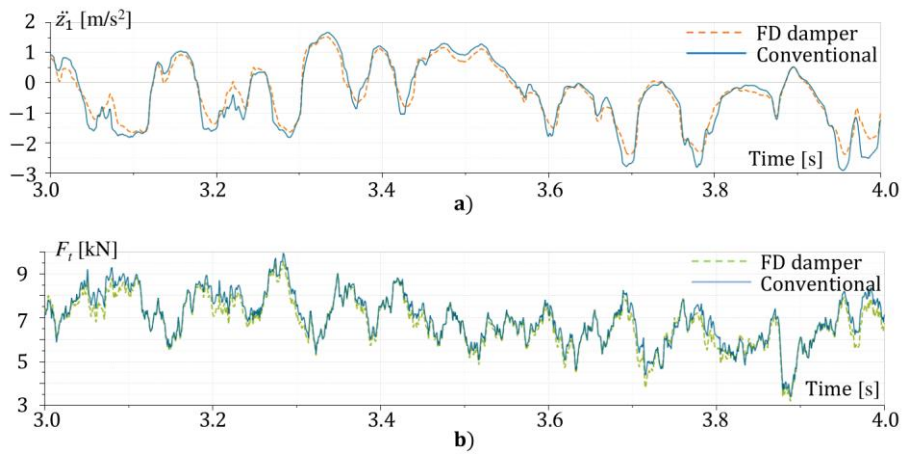
Fig. 15. Time history of (a) body vertical acceleration $\ddot{z}_1$ and (b) tire normal force Ft: conventional damper vs FD damper; profile – ISO C–D at 72 km/h
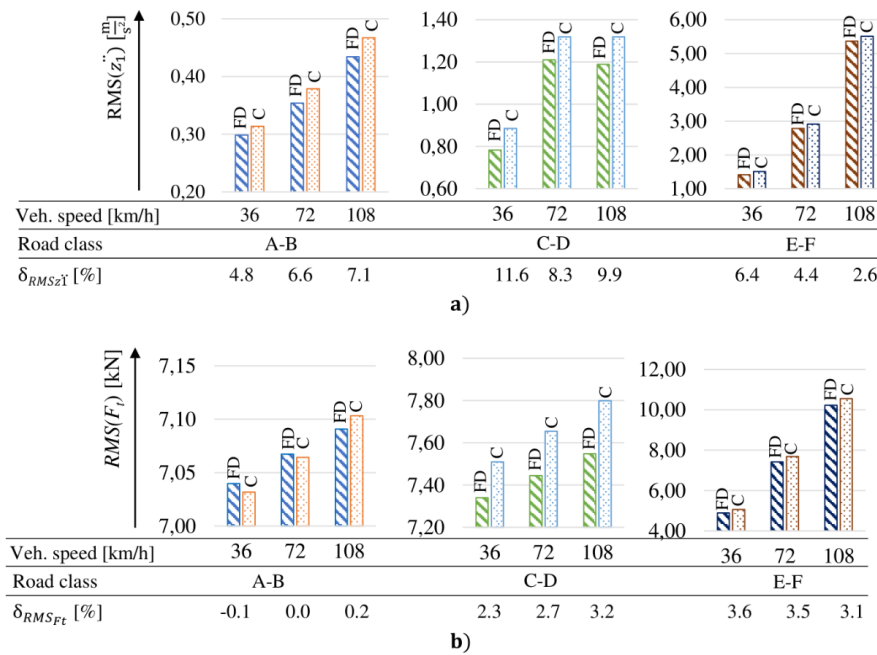
Fig. 16. RMS and δRMS metrics: (a) vehicle body acceleration $\ddot{z}_1$, (b) tire normal force Ft

4. CONCLUSIONS

The purpose of the present study was to provide a dynamic, experimentally verified, functional model of a twin-tube FD damper and to examine its impact on vehicle ride comfort and safety. To the best knowledge of the authors, the hydraulic circuit of the presented model of the FD damper corresponds to the existing solutions available on the automotive market; the model assumes the presence of the FD valve housed within the piston valve assembly and operating on the rebound side of the piston valve in parallel to the main valve assembly. The developed model was verified experimentally across a wide range of frequencies and velocity inputs. It copies the essential features of FD type valving systems, that is, reduction of pressures with the increase of the excitation input frequency. The model is of functional character and presents the principle of operation on which the FD damper technologies are based. It is not design nor configuration oriented.

Furthermore, the simulation studies conducted by means of simple QCM models indicate that the use of the FD damper in vehicles results in improved driving comfort levels with a negligible impact on road holding performance. Depending on the type of surface and driving conditions, the improvement in comfort levels is up to 12%, while the impact on safety does not exceed 4%. In all examined cases, the trade-off between comfort and safety favours the use of FD dampers. The improvements in comfort metrics are supported not only by the obtained RMS values of the body acceleration, but also by the fact that their peak values are smoothed and significantly reduced. Therefore, further improvement in terms of noise and harshness are likely. It must be noted, however, that the research presented in this paper covers one specific case of a vehicle and damper and so broader research must be conducted in order to generalise the conclusions.

The presented model is novel. To our knowledge, no math models of FD type suspension dampers capable of copying the damping force variation with the frequency of the excitation have been developed so far. The FD technology has been present on the automotive market for the last 20 years (mostly as aftermarket solutions) and it has gained the approval of the OEMs recently as the industry tends to seek new solutions in on-going efforts to meet the customer expectations in terms of comfort, handling and NVH using passive damping technologies. The FD dampers may fill in the gap between the conventional passive suspension sys-

tems and more sophisticated state-of-the-art electronically controlled real-time ones. The technology is simple yet effective and easy to integrate into the existing passive valve assemblies. As such, a model that is capable of copying the dynamic behaviour of such valves may be a valuable asset in the vehicle development work.

Finally, the authors plan to develop a parametric model of the FD valve linked to the geometry of a specific FD valve concept so that the model can be an aid in the valve development and engineering process.

REFERENCES

1. Neal MW, Cwycyshyn W, Badiru I. Tuning Dampers for Ride and Handling of Production Vehicles. SAE International. 2015;8(1):152–9.
2. Badiru I, Cwycyshyn WB. Customer focus in ride development. SAE International. 2013;2013(1):1355.
3. Rajmani R. Vehicle Dynamics and Control, 2nd ed. Springer New York Dordrecht Heidelberg London. 2012.
4. Sekulić D, Dedović V. The Effect of Stiffness and Damping of the Suspension System Elements on the Optimisation of the Vibrational Behaviour of a Bus. International Journal for Traffic and Transport Engineering. 2011;1:231–44.
5. Solmaz S, Afatsun AC, Başlamışlı SÇ. Parametric analysis and compensation of ride comfort for electric drivetrains utilizing in-wheel electric motors. International Journal of Simulation: Systems, Science and Technology. 2016;17(33).
6. Nguyen LH, Hong KS, Park S. Road-frequency adaptive control for semi-active suspension systems. Int J Control Autom Syst. 2010;8(5):1029–38.
7. Slaski G. Simulation and experimental testing of adaptive suspension damping control depending on the frequency of a sinusoidal kinematic input. The Archives of Automotive Engineering. 2014;2(64): 165–78.
8. Pletschen N, Spirk S, Lohmann B. Frequency-selective adaptive control of a hybrid suspension system. IFAC Proceedings Volumes. 2013;46(21):237–42.
9. Zhang Y, Guo K, Li SE, Shao X, Zheng M. Prototyping design and experimental validation of membranous dual-cavity based amplitude selective damper. Mech Syst Signal Process. 2016;76–77:810–22.
10. Franczyk B, Maniowski M, Gołdasz J. Frequency-dependent automotive suspension damping systems: State of the art review. Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering [Internet]. 2023;0(0). Available from: https://doi.org/10.1177/09544070231174280
11. Lee CT, Moon BY. Simulation and experimental validation of vehicle dynamic characteristics for displacement-sensitive shock absorber using fluid-flow modelling. Mech Syst Signal Process. 2006;20(2):373–88.
12. Hazaveh NK, Rodgers GW, Chase JG, Pampanin S. Experimental Test and Validation of a Direction- and Displacement-Dependent Viscous Damper. J Eng Mech. 2017;143(11).
13. Ilbeigi S, Jahanpour J, Farshidianfar A. A novel scheme for nonlinear displacement-dependent dampers. Nonlinear Dyn. 2012;70(1): 421–34.
14. Łuczko J, Ferdek U. Nonlinear dynamics of a vehicle with a displacement-sensitive mono-tube shock absorber. Nonlinear Dyn. 2020 Mar 1;100(1):185–202.
15. Łuczko J, Ferdek U, Łatas W. Nonlinear analysis of shock absorbers with amplitude-dependent damping. AIP Conf Proc. 2018;1922(1):100011.
16. Goldasz J. Modelling of amplitude-selective-damping valves. Mechanics and Control. 2011;30(2):60–4.
17. Nie S, Zhuang Y, Wang Y, Guo K. Velocity & displacement-dependent damper: A novel passive shock absorber inspired by the semi-active control. Mech Syst Signal Process. 2018 Jan 15;99:730–46.
18. Xu T, Liang M, Li C, Yang S. Design and analysis of a shock absorber with variable moment of inertia for passive vehicle suspensions. J Sound Vib. 2015 Oct 27;355:66–85.
19. Sikora M. Modeling and Operational Analysis of an Automotive Shock Absorber with a Tuned Mass Damper. Acta Mechanica et Automatica. 2018;12(3).
20. De Kock P, De Ruiter AAW. Shock absorber with frequency-dependent damping. WO 03/040586 A1 (Patent), 2003.
21. Nowaczyk M, Van de Plas J, Vochten J. Shock Absorber With Frequency Dependent Passive Valve. United States; US9638280B2, 2017.
22. Dixon JC. The Shock Absorber Handbook. 2nd ed.. John Wiley & Sons, Ltd; 2007.
23. Sikora M. Study of Flow-Induced Vibration Phenomena in Automotive Shock Absorbers. Mechanics and Control. 2014;33(2).
24. Skačkauskas P, Žuraulis V, Vadluga V, Nagurnas S. Development and verification of a shock absorber and its shim valve model based on the force method principles. Eksploatacja i Niezawodność - Maintenance and Reliability. 2017;19(1):126–33.
25. Xu J, Chu J, Ma H. Hybrid modeling and verification of disk-stacked shock absorber valve. Advances in Mechanical Engineering. 2018;10(2).
26. Czop P, Sławik D, Śliwa P, Wszołek G. Simplified and advanced models of a valve system used in shock absorbers Analysis and modelling. Journal of Achievements in Materials and Manufacturing Engineering. 2009;33(2).
27. Farjoud A, Ahmadian M. Shim stack deflection analysis in hydraulic dampers using energy methods. In: Active and Passive Smart Structures and Integrated Systems 2010. 2010.
28. Siemens Industry Software NV. Simcenter Amesim. Siemens; 2018.
29. Mahmood M, Nassar A, Mohammad H. Analysis and Study Indicators for Quarter Car Model with Two Air Suspension System. Basrah journal for engineering science. 2022;22(2).
30. Agostinacchio M, Ciampa D, Olita S. The vibrations induced by surface irregularities in road pavements - a Matlab® approach. European Transport Research Review. 2014;6(3).
31. Wang P. Effect of electric battery mass distribution on electric vehicle movement safety. In: Vibroengineering Procedia. 2020.
32. Huang S, Nguyen V. Influence of dynamic parameters of electric-vehicles on the ride comfort under different operation conditions. Journal of Mechanical Engineering, Automation and Control Systems. 2021;2(1).
33. Sharma SK, Kumar A. Ride comfort of a higher speed rail vehicle using a magnetorheological suspension system. Proceedings of the Institution of Mechanical Engineers, Part K: Journal of Multi-body Dynamics. 2018;232(1):32-48.
34. Deubel C, Ernst S, Prokop G. Objective evaluation methods of vehicle ride comfort — A literature review. Journal of Sound and Vibration. 2023; 548:117515.
35. Burkhard G, Berger T, Enders E, Schramm D. An extended model of the ISO-2631 standard to objectify the ride comfort in autonomous driving. Work. 2021; 68(s1):37-45.

Bartłomiej Franczyk: https://orcid.org/0000-0002-7566-996X

Janusz Gołdasz: https://orcid.org/0000-0002-0226-3360

EFFECT OF PORE ARCHITECTURE OF 3D PRINTED OPEN POROSITY CELLULAR STRUCTURES ON THEIR RESISTANCE TO MECHANICAL LOADING: PART I – EXPERIMENTAL STUDIES

Monika BERNACKA*/** , **Mehmet ALADAG***, **Adrian DUBICKI***, **Izabela ZGŁOBICKA***

*Faculty of Mechanical Engineering, Bialystok University of Technology, Wiejska 45C, 15-351 Bialystok
**Technology Applied Sp. z o.o., Wiejska 42/3, 15-509 Sobolewo, Poland

monika.bernacka@sd.pb.edu.pl, mehmet.aladag@sd.pb.edu.pl, a.dubicki@pb.edu.pl, i.zglobicka@pb.edu.pl

Abstract: The development of additive manufacturing (AM) techniques has sparked interest in porous structures that can be customized in terms of size, shape, and arrangement of pores. Porous lattice structure (LS, called also lattice struct) offer superior specific stiffness and strength, making them ideal components for lightweight products with energy absorption and heat transfer capabilities. They find applications in industries such as aerospace, aeronautics, automotive, and bone ingrowth applications. One of the main advantages of additive manufacturing is the freedom of design, control over geometry and architecture, cost and time savings, waste reduction, and product customization. However, the designation of appropriate struct/pore geometry to achieve the desired properties and structure remains a challenge. In this part of the study, five lattice structs with various pore sizes, with two volume fractions for each, and shapes (ellipsoidal, helical, X-shape, trapezoidal, and triangular) were designed and manufactured using selective laser sintering (SLS) additive manufacturing technology. Mechanical properties were tested through uniaxial compression, and the apparent stress-strain curves were analyzed. The results showed that the compression tests revealed both monotonic and non-monotonic stress-strain curves, indicating different compression behaviors among the structures. The helical structure exhibited the highest resistance to compression, while other structures showed similarities in their mechanical properties. In Part II of this study provides a comprehensive analysis of these findings, emphasizing the potential of purpose-designed porous structures for various engineering applications.

Key words: Additive manufacturing (AM), porous structures, lattice structs, selective laser sintering, mechanical properties

1. INTRODUCTION

Development of additive manufacturing (AM) (also known as 3D printing) techniques stimulated interest in porous structures which can be designed in terms of size, shape, and arrangement of pores. Of special importance are porous structures composed of repeating cellular units - typically generated by a periodic arrangement of beams/struts forming a three-dimensional framework. Superior specific stiffness and strength of lattice structures (LS, called lattice struct) make them ideal components of lightweight products, which may have the capacity for energy absorption and efficient heat transfer. They are thus widely used in industries such as aerospace, aeronautic and automotive. Their low mass, tunable elastic modulus, and porous nature make them suitable for bone ingrowth applications (1–3). Some of the key advantages of these methods over standard techniques of fabrication of porous materials are design freedom, full control of geometry and architecture, time and cost saving, waste reduction, and the possibility of product customization. Currently, available 3D printing devices allow for the fabrication of products from materials such as metal, ceramics, polymers, and composites. The challenge remains the selection of appropriate technology and process parameters assuring the required properties and structure of printed artifacts (4–7).

Selective laser sintering (SLS) is one of the common methods of additive manufacturing. The most commonly employed material in SLS printing technology offered by EOS is Polyamide (PA-2200). The other materials include TPU 1301, PA-3200 GF (blends polyamide with glass fiber), Alumide (based on PA12 with aluminum), and PA 2210 FR. Its distinct characteristic is that particles of powdered substrate are sintered by a laser beam with no formation of liquid phase. This allows for the fabrication of material-wise highly uniform structures of highly complex architecture (8). Ultra-lightweight elements of controlled porosity can be printed by SLS, which are the subject of the current paper, in which we focus on attention on fine-tuning mechanical properties of lattice-type porous structure by adjusting the size, shape, and arrangement of pores.

Numerous studies have been conducted by researchers to explore the design and fabrication of porous structures. Dong et al (9) made a wide review to collect the affecting parameters on the fabrication of LS. Mantachie et al (10) conducted a research study focusing on the mechanical response of Selective Laser Melting (SLM) lattice structs. Their work aimed to address the lack of an overarching analysis of the mechanical properties of the particular lattice structures. Through a comprehensive review of experimental data, Mantachie provided valuable insights into the design, fabrication, and performance of SLM lattice structs, encompassing their mechanical, electrical, thermal, and acoustic properties. The studies can be replicated with the various researches (11–14). Some of these studies are focused on limitations (capacity) fabrication of porous structures. Yuan et al. (15) prepared a review of additive manufacturing of polymeric composites from material

processing to structural design. It may be also a comprehensive guide to the stakeholders who want to utilize or develop an AM process for polymeric composites. Supporting this study, Hossain et al. (16) addressed the limitations of typical periodic unit cells in porous structures created through laser powder bed fusion (PBF) additive manufacturing. They introduced modifications to create isotropic stochastic porous structures, and identified a remaining limitation in the standard deviation of the elastic modulus values. While the structures showed an improved isotropy and potential for lightweighting and biomaterial applications, further refinement is needed to enhance consistency in their mechanical performance. Both studies contribute valuable knowledge and fill existing gaps in the field. Efforts have been made by researchers to overcome these limitations by refining the processing parameters and improving the porous geometry as well as the mechanical characteristics. This represents an improvement in the thermal and mechanical characteristics of porous structures. To date, several studies have been conducted in this area. Han et al. (17) conducted a thorough review of LS, encompassing their properties, applications, and fabrication methods. They categorized lattice structures as uniform and non-uniform and examine different design approaches, including geometric unit cell-based, mathematical algorithm-generated, and topology optimization methods. The authors also investigated gradient and topology optimization techniques for non-uniform lattice structs and offer insights into the future development of this field. Supporting this research, several studies were conducted to optimize/characterize the structures (18–23). In particular, the compressive behavior of the lattice structs (24–26).

In this study, we designed lattice structures (LS) with two different pore sizes, 1.4 mm, and 1.9 mm, and incorporated five different pore shapes; ellipsoidal, helix, x-shapes, trapezoid, triangle in terms of unit cells. Subsequently, these designs were fabricated using SLS printing technology. To elucidate the impact of pore shape on mechanical properties, compression tests were conducted for each structure.

Initially, the density of the fabricated structures was computed. Following this, a microscopic examination was conducted to identify any manufacturing anomalies. Subsequently, uniaxial compression was applied to the structures. Notably, characteristic features of the apparent stress-strain curves derived from the compression tests were analyzed for each structure. Finally, the applied load, in tandem with the resulting strains, was quantified as a function of their relative density. The results were detailly discussed.

2. MATERIALS AND METHODS

2.1. Modelling the Lattices

Five types of 3D lattices with different geometry of the pores have been designed. SolidWorks was employed to create these lattices structures. The types can be described as having cell units based on and illustrated in **Fig 1**:

− ellipsoidal,
− helix,
− x-shape,
− trapezoid,
− triangle.

Each type specimens were printed in the form of cubes 30 x 30 x 30 mm. Two volume fractions for each structure were obtained under conditions of constant size of the pores of 1.4 and 1.9 mm. The corresponding specimens for each structure type are designated as HP (high porosity) and LP (low porosity), respectively.
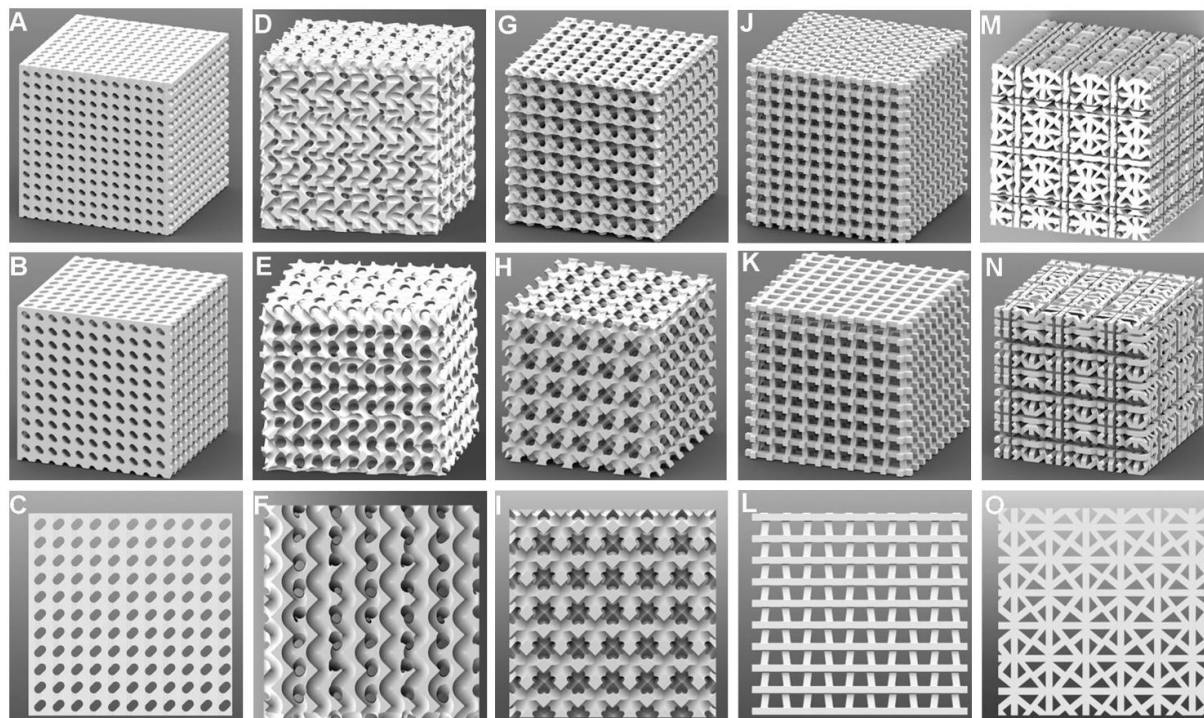
Fig 1. Designed CAD models of lattice structs: (A-C) ellipsoidal , (D-F) helical, (G-I) X-shape, (J-L) trapezoidal, (M-O) triangular, LP (A, D, G, J, M), HP (B, E, H, K, N), representative cross sections (C, F, I, L, O)
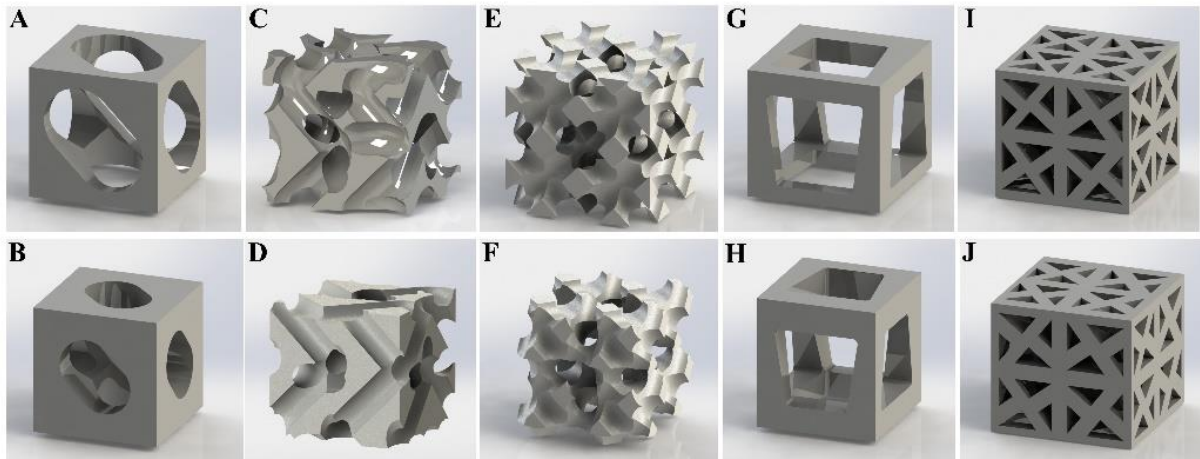
Fig 2. Single cells of lattice structs: (A-B) ellipsoidal , (C-D) helical, (E-F) X-shape, (G-H) trapezoidal, (I-J) triangular, LP (A, C, E, G, I), HP (B, D, F, H, J)

2.2. Fabrication

Lattice structs have been fabricated with SLS 3D printer EOS P395 (EOS GmbH Electro Optical Systems, Germany). The structures were positioned with the same orientation on the bed. The schematical illustration of the printing process is given in Fig 3. Printing process parameters, given in Tab. 1, were kept constant.

Polyamide (PA 2200), used in the present study is one of the most commonly employed material in Selective Laser Sintering (SLS) printing technology offered by EOS. This synthetic thermoplastic polymer has a high biocompatibility, flexibility, hygroscopicity, good chemical resistance as well as high strength and hardness. Selected mechanical properties of polyamide are listed in Tab. 2.
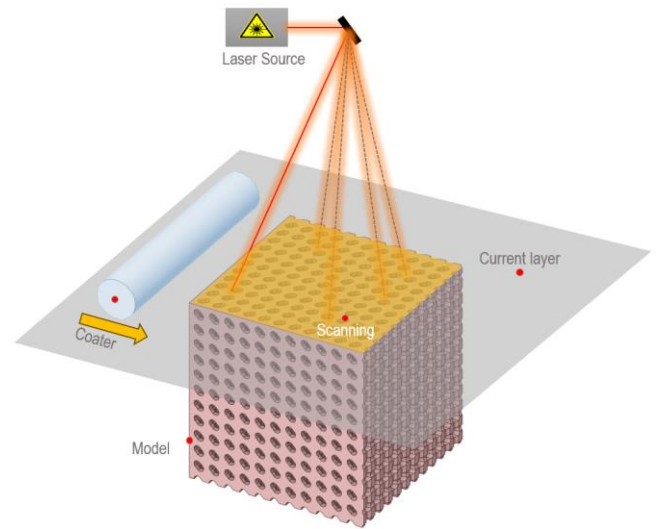
Tab. 1. Parameters of the printing process

Parameters		Unit	Value
Proportions of the virgin/recycled powder		%	50/50
Laser power		W	30
Process temperature		°C	175.5
Temperature of the working chamber		°C	130
Layer thickness		mm	0.12
Material shrinkage	along the X axis	%	3.23
	along the Y axis		3.24
	along the Z axis (0)		2.55
	along the Z axis (600)		1.4

Tab. 2. Mechanical properties of PA 2200

Mechanical properties	Value	Unit
Density	930	kg/m^3
Tensile strength	48	MPa
Tensile Modulus	1650	MPa
Strain at break	18	%
Melting temperature (20°C/min)	176	°C
Shore D hardness	75	—
Powder size	60	µm

Fig 3. The schematic illustration of the printing process

2.3. Compression Test

The compression test was conducted utilizing 322 MTS Load Unit testing machine, incorporating a video extensometer, Aramis 3D (4M), to capture strain localization precisely. The lattice structures were positioned on the platform of the testing machine, and deformation was initiated by imparting a motion of 0.05 mm/sec along the -Y direction. A schematic illustration of the compression test is presented in Fig. 4. Data acquisition during the test was performed by the video extensometer, recording at a rate of 5 images per second to ensure accurate deformation measurement.

Fig. 4. Schematical illustration of the compression test and a view from the Aramis video extensometer.

3. RESULTS AND DISCUSSION

Surface quality of printed specimens was verified with digital microscope (Tagarno Magnus Prestige FHD). Single, un-sintered powder particles were observed between the lattice struts (indicated with arrows in Fig. 5). Mathematical statistical tools were not used due to their very small numbers. It has been assumed that these rarely found printing flaws have negligible effect under conditions of compression tests.The arithmetical value of 3 measurements was taken as the tests results.

The mass of the printed structures, apparent density and porosity are listed in Tab. 3 and shown in Fig. 6. The values of density and porosity of the manufactured samples in Tab.3 was based on the Archimedes law. The highest density/lowest porosity features X-shape LP – 1.03 g/mm3. The HP trapezoidal has the lowest apparent density of 0.647g/ mm3.

As can be seen from Fig. 6, experimental porosity of the printed specimens ranges from 50.46 to 82% and theoretical porosity from 45 to 72%. Experimental porosities of the printed structures have been compared with the theoretical values obtained from CAD modelling – see Fig. 6. It can be concluded that experimental values were higher than predicted from CAD data, apart from helical. Higher than predicted porosity of print-outs can be explained in terms of the effect of struts surface roughness.

Forces applied in compression tests have been analyzed in terms of apparent stress, defined as force divided by cube surface area, as function of strain, defined as reduction in cube height. The stress and strain curves for each structure were plotted in Fig. 7. The solid curves present for LP variants and the dashed for HP ones.
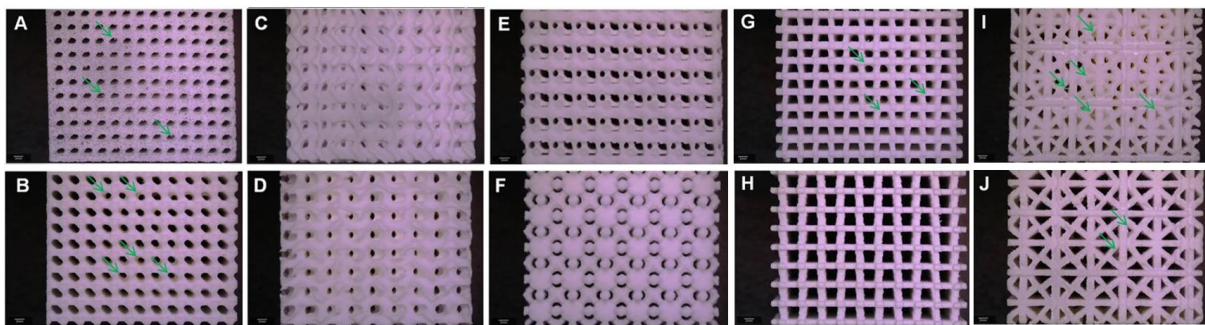
Fig. 5: Digital microscope images of manufactured structures: (A-B) ellipsoidal, (C-D) helical, (E-F) X-shape, (G-H) trapezoidal, (I-J) triangular, where: LP (A, C, E, G, I), HP (B, D, F, H, J); green arrows: single, un-sintered powder particles

Tab. 3. The average mass, density and porosity of the manufactured samples

Structure		Mass [g]	Density [g/mm³]	Porosity [%]
Type of structure	Porosity			
Ellipsoidal	LP	11.23 ± 0.01	0.777	59.15
	HP	9.22 ± 0.14	0.750	62.90
Helical	LP	13.11 ± 0.68	1.03	66.24
	HP	12.33 ± 0.22	0.879	66.71
X-shape	LP	8.77 ± 0.20	0.819	50.46
	HP	8.43 ± 0.06	0.799	51.57
Trapezoidal	LP	6.4 ± 0.22	0.675	73.32
	HP	4.4 ± 0.07	0.647	82.82
Triangular	LP	10.09 ± 0.36	0.828	60.35
	HP	7.33 ± 0.10	0.704	65.99
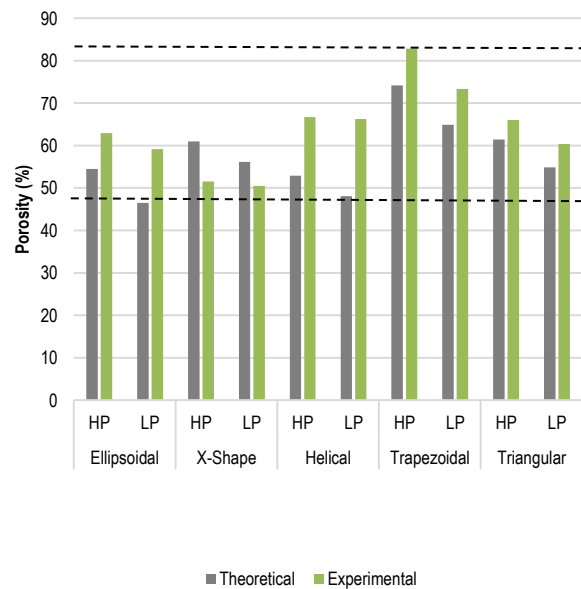
Fig. 6. Porosities of the examined structures
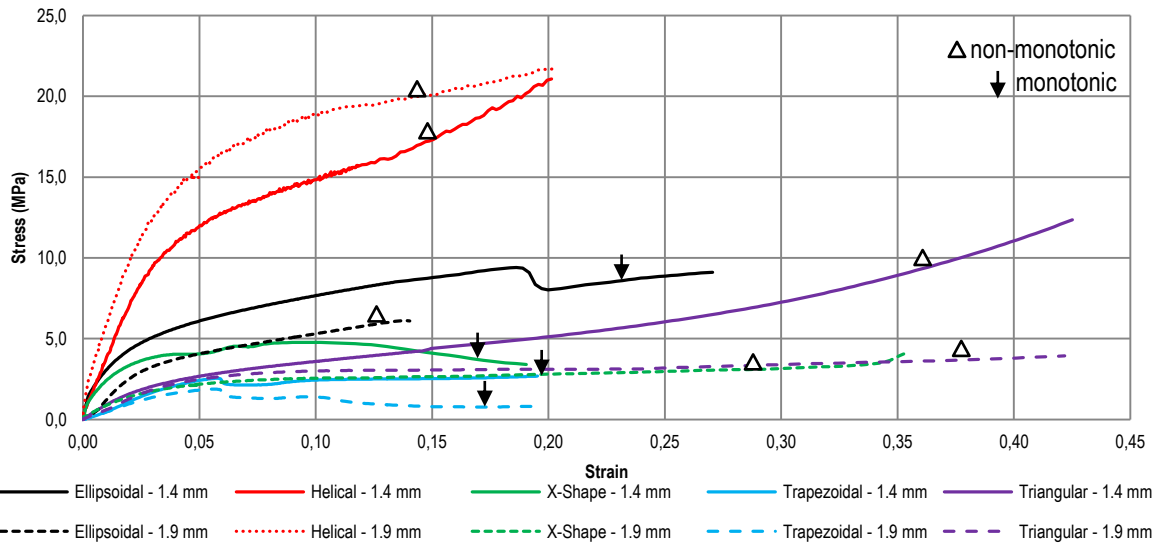
Fig. 7. Stress-strain curves for the samples of the investigated structures

Two types of curves can be distinguished in Fig. 7: (a) monotonic (arrows) and (b) non-monotonic (triangle). To elucidate differences in the compression behavior of various structures, analyses were carried out of apparent stresses at 5 and 15% of apparent strain. The stress values were normalized by the density of a given samples. The results are presented in Tab. 4 and

Tab. 5, for monotonic and non-monotonic curves.

It can be concluded that the tested structures show comparable mechanical properties to previously tested structures with similar parameters, however different unit cells made by using SLS printing technology. Change in pore size and architecture contributed to a decrease/increase in the analyzed mechanical values. The experimental Young's modulus of lattice structs in previous studies ranged from 102.63 to 335.14 MPa. The yield strength of lattice structs ranges from 2.07 to 13.73 MPa. The structures examined in the article are characterized by a Young's modulus of 47.54 -544.12 MPa and a yield strength of 0.935-9.95 MPa (27). Han et al. (28) developed a graded structure through. They established a mathematical correlation between graded porosity and compressive properties. By adjusting the graded volume fraction, the elastic modulus and yield strength of the structures can be customized within the range of 0.28–0.59 GPa and 3.79–17.75 MPa, respectively.

Tab. 4. Normalized values of forces for monotonic parts

Structure	Porosity	$F_{5\%}$ $[\frac{kN}{g/mm^3}]$	$F_{15\%}$ $[\frac{kN}{g/mm^3}]$	Young's modulus $[\frac{GPa}{g/mm^3}]$	Initial stiffness $[\frac{1}{g/mm^3}]$
Ellipsoidal	LP	9.27	12.32	0.72	1413.71
	HP	5.2	-	0.47	463.28
Helical	LP	14.18	19.58	1.16	2395.3
	HP	16.49	22.08	1.44	2025.1
X-shape	LP	5.19	5.58	0.56	1109.86
	HP	2.75	3.17	0.23	299.73
Triangular	LP	3.31	5.21	0.25	339.44
	HP	2.83	4.08	0.38	235.56

Tab. 5. Normalized values of forces for non-monotonic parts

Structure	Porosity	F_{max} $\frac{kN}{g/mm^3}$	F_{min} $\frac{kN}{g/mm^3}$	$F_{5\%}$ $\frac{kN}{g/mm^3}$	$F_{15\%}$ $[\frac{kN}{g/mm^3}]$	Initial stiffness $[\frac{1}{g/mm^3}]$
Trapezoidal	LP	4.22	1.79	3.32	2.00	633.66
	HP	2.55	0.87	2.22	1.07	496.31

Tab. 6. Values of forces for all tested structure

Structure	Porosity	Young's modulus [MPa]	Yield strength [MPa]
Ellipsoidal	LP	492,23	3,05
	HP	126,09	2,63
Helical	LP	352,92	9,09
	HP	544,12	9,95
X-shape	LP	306,42	2,422
	HP	70,176	1,57
Trapezoidal	LP	56,247	2,066
	HP	47,536	1,368
Triangular	LP	83,972	1,856
	HP	59,64	0,935

The highest value of the force for 5% deformation was obtained for the helical with LP porosity - 16.49 kN/(g/mm³). The non-monotonic structure helical and trapezoidal have similar values of the force F_{max}. The highest values of the F_{max} were obtained for LP X-shape, equal to 6.77 kN/(g/mm³). Additionally, for non-monotonic structures, the minimal force F_{min} was calculated. HP trapezoidal has the highest value of this parameter, equal to 0.87 kN/(g/mm³).

Structures ellipsoidal, X-shape, trapezoidal and triangular were characterized by a high similarity of the obtained values of the initial stiffness (see in Fig. 8). The highest values of the initial stiffness were obtained for monotonic LP helical equal to 2395.3 (1/g/mm³). In all specimens the initial stiffness dropped with the increasing porosity.

By analyzing the values of normalized forces obtained for 15% deformation a high similarity of the obtained values for porous X-shape and trapezoidal was observed. The highest value of the force for 15% deformation was measured for LP helical, equal to 22.08 kN/(g/mm³) (see in Fig. 9). All obtained values of compres-

sion force under 15% of strain were plotted in Fig. 9. This plot allows to notice distinct compression resistance of helical. Also, it can be noted that unlike in the case of structures X-shape, trapezoidal, triangular, higher compression force has been measured for HP case.
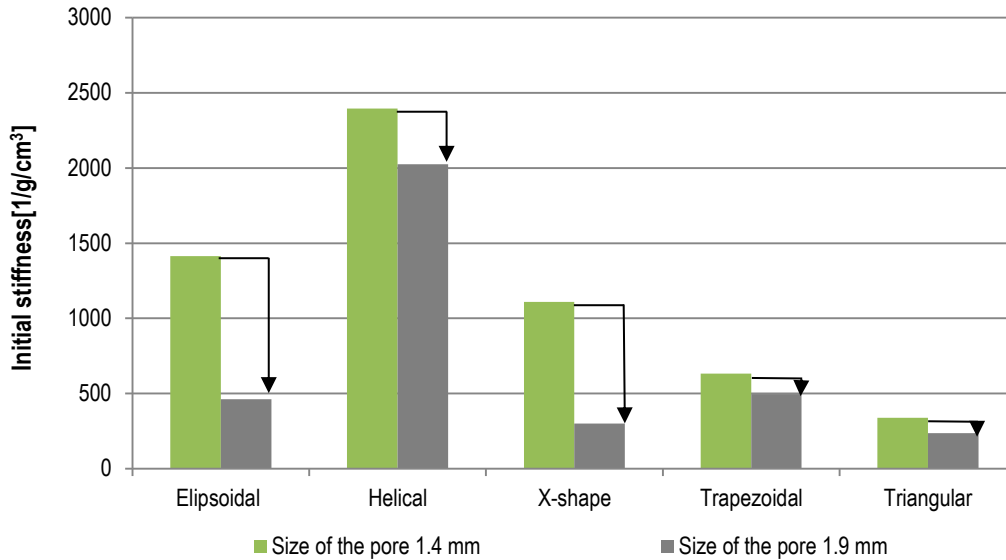
Fig. 8. Initial stiffness for structure with size of the pore 1.4 and 1.9 mm; black arrows: increase/decrease in values
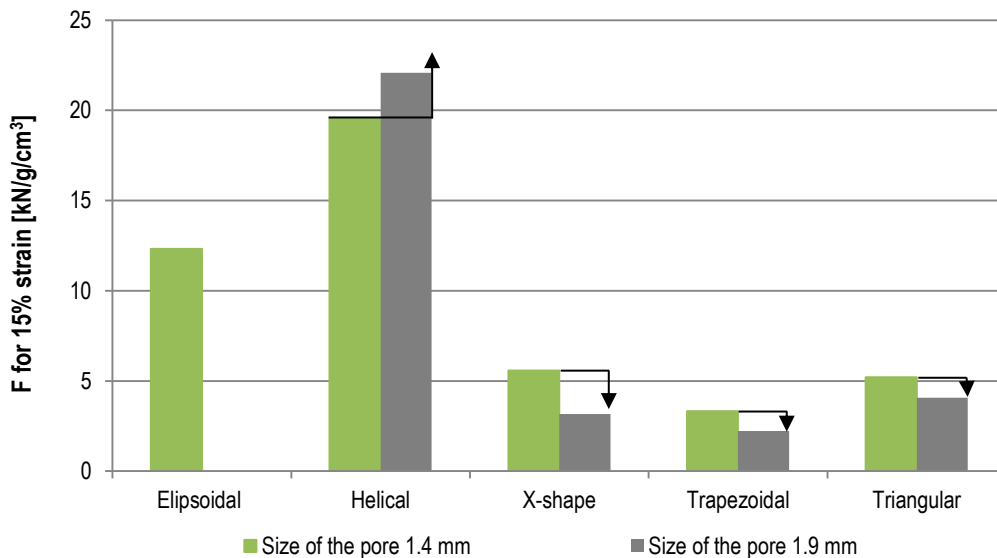
Fig. 9. F for 15% strain for structure with size of the pore 1.4 and 1.9 mm; black arrows: increase/decrease in values

In Tab. 7 - 16, the strains as a result of deformation in both direction , X and Y, are given for each structure. To measure the displacement, the two points were selected in X and Y direction and the distance between these points is given in L_{iy} and L_{ix} columns for unloaded condition, L_{fy} and L_{fx} for after 10% of deformation. For each structure, the displacement values in both the Y and X directions and the corresponding stresses give information about their mechanical behavior. Under 0% strain, the measurements represent the initial configurations of the structures and show no deformation. In contrast, under 10% strain, the displacement values show the response of the shapes to external forces and show the extent

of deformation, with strain values measuring this change. It can be noted that all structures showcase a positive Poisson's ratio in response to deformation. Under a 10% strain, the strains in the X-direction remain uniform among the ellipsoidal, X-shaped, trapezoidal, and triangular structures, each registering a 1% strain. However, the helical structure differs from other structures, showing a difference of 1% between LP and HP configurations. Specifically, the LP helical structure shows a 2% strain, while the HP exhibits a 3% strain. As previously noted, in this comparative analysis table, the helical structure carried a remarkable load of approximately 17 kN compared to other structures.

Tab. 7. Results of displacement measurements for ellipsoidal – 0% strain

Ellipsoidal	Force [N]	L_{iy} [mm]	Strain Y	L_{ix} [mm]	Strain X
LP	0	17.75	0	19.91	0
HP		13.64		20.59	

Tab. 8. Results of displacement measurements for ellipsoidal – 10% strain

Ellipsoidal	Force [N]	L_{fy} [mm]	Strain Y	L_{fx} [mm]	Strain X
LP	8825.38	15.96	0.1	20.22	0.01
HP	4637.57	12.28		20.75	

Tab. 9. Results of displacement measurements for helical – 0% strain

Helical	Force [N]	L_{iy} [mm]	Strain Y	L_{ix} [mm]	Strain X
LP	0	18.21	0	17.69	0
HP		14.76		19.97	

Tab. 10. Results of displacement measurements for helical – 10% strain

Helical	Force [N]	L_{fy} [mm]	Strain Y	L_{fx} [mm]	Strain X
LP	17482.67	16.4	0.1	18.01	0.02
HP	18034.79	13.25		20.65	0.03

Tab. 11. Results of displacement measurements for X-shape – 0% strain

X-Shape	Force [N]	L_{iy} [mm]	Strain Y	L_{ix} [mm]	Strain X
LP	0	17.85	0	14.02	0
HP		24.49		24.89	

Tab. 12. Results of displacement measurements for X-shape – 10% strain

X-Shape	Force [N]	L_{fy} [mm]	Strain Y	L_{fx} [mm]	Strain X
LP	5537.68	16.03	0.1	14.09	0.01
HP	2517.55	22.07		25.24	

Tab. 13. Results of displacement measurements for trapezoidal – 0% strain

Trapezoidal	Force [N]	L_{iy} [mm]	Strain Y	L_{ix} [mm]	Strain X
LP	0	27.22	0	22.6	0
HP		21.22		20.54	

Tab. 14. Results of displacement measurements for trapezoidal – 10% strain

Trapezoidal	Force [N]	L_{fy} [mm]	Strain Y	L_{fx} [mm]	Strain X
LP	3931.84	24.49	0.1	22.72	0.01
HP	868.04	19.12		20.69	

Tab. 15. Results of displacement measurements for triangular – 0% strain

Triangular	Force [N]	L_{iy} [mm]	Strain Y	L_{ix} [mm]	Strain X
LP	0	18.44	0	20.77	0
HP		20.79		20.87	

Tab. 16. Results of displacement measurements for triangular – 10% strain

Triangular	Force [N]	L_{fy} [mm]	Strain Y	L_{fx} [mm]	Strain X
LP	3661.37	16.59	0.1	21.01	0.01
HP	3560.75	18.71		21.15	

4. CONCLUSION

Based on the obtained results for give types of lattice structures were designed with different pore shapes, with ellipsoidal, X-shaped, helical, trapezoidal, and triangular cell units, the following conclusions were drawn.

- The surface quality of the printed lattice structures was examined using a digital microscope, revealing the presence of occasional un-sintered powder particles between the lattice struts. These anomalies have a negligible impact on compression test results.
- The tests revealed characteristic stress-strain curves, with apparent porosities ranging from 50.46% to 82%, surpassing the theoretical values obtained from CAD modeling, except for the helical structure.
- The apparent density and porosity of the structures varied across different designs, with the X-shape LP variant exhibiting the highest density (1.03 g/mm³) and the HP trapezoidal variant having the lowest apparent density (0.647 g/mm³).
- Detailed analyses of compression behavior involved distinguishing between monotonic and non-monotonic stress-strain curves. Apparent stress values at 5% and 15% apparent strain were normalized by the density of each sample. The helical structure, particularly in the LP variant, exhibited the highest force for 5% deformation, 16.49 kN/(g/mm³), while X-shape LP demonstrated the highest force for 15% deformation, 22.08 kN/(g/mm³).
- Initial stiffness values revealed a general decline with increasing porosity across all specimens. Notably, the helical LP variant displayed the highest initial stiffness (2395.3 1/g/mm³). Comparative analyses of normalized forces for 15% deformation highlighted similarities between porous X-shape and trapezoidal structures.
- It has been shown that porous structure consisting of cell units with a helix structure has the most favorable mechanical properties. From this end, geometry of the designed structures is the main determinant of their mechanical properties in terms of the resistance to compression force.
- Finally, helical structure is characterized by the highest resistance to compression. To elucidate this, a comprehensive numerical analysis of this finding is extensively investigated in Part II.

As a summary of the study, it can be noted that the results indicate a high potential for purpose-designed porous structures to meet the requirements of certain applications in areas that require low-weight, high-stiffness such as biomedical and aerospace.

REFERENCES

1. Ruiz de Galarreta S, Jeffers JRT, Ghouse S. A validated finite element analysis procedure for porous structures. Mater Des. 2020 Apr 1;189:108546.
2. Guerra Silva R, Torres MJ, Zahr Viñuela J. A Comparison of Miniature Lattice Structures Produced by Material Extrusion and Vat Photopolymerization Additive Manufacturing. Polymers (Basel) [Internet]. 2021 Jul 1 [cited 2023 Jan 30];13(13). Available from: https://pubmed.ncbi.nlm.nih.gov/34208960/
3. Cipriani CE, Ha T, Martinez Defilló OB, Myneni M, Wang Y, Benjamin CC, et al. Structure–Processing–Property Relationships of 3D Printed Porous Polymeric Materials. ACS Mater Au. 2021 Sep 8;1(1): 69–80.
4. Chen H, Han Q, Wang C, Liu Y, Chen B, Wang J. Porous Scaffold Design for Additive Manufacturing in Orthopedics: A Review. Front Bioeng Biotechnol. 2020 Jun 17;8:609.
5. Tofail SAM, Koumoulos EP, Bandyopadhyay A, Bose S, O'Donoghue L, Charitidis C. Additive manufacturing: scientific and technological challenges, market uptake and opportunities. Mater Today. 2018 Jan 1;21(1):22–37.

6. Gadomska-Gajadhur A, Łojek K, Szymaniak M, Gadomska A. Materiały porowate do regeneracji tkanki chrzęstnej i kostnej. Wyr Med. 2018;3.

7. Kruk A, Gadomska-Gajadhur A, Ruskowski P, Chwojnowski A, Synoradzki L. Otrzymywanie polilaktydowych rusztowań komórkowych o strukturze gąbczastej – badania wstępne i optymalizacja procesu. Polimery. T. 62. 2017;2(2):118–26.

8. Mierzejewska Ż. Technologia SLS – charakterystyka i zastosowanie selektywnego spiekania laserowego w inżynierii biomedycznej. J Technol Exploit Mech Eng. 2015;1:178–90.

9. Maconachie T, Leary M, Lozanovski B, Zhang X, Qian M, Faruque O, et al. SLM lattice structures: Properties, performance, applications and challenges. Mater Des [Internet]. 2019 Dec 1 [cited 2023 Jun 16];183(10):108137.
Available from: https://dx.doi.org/10.1115/1.4037305

10. Maconachie T, Leary M, Lozanovski B, Zhang X, Qian M, Faruque O, et al. SLM lattice structures: Properties, performance, applications and challenges. Mater Des [Internet]. 2019;183:108137. Available from: https://doi.org/10.1016/j.matdes.2019.108137

11. Uribe-Lam E, Treviño-Quintanilla CD, Cuan-Urquizo E, Olvera-Silva O. Use of additive manufacturing for the fabrication of cellular and lattice materials: a review.
https://doi.org/101080/1042691420201819544 [Internet]. 2020 [cited 2023 Jun 16];36(3):257–80. Available from: https://www.tandfonline.com/doi/abs/10.1080/10426914.2020.1819544

12. Tao W, Leu MC. Design of lattice structure for additive manufacturing. Int Symp Flex Autom ISFA 2016. 2016 Dec 16;325–32.

13. Bhat C, Kumar A, Lin SC, Jeng JY. Design, fabrication, and properties evaluation of novel nested lattice structures. Addit Manuf. 2023 Apr 25;68:103510.

14. Kantaros A, Piromalis D. Fabricating Lattice Structures via 3D Printing: The Case of Porous Bio-Engineered Scaffolds. Appl Mech 2021, Vol 2, Pages 289-302 [Internet]. 2021 May 25 [cited 2023 Jun 16];2(2):289–302. Available from: https://www.mdpi.com/2673-3161/2/2/18/htm

15. Yuan S, Li S, Zhu J, Tang Y. Additive manufacturing of polymeric composites from material processing to structural design. Compos Part B Eng [Internet]. 2021;219(April):108903. Available from: https://doi.org/10.1016/j.compositesb.2021.108903

16. Hossain U, Ghouse S, Nai K, Jeffers JR. Controlling and testing anisotropy in additively manufactured stochastic structures. Addit Manuf. 2021 Mar 1;39:101849.

17. Pan C, Han Y, Lu J. Design and Optimization of Lattice Structures: A Review. Appl Sci 2020, Vol 10, Page 6374 [Internet]. 2020 Sep 13 [cited 2023 Jun 16];10(18):6374. Available from: https://www.mdpi.com/2076-3417/10/18/6374/htm

18. Wang P;, Yang F;, Zhao J, Wang P, Yang F, Zhao J. Compression Behaviors and Mechanical Properties of Modified Face-Centered Cubic Lattice Structures under Quasi-Static and High-Speed Loading. Mater 2022, Vol 15, Page 1949 [Internet]. 2022 Mar 6 [cited 2022 Aug 5];15(5):1949. Available from: https://www.mdpi.com/1996-1944/15/5/1949/htm

19. Beloshenko V, Beygelzimer Y, Chishko V, Savchenko B, Sova N, Verbylo D, et al. Mechanical Properties of Thermoplastic Polyurethane-Based Three-Dimensional-Printed Lattice Structures: Role of Build Orientation, Loading Direction, and Filler. 3D Print Addit Manuf [Internet]. 2021 May 14 [cited 2021 Nov 4];3dp.2021.0031. Available from: https://www.liebertpub.com/doi/abs/10.1089/3dp.2021.0031

20. Li S, Yuan S, Zhu J, Zhang W, Tang Y, Wang C, et al. Optimal and adaptive lattice design considering process-induced material anisotropy and geometric inaccuracy for additive manufacturing. Struct Multidiscip Optim. 2022 Jan 1;65(1):1–16.

21. Song J, Wang Y, Zhou W, Fan R, Yu B, Lu Y, et al. Topology optimization-guided lattice composites and their mechanical characterizations. Compos Part B Eng. 2019 Mar 1;160:402–11.

22. Bahrami Babamiri B, Askari H, Hazeli K. Deformation mechanisms and post-yielding behavior of additively manufactured lattice structures. Mater Des. 2020 Mar 1;188.

23. Yavas D, Liu Q, Zhang Z, Wu D. Design and fabrication of architected multi-material lattices with tunable stiffness, strength, and energy absorption. Mater Des [Internet]. 2022 May 1 [cited 2022 Nov 3];217:110613. Available from:
https://linkinghub.elsevier.com/retrieve/pii/S0264127522002349

24. Yang L, Yan C, Cao W, Liu Z, Song B, Wen S, et al. Compression–compression fatigue behaviour of gyroid-type triply periodic minimal surface porous structures fabricated by selective laser melting. Acta Mater. 2019 Dec 1;181:49–66.

25. Park JH, Park K. Compressive behavior of soft lattice structures and their application to functional compliance control. Addit Manuf. 2020 May 1;33:101148.

26. Zhang L, Lifton J, Hu Z, Hong R, Feih S. Influence of geometric defects on the compression behaviour of thin shell lattices fabricated by micro laser powder bed fusion. Addit Manuf [Internet]. 2022 Oct 1 [cited 2022 Dec 20];58:103038. Available from:
https://linkinghub.elsevier.com/retrieve/pii/S2214860422004304

27. Zhao Z, Wu Z, Yao D, Wei Y, Li J. Mechanical properties and failure mechanisms of polyamide 12 gradient scaffolds developed with selective laser sintering. J Mech Behav Biomed Mater. 2023 Jul 1;143:105915.

28. Han C, Li Y, Wang Q, Wen S, Wei Q, Yan C, et al. Continuous functionally graded porous titanium scaffolds manufactured by selective laser melting for bone implants. J Mech Behav Biomed Mater. 2018 Apr 1;80:119–27.

Monika Bernacka: https://orcid.org/0000-0003-3481-0768

Mehmet Aladag: https://orcid.org/0000-0002-2484-7519

Adrian Dubicki: https://orcid.org/0000-0002-3994-2957

Izabela Zglobicka: https://orcid.org/0000-0002-4432-9196

EFFECT OF PORE ARCHITECTURE OF 3D PRINTED OPEN POROSITY CELLULAR STRUCTURES
ON THEIR RESISTANCE TO MECHANICAL LOADING:
PART II – NUMERICAL ANALYSIS

Mehmet ALADAG*, Monika BERNACKA, Adrian DUBICKI*, Izabela ZGŁOBICKA***

*Faculty of Mechanical Engineering, Bialystok University of Technology, Wiejska 45C, 15-351 Bialystok, Poland
**Technology Applied Sp. z o.o., Wiejska 42/3, 15-509 Sobolewo, Poland

mehmet.aladag@sd.pb.edu.pl, monika.bernacka@sd.pb.edu.pl, a.dubicki@pb.edu.pl, i.zglobicka@pb.edu.pl

Abstract: The objective of this study was to investigate pore shape effect on resistance to compression of open porosity lattice structures obtained with 3D printing. To this end, three distinct pore architectures were investigated: ellipsoidal, helical and X-shape. Open porosity of these structures was 54%, 50% and 60%, respectively. Their mechanical properties were evaluated through compression tests, and their behaviours were analysed using finite element modelling (FEM). The results indicated that the pore size has a significant effect on the stiffness of cellular structures. It was observed that the helical structure exhibited superior properties among the structures tested due to homogenous stress distribution. It was also found that ellipsoidal and x-shape structures are much more sensitive to localisation of the deformation.

Key words: lattice structures, symmetry, 3D printing, porous elements, compression, finite element method

1. INTRODUCTION

The recent development of 3D printing methods has opened new opportunities for the fabrication of open porosity lattice structures (OPLSs). Such structures generally are characterised by a low weight and find applications in a number of products ranging from tissue engineering to motor parts (1).

Generally, the strength of OPLSs depends on the properties of printing materials and total porosity (2). Architected porous materials, developed to meet strength, stiffness, and toughness needs, offer improved mechanical properties, saving on weight and costs without compromising structural integrity in various engineering domains (3). However, as it is already common knowledge for composite materials, OPLSs can be viewed as pore-interpenetrated composites; also size, shape and spatial arrangement of pores influence their mechanical properties. Thirunavukkarasu et al. (4) demonstrated that manipulating topology unit cell designs with consistent aspect ratios can enhance mechanical performance, resulting in increased critical load tolerance, optimised buckling resistance and improved energy dissipation rates for versatile applications. The influence of pore size/shape/distribution, in short, pore architecture, can be analysed directly by considering the size/shape connectivity of struts, which are counterparts to pores in OPLSs. Indirectly, the strength of OPLSs can be approached with stereological parameters such as the volume fraction of pores, Vv, and their specific surface Sv (5). Both approaches are adopted in the present study to infer the effect of pore shape on the mechanical properties of OPLSs.

Analysis of the effect of pore architecture on properties of OPLSs presented here is stimulated by results recently reported by Bernacka et al. (6). In part one, five different lattice structures with various pore sizes and shapes, with two volume fractions for each, and shapes (ellipsoidal, helical, X-shape, trapezoidal and triangular) were designed and fabricated using the selective laser sintering (SLS) additive manufacturing (AM) method. Mechanical properties were tested through uniaxial compression, and the apparent stress-strain curves were analysed. In part two, we selected three lattice structures, ellipsoidal, helical and X-shape, that show outstanding compressive strength properties in the compression test. From this end, we include in the analysis as reference structures the ones discussed in Ref. (7) and take into account the results presented in Refs (8–10). Mechanical properties we analyse using finite element analysis (FEA) are presented in Ref. (11–15).

2. EXPERIMENTAL MODELING

In order to investigate the effect of pore size on the properties of OPLSs, we studied the properties of three structures with porosity in the range of 50%–60%, as shown in Fig. 1. These structures have been printed by the SLS method with polyamide (PA-2200) polymer and tested in compression tests. Results of the compression tests have been analysed by finite element modelling (FEM), which is nearly routinely used in analyses of mechanical properties of OPLSs – see for example Refs (15–31).

The SLS method used here for printing OPLSs is one of AM technologies (33,34) Its distinct characteristic is that particles of powdered substrate are sintered by a laser beam with no formation of liquid phase. This allows for the fabrication of uniform structures of highly complex architecture. Ultra-light weight elements can be printed of controlled porosity by SLS, which is of

prime importance for the investigations carried out in the present study.

2.1. Material

Polyamide (PA-2200), used in the present study, is one of the most commonly employed materials in SLS printing technology offered by EOS. This synthetic thermoplastic polymer has high biocompatibility, flexibility, hygroscopicity, good chemical resistance as well as high strength and hardness. Selected mechanical properties of PA 2200 are listed in Tab. 1.

Tab. 1. Selected mechanical properties of PA 2200 (35)

Mechanical properties	Value	Unit
Density	930	kg/m^3
Tensile strength	48	MPa
Tensile modulus	1,650	MPa
Strain at break	18	%
Melting temperature (20°C/min)	176	°C
Shore D hardness	75	—
Powder size	60	μm

2.2. Stereological parameters of designed structures

Three 3D lattice structures with distinctly different shapes of pores and degree of symmetry have been designed using CAD software SolidWorks. The size of the porosities was 1.9 mm for each lattice. The geometry of the designed lattices is depicted in Fig. 1.
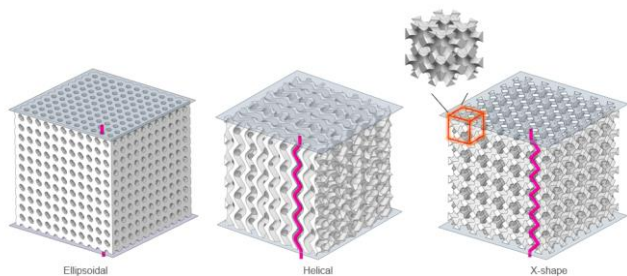
Fig. 1. Lattice structures investigated in this study and pore pathways along compression direction

Based on the shape of pores, one can be describe the structures presented in Fig. 1 as having cell units of:

- ellipsoidal,
- helical,
- X-shape.

Visually recognisable differences in the geometry of the three cell structures shown in Fig. 1 can quantified using principles of stereology – see for example Ref. (5). For cell structures like the ones analysed here, it is rational to focus attention on the pores, which account for more than 50% of their volume. Since pores are 3-dimensional features, their geometry can be quantified in terms of volume fraction, VV, size (e.g., equivalent diameter) and shape (using any of shape factors described in the literature). It should

be noted, however, that unlike in the case of close pores, size of pores in OPLSs from stereological point of view is an ill-defined parameter. This is because in OPLS there is one pore percolating throughout the structure of interest, and geometrical dimension of such pores depends on the size of structure. In this situation, pore surface to pore volume, SV, is used, the value of which does not depend on the physical dimensions of the open porosity structure of interest. It does not depend on the size of the structure. Its value increases with the decreasing distance in between the struts forming the lattice structure.

Stereological considerations show that parameter SV has can also be used to calculate a mean intercept length of porosity, i.e. average length of randomly oriented and positioned secants drawn across porosity in a given structure, $\bar{l}$, using the following stereological relationship:

$$\bar{l} = 2/S_V$$

Regarding shape of open porosity pores, it can be described by tortuosity; it is defined as the ratio of actual flow path length to the straight distance between the ends of the flow path. Taking into account symmetry of the structures of interest, we defined tortuosity in the way as shown in Fig. 1, i.e., by the ratio of actual vertical channel length to edge length.

The designed lattice structures were printed into cubes of dimensions 30 mm × 30 mm × 30 mm with similar porosity. In the analyses of their resistance to compression, applied force and average stress were normalised by porosity to set a stage for extracting a possible pore size effect. Values of the above-described stereological parameters for the designed structures are listed in Tab. 2.

Tab. 2. Stereological parameters of designed lattice structures

Structure	Porosity [%]	S_v [1/mm]	$\bar{l}$ [mm]	L/L_0
Ellipsoidal	54	3.62	7.24	1
Helical	50	2.22	4.44	1.2
X-shape	60	2.44	4.88	1.4

It can be noted from Tab. 1 that all the structures differ in pores volume fraction, which is highest for X-Shape, pore size, which is lowest for helical structure and shape/tortuosity, which is highest for ellipsoidal.

2.3. Compression tests and FEM

Compression tests of the lattice structures were carried out using the SHIMADZU 322 MTS Load Unit with a deflection rate of 1 mm/min, and the video extensometer ARAMIS 3D 4M was used to investigate the deformation. The recorded data were implemented by GOM Correlate software. Compression force-displacement curves were normalised to compression stress-compression strain ones. To this end, applied force was divided by surface of cube (30 mm × 30 mm) and induced displacement by cube edge length. Further experimental details can be found in a paper by M. Bernacka (6).

The compression behaviours of the printed lattice structures were analysed using Marc Mentat FEA software (36). Schematic explanation of modelling is given in Fig. 2. A tetrahedron element

type was utilised to generate the mesh of the lattice structures. The average number of elements was 250,000 for each structure. Compression of the structures was modelled by applying up to 20% reduction of height induced with rigid plates. FEM results were analysed in terms of macroscopic stress-strain relationships for three geometries of interest.
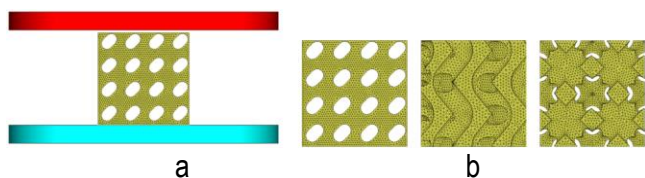
Fig. 2. Schematic explanation of: (a) compression test FEA model, (b) mesh view of each structures. FEA, finite element analysis

Nonlinear elastic-plastic properties evaluated of the printed structure were established through the curve a fitting method. The experimental results obtained in tensile tests of bulk coupled with PA 2200 as reported in Ref. (37). Results of this curve fitting exercise are shown in Fig. 3. A fully satisfactory agreement between experimental and modelled tensile curves was obtained, which rationalized the use of the numerical approximation of elastic-plastic properties obtained in further computations.

In FEA, we used 10 mm × 10 mm × 10 mm cells representative for lattice structures of interest. In modelling compression tests, these cells were placed between two rigid stamps.
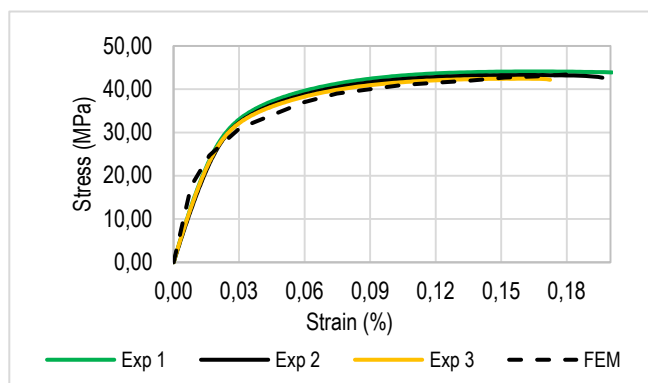
Fig. 3. Experimental and FEM tensile test curves. FEM, finite element modelling, Exp, experimental test

3. RESULTS

The stress-strain curves obtained in compression tests for the three structures investigated here are shown in Fig. 4 together with FEM modelling.

The results shown in Fig. 4 indicate that the helical lattice structure exhibits significantly higher resistance to compression than the other two structures. The compressive strength and stress were normalised by dividing the density of the lattice structure. Therefore, specific strength and specific stress were obtained (see Tab. 3). Based on the calculation, the helical structure has outstanding properties, in terms of specific stress 33,850 MPa/g/mm³ and specific strength 3,386 kN/g/mm³, compared with ellipsoidal and X-shape (see in Fig. 5).
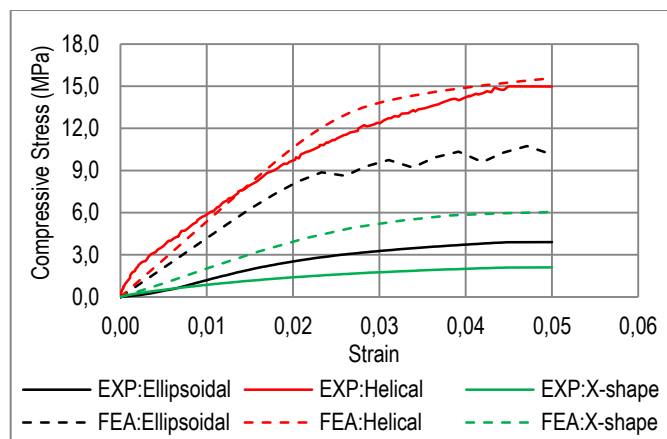
Fig. 4. Experimental compression stress-strain curves – solid lines, broken lines show results of FEA. FEA, finite element analysis and, EXP, experimental analysis

Tab. 3. Compressive stress, at 2% of deformation and volume of lattice structures

Structure property	Ellipsoidal	Helical	X-Shape
Volume (mm³)	408.161	495.531	390.222
Compressive stress at the strain 2% (MPa)	8.2	11.0	4.1
Compressive force at the strain 2% (kN)	0.8	1.1	0.4
Density of structures (gr/cm³)	0.38	0.46	0.36
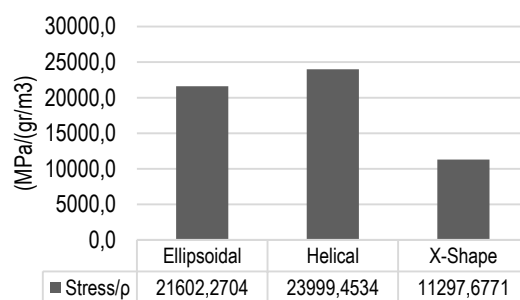
a
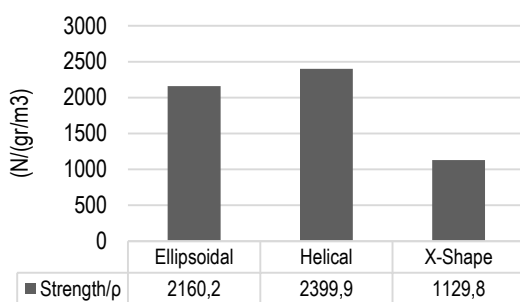
b

Fig. 5. (a) Specific stress and (b) specific strength of OPLSs.

An interesting observation can be made when analysing data in Fig. 4 that there is a major difference as far as agreement between the experimental values and the one obtained with FEM is concerned. FEM of compression of helical structure agrees well with experimental stress-strain curve. On the other hand, in the case of the other two structures, FEM grossly overestimated experimentally measured compression stress. Since FEM is an established tool for modelling deformation of composite materials, we consider the disagreement of numerical and experimental data for two OPLSs studied here as having background in physical properties and not as numerical error. In the following section, we provide explanation of these disparities.

4. DISSCUSION

One can discuss results presented here in terms of a relationship between stereological parameters of the investigated structures and their resistance to compression. Although with three structures differing in volume fraction, size and shape of pores, it is not possible to precisely determine character and parameters of the relationship governing their properties; there is a clear indication of the effect of pore size on the compression test. As can be noted, the ellipsoidal structure with the largest size of pores shows the lowest resistance to the compression stress.

To elucidate the effect of pore size on compression of the structures of interest, normalized compression stress for 20% deformation (σ') is plotted in Fig. 6 against $1/\sqrt{l}$, Fig. 6a, and $1/l$ Fig. 6b. These two plots are expected to linearise relationship between l and compression stress under two possible approaches. First, one can expect flow stress – size of pore relationship – as predicted by Hall-Petch who derived dependence of flow stress on the size of grains in metals. The other possible approach is based on mechanics of composites, under assumption that the struts of cellular structures can be as consisting of near surface layer of some physically defined thickness and the core. If this is the case, $1/\bar{l} = SV$ and determines the relative content of near surface zone.

It can be noted from the plot in Fig. 6 that both model relationships give reasonable agreement with experimental data. Obviously, with only three data points, it is difficult to draw conclusions; however, we suggest that the composite approach as shown in Fig. 6b gives a better agreement.

In the further discussion of the results presented in Fig. 4, we concentrate attention on two issues: (1) experimentally measured lower compression stress of elliptical and X-shape structures and (2) disparity between experimental and numerical results.

To explain differences in resistance to compression and the reasons for disagreement between experimental and modelled response to compression in the case of elliptical and X-shape OPLSs, we have analysed stress and strain distributions in the OPLS of interest and re-visited experimental detail of their compression tests.

Stress and strain distributions in the ellipsoid, helical and X-shaped structures computed with finite element method are shown in Fig. 7 and 8, respectively, for strain of 20%, (a reduction in height of the compressed cube). In analysing both stress and strain, we focused attention of homogeneity of their spatial distribution over the struts forming lattice structures.

a

b
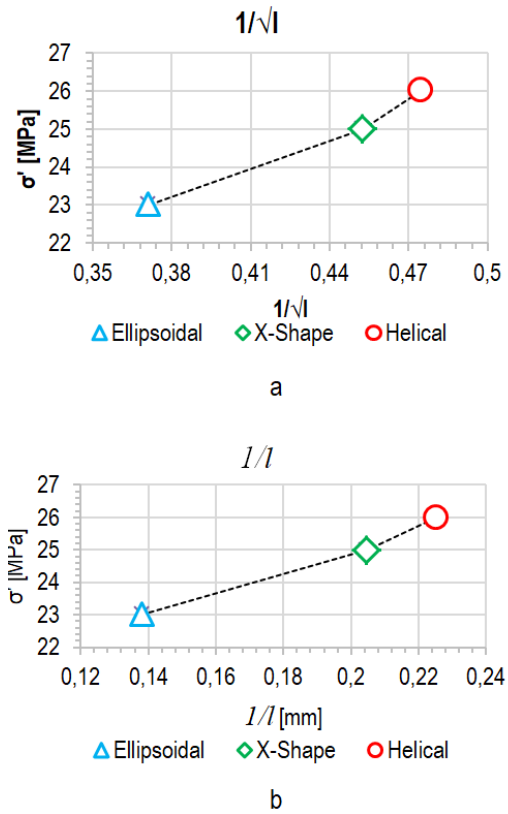
Fig. 6. Normalised stress with respect to average intercept length of porosity: (a) $1/\sqrt{\bar{l}}$ and (b) $1/\bar{l}$
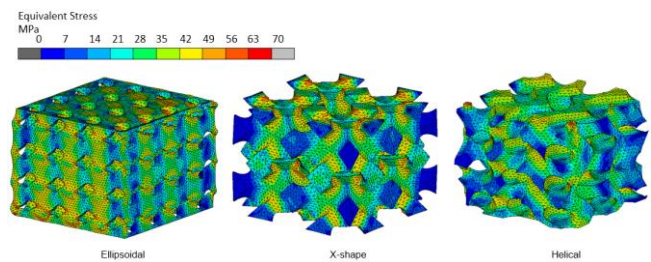
Fig. 7. Equivalent stress distributions for 20% of deformation
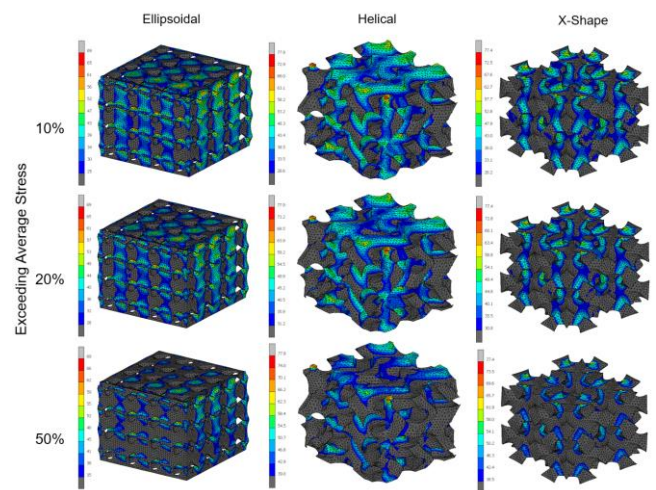
Fig. 8. Stress distribution by the value of equivalent stress exceeding the average equivalent by 10%, 20% and 50%

Graphical illustration (mapping) of equivalent stresses shown in Fig. 7 allows us to conclude that in the case of helical structures, stresses generated by compression are more uniformly distributed over the polymeric material forming the respective lattices. Helical type of cells increased the absolute stiffness and strength against the applied load ( ). One should also note that the stress concentrations are particularly high in the part of the structures in direct contact with load applying rigid stamps (see in Fig. 7).

In Fig. 8, the equivalent stress distribution on the OPLSs exceeding by 10%, 20% and 50% values of average stress and volume of the lattice elements carrying stress exceeding average value were calculated accordingly. The grey regions remained below the threshold stress values, and these regions were not included in the volume calculation.

These qualitative observations can be quantified by providing relative volume of the material with equivalent stress, exceeding the average equivalent stress by 10%, 20% and 50% values which are also listed in Fig. 9.

Based on the calculation of the volume that has the range of specified stress percentage, the concentrated stress per volume was plotted in Fig. 9. As can be seen from the data in the figure, elliptical and X-shape structures are characterised by significantly lower load carrying contribution of various parts of the structures.

The results of FEM of equivalent strain distribution are shown in Fig. 10. Generally, the same conclusions can be drawn with regard to load distribution over the materials of OPLSs analysed here. Visibly, strains are less uniform in ellipsoidal and X-shape structures, both in terms of inter-pore "bridges" being more strained and strain concentrations in the near-stamp zone.

Fig. 10 shows also images captured by Aramis video extensometer system during compression tests. These are images recorded for side-walls. As a result, in the case of elliptical structure, pores are clearly visible, unlike in the case the helical and X-shape ones. Thus, it was possible to investigate shape of some of the pores, e.g. the two indicated with arrows in the lower row in Fig. 10. One can clearly see that the pore next to the upper stamp

is much more elongated (deformed) in comparison with the other one. An important conclusion can be drawn that plastic deformation in elliptical structure is non-uniform and higher in the near-stamp zone. Other ARAMIS images confirm that the same applies to X-shape one, while in the case of helical structure, no localisation in near-surface zone was observed. This uniformity of strain distribution in the helical structure against strain localisation near to the compressed surface in the other two is the reason for the disparity between FEM modelling and experimental stress-strain curves shown in Fig. 4.

In explaining the disparity between experimental stress-strain curves and the ones obtained numerically, it should be noted that, experimental results obtained with ARAMIS indicate much higher strain localisation than estimated from FEM modelling. Thus, it concluded that within modelling carried out in the present paper, we can provide rational for propensity of elliptical and X-shape structures to strain localisation in the near to the compression stamp zone. However, we were not able fully capture in our models strain localization intensity.

FEM modelling was also used to discern details of stress distributions in the respective lattice structures under compression. The results presented in Fig. 11 and 12 concern stress vectors and flow lines, respectively. The stress vector distributions in Fig. 11 reveal that in the case of ellipsoidal and X-shaped structures, the stress vectors form a less regular pattern compared to that of the helical structure. Specifically, in the ellipsoidal structure, the stress vectors are concentrated in vertically oriented struts, leading to their buckling. In contrast, in the helical structure, the stress vectors are concentrated in horizontally oriented struts, with the concentration being perpendicular to the applied load. As a result of the pore shape in this structure, it can be noted that significant distortion or buckling-induced deformation does not occur, and stress concentrations are also minimised. In the X-shaped structure, the vectors are primarily perpendicular to the outer surface, with the highest concentration of stress vectors occurring at the point of highest strain, as seen in Fig. 11.

	Ellipsoid al	Ellipsoid al	Ellipsoid al	Helical	Helical	Helical	X-Shape	X-Shape	X-Shape
Volume Fraction (mm3)	246,52	221,00	148,01	353,35	322,82	228,72	125,86	97,44	38,16
Stress threshold (MPa)	25,3	27,6	34,5	28,6	31,2	39	28,2	30,8	38,5
Exceeding Stress Level (%)	10	20	50	10	20	50	10	20	50
Volume of Lattice (mm3)	408,161	408,161	408,161	495,531	495,531	495,531	390,222	390,222	390,222
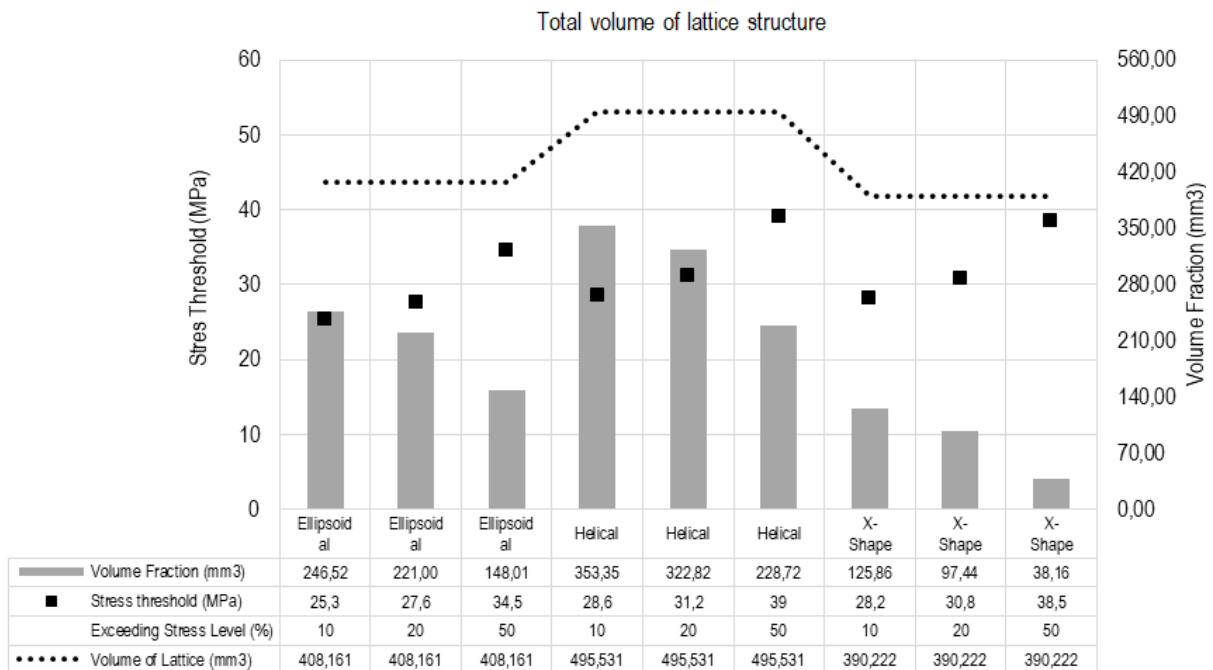
Fig. 9. The exceeding stress (exceeding average equivalent stress by 10%, 20% and 50%) per volume fraction
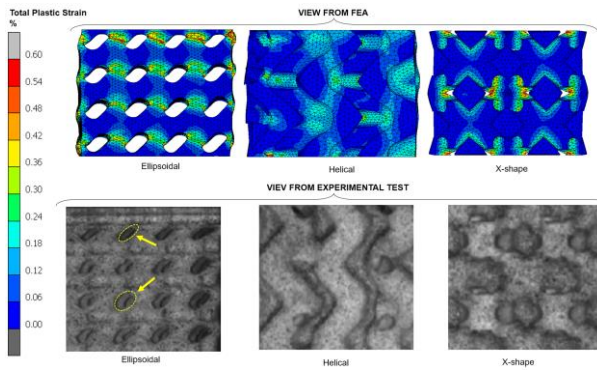
Fig. 10. Plastic strains distribution for 20% compression
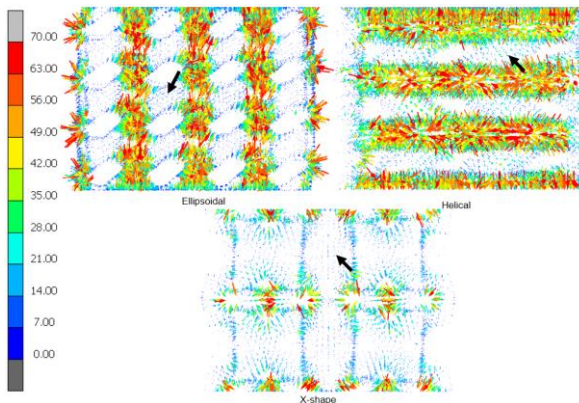
Fig. 11. Stress vectors, their direction and magnitudes of the structures

The material deformation flow patterns, depicted in Fig. 12, reveal that flow lines in helical and X-shape geometries are predominantly vertical, running from the top to the bottom of the structure. In the case of ellipsoidal structure, a distinct flow pattern is formed, with the pores altering the flow direction during loading. This deviation in the flow pattern may result in inhomogeneous deformation across the structure and may also lead to a change in the direction of loading, resulting in unsymmetrical deformation and a decrease in stiffness.
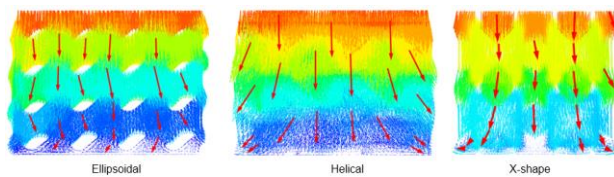
Fig. 12. Material flowlines during deformation

Results presented in Fig. 12 show similarity in flowline patterns in helical and X-shape structures. However, other analyses clearly indicate that X-shape differs from helical structures in terms of stress/strain and stress vector distribution.

5. CONCLUSIONS

The results obtained can be concluded in the following points, which are categorised into one based on stereological and micro-mechanical considerations.

Within the stereological approach to predicting properties of OPLSs, it has been demonstrated that:

— pore size has an important effect on compression stress of OPLSs

— outstanding stiffness of helical structures can also be attributed to the shape of pores

Within the micromechanics approach, the outstanding stiffness of helical OPLSs stems from much more uniform stress and strain distribution over the material of struts forming the respective structures. One may also note that in helical structures, stress vectors are well aligned and predominantly normal to the compression axis; hence, the presence of a helical structure amplified the overall stiffness and strength in response to applied loads.

An attempt to explain the disagreement between results of FEM modelling and experimental data brought attention to the phenomenon of strain localisation in the zone near to compression stamp, in fact captured with Aramis video extensometer system for monitoring progress in deformation in compression tests. Susceptibility of OPLS structure to localisation of strain is subject of a separate paper.

Results of the present study highlight importance of pore shape on the properties of high porosity structures. They also provide an insight into deformation of OPLSs and limitations in deformation modelling with FEM.

REFERENCES

1. Korkmaz ME, Gupta MK, Robak G, Moj K, Krolczyk GM, Kuntoğlu M. Development of lattice structure with selective laser melting process: A state of the art on properties, future trends and challenges. J Manuf Process. 2022 Sep 1;81:1040–63.
2. Tian L, Lu L, Chen W, Xia Y, Wang CCL, Wang W. Organic Open-cell Porous Structure Modeling. In: Symposium on Computational Fabrication [Internet]. New York, NY, USA: ACM; 2020. p. 1–12. Available from: https://dl.acm.org/doi/10.1145/3424630.3425414
3. Mora S, Pugno NM, Misseroni D. 3D printed architected lattice structures by material jetting. Mater Today [Internet]. 2022;59: 107–32. Available from: https://doi.org/10.1016/j.mattod.2022.05.008
4. Thirunavukkarasu N, Gao J, Peng S, Laroui A, Wu L, Weng Z. Mechanically robust 3D printed elastomeric lattices inspired by strong and tough hierarchical structures. Addit Manuf [Internet]. 2023;66(November 2022):103451. Available from: https://doi.org/10.1016/j.addma.2023.103451
5. Kurzydłowski KJ, Ralph B. The quantitative description of the microstructure of materials. CRC Press; 1995. 418 p.
6. Bernacka M, Aladag M, Dubicki A, Zgłobicka I. Effect of Pore Architecture of 3d Printed Open Porosity Cellular Structures on their Resistance to Mechanical Loading: Part I – Experimental Studies. Acta Mech Autom. 2024;18(3):419–26.
7. Montazerian H, Davoodi E, Asadi-Eydivand M, Kadkhodapour J, Solati-Hashjin M. Porous scaffold internal architecture design based on minimal surfaces: A compromise between permeability and elastic properties. Mater Des [Internet]. 2017 Jul;126(April):98–114. Available from: http://dx.doi.org/10.1016/j.matdes.2017.04.009
8. Li S, Yuan S, Zhu J, Zhang W, Tang Y, Wang C, et al. Optimal and adaptive lattice design considering process-induced material anisotropy and geometric inaccuracy for additive manufacturing. Struct Multidiscip Optim. 2022 Jan 1;65(1):1–16.
9. Barbaros I, Yang Y, Safaei B, Yang Z, Qin Z, Asmael M. State-of-the-art review of fabrication, application, and mechanical properties of functionally graded porous nanocomposite materials. Nanotechnol Rev [Internet]. 2022 Jan 1 [cited 2022 Oct 22];11(1):321–71. Available from: https://www.degruyter.com/document/doi/10.1515/ntrev-2022-0017/html

10. Liu R, Ma L, Liu H, Xu B, Feng C, He R. Effects of pore size on the mechanical and biological properties of stereolithographic 3D printed HAp bioceramic scaffold. Ceram Int. 2021 Oct 15;47(20):28924–31.

11. Jiao C, Xie D, He Z, Liang H, Shen L, Yang Y, et al. Additive manufacturing of Bio-inspired ceramic bone Scaffolds: Structural Design, mechanical properties and biocompatibility. Mater Des. 2022 May 1;217.

12. Cerardi A, Caneri M, Meneghello R, Concheri G, Ricotta M. Mechanical characterization of polyamide cellular structures fabricated using selective laser sintering technologies. Mater Des. 2013 Apr 1;46:910–5.

13. Cipriani CE, Ha T, Martinez Defilló OB, Myneni M, Wang Y, Benjamin CC, et al. Structure–Processing–Property Relationships of 3D Printed Porous Polymeric Materials. ACS Mater Au [Internet]. 2021 Sep 8 [cited 2022 Oct 22];1(1):69–80. Available from: https://pubs.acs.org/doi/full/10.1021/acsmaterialsau.1c00017

14. Zeleniakienė D, Kleveckas T, Liukaitis J, G. M. The Influence of Porosity on Stress and Strain State of Porous Polymer Materials. Mater Sci. 2003;9(4):358–62.

15. Chen X, Ji Q, Wei J, Tan H, Yu J, Zhang P, et al. Light-weight shell-lattice metamaterials for mechanical shock absorption. Int J Mech Sci [Internet]. 2020;169(August 2019):105288. Available from: https://doi.org/10.1016/j.ijmecsci.2019.105288

16. White DA, Kudo J, Swartz K, Tortorelli DA, Watts S. A reduced order model approach for finite element analysis of cellular structures. Finite Elem Anal Des [Internet]. 2023 Feb [cited 2022 Oct 26];214:103855. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0168874X22001287

17. Ni Y, Liao H, Zhao Q, Wu W, Shi Y, Wu S. Investigations of the failure behaviors of open-cell copper foam based on in-situ X-ray tomography compression experiments and image reconstructed finite element modeling. Eng Fract Mech. 2022 Mar 15;263:108323.

18. Ruiz de Galarreta S, Jeffers JRT, Ghouse S. A validated finite element analysis procedure for porous structures. Mater Des. 2020 Apr 1;189.

19. Abueidda DW, Elhebeary M, Shiang CS (Andrew), Pang S, Abu Al-Rub RK, Jasiuk IM. Mechanical properties of 3D printed polymeric Gyroid cellular structures: Experimental and finite element study. Mater Des. 2019 Mar 5;165:107597.

20. Smith M, Guan Z, Cantwell WJ. Finite element modelling of the compressive response of lattice structures manufactured using the selective laser melting technique. Int J Mech Sci. 2013 Feb 1;67: 28–41.

21. Zhong L, Li X. Simulation analysis of lightweight cylindrical lattice materials with different unit cells. J Coast Res. 2015 Mar; (73 (10073)):155–9.

22. Sadeghzade M, Gharehbaghi H, Farrokhabadi A. Experimental and analytical studies of mechanical properties of additively manufactured lattice structure based on octagonal bipyramid cubic unit cell. Addit Manuf. 2021 Dec 1;48:102403.

23. Chen W, Zheng X, Liu S. Finite-Element-Mesh Based Method for Modeling and Optimization of Lattice Structures for Additive Manufacturing. Mater 2018:23;11(11):2073.

24. Song J, Wang Y, Zhou W, Fan R, Yu B, Lu Y, et al. Topology optimization-guided lattice composites and their mechanical characterizations. Compos Part B Eng. 2019 Mar 1;160:402–11.

25. Leon-Patiño CA, Drew RAL. Role of metal interlayers in the infiltration of metal-ceramic composites. Curr Opin Solid State Mater Sci. 2005 Aug;9(4–5):211–8.

26. Lim YE, Park JH, Park K. Automatic Design of 3D Conformal Lightweight Structures Based on a Tetrahedral Mesh. Int J Precis Eng Manuf Technol 2018 54 [Internet]. 2018 Aug 23 [cited 2022 Oct 26];5(4):499–506. Available from: https://link.springer.com/article/10.1007/s40684-018-0053-2

27. Salonitis K, Chantzis D, Kappatos V. A hybrid finite element analysis and evolutionary computation method for the design of lightweight lattice components with optimized strut diameter. Int J Adv Manuf Technol 2016 909 [Internet]. 2016 Oct 21 [cited 2022 Oct 26];

90(9):2689–701. Available from: https://link.springer.com/article/10.1007/s00170-016-9528-x

28. Yang L, Yan C, Cao W, Liu Z, Song B, Wen S, et al. Compression–compression fatigue behaviour of gyroid-type triply periodic minimal surface porous structures fabricated by selective laser melting. Acta Mater. 2019 Dec 1;181:49–66.

29. Shin J, Kim S, Jeong D, Lee HG, Lee D, Lim JY, et al. Finite Element Analysis of Schwarz P Surface Pore Geometries for Tissue-Engineered Scaffolds. Math Probl Eng [Internet]. 2012;2012:1–13. Available from: http://www.hindawi.com/journals/mpe/2012/694194/

30. Gonzalez FJQ, Nuno N. Finite element modeling of manufacturing irregularities of porous materials. Biomater Biomech Bioeng [Internet]. 2016 Mar 25;3(1):1–14. Available from: http://koreascience.or.kr/journal/view.jsp?kj=E1TPEG&py=2016&vnc=v3n1&sp=1

31. Deliormanlı AM, Deliormanlı AH. Finite element method simulation for the prediction of mechanical properties of three-dimensional periodic bioactive glass scaffolds. J Aust Ceram Soc [Internet]. 2017 Oct 13;53(2):299–307. Available from: http://link.springer.com/10.1007/s41779-017-0037-7

32. Khrapov D, Koptyug A, Manabaev K, Léonard F, Mishurova T, Bruno G, et al. The impact of post manufacturing treatment of functionally graded Ti6Al4V scaffolds on their surface morphology and mechanical strength. J Mater Res Technol. 2020;9(2):1866–81.

33. Gardan J. Additive manufacturing technologies : State of the art and trends [Internet]. Additive Manufacturing Handbook. CRC Press; 2017 [cited 2022 Oct 26]. Available from: https://www.taylorfrancis.com/chapters/edit/10.1201/9781315119106-10/additive-manufacturing-technologies-julien-gardan

34. Ngo TD, Kashani A, Imbalzano G, Nguyen KTQ, Hui D. Additive manufacturing (3D printing): A review of materials, methods, applications and challenges. Compos Part B Eng. 2018 Jun 15;143:172–96.

35. PA 12 - PA2200: Nylon for Industrial 3D Printing | EOS GmbH [Internet]. [cited 2022 Oct 10]. Available from: https://www.eos.info/en/additive-manufacturing/3d-printing-plastic/sls-polymer-materials/polyamide-pa-12-alumide

36. MSC Software. Marc/Mentat 2020 Advanced Nonlinear FEA Software.

37. Bernacka ME. Bio-inspired design and experimental verification properties of porous elements obtained by 3D printing. [Bialystok]: Bialystok University of Technology; 2021.

38. Teng F, Sun Y, Guo S, Gao B, Yu G. Topological and Mechanical Properties of Different Lattice Structures Based on Additive Manufacturing. Micromachines. 2022;13(7).

Mehmet Aladag: https://orcid.org/0000-0002-2484-7519

Monika Bernacka: https://orcid.org/0000-0003-3481-0768

Adrian Dubicki: https://orcid.org/0000-0002-3994-2957

Izabela Zglobicka: https://orcid.org/0000-0002-4432-9196

Mohsen Soori, Fooad Karimi Ghaleh Jough, Roza Dastres, Behrooz Arezoo
Robotical Automation in CNC Machine Tools: A Review

DOI 10.2478/ama-2024-0048

ROBOTICAL AUTOMATION IN CNC MACHINE TOOLS: A REVIEW

Mohsen SOORI*, Fooad Karimi Ghaleh JOUGH*, Roza DASTRES, Behrooz AREZOO***

*Department of Civil Engineering, Final International University, AS128, Kyrenia, North Cyprus, Via Mersin 10, Turkey
**Department of Computer Engineering, Cyprus International University, North Cyprus, Via Mersin 10, Turkey
***CAD/CAPP/CAM Research Center, Department of Mechanical Engineering, Amirkabir University of Technology (Tehran Polytechnic),
424 Hafez Avenue, Tehran 15875-4413, Iran

mohsen.soori@final.edu.tr, fooad.karimi@final.edu.tr, roza.dastres@yahoo.com, arezoo@aut.ac.ir

Abstract: Robotics and automation have significantly transformed Computer Numerical Control (CNC) machining operations, enhancing productivity, precision, and efficiency. Robots are employed to load and unload raw materials, workpieces, and finished parts onto CNC machines. They can efficiently handle heavy and bulky components, reducing the demand of manual labour and minimizing the risk of injuries. Robots can also be used in CNC machine tools to perform tasks such as automatic tool changing system, part inspection, and workpiece positioning. Automation technologies, including in-line inspection systems and Non-Destructive Testing (NDT) methods, can be integrated into CNC machining cells to enhance accuracy and reduce scrap and rework in machining operations. These systems collect real-time data on process parameters and machine tool performance to predict maintenance, optimize machining parameters, and improve overall efficiency. In the current study, applications of robotics and automation in the modification of CNC machine tools are reviewed and discussed. Different applications of robotics and automation in CNC machine tools, such as automated material handling, automatic tool changing, robotic work cells, adaptive machining, machine tending, quality inspection, data monitoring and analysis, and production line integration, are discussed. Thus, by analysing recent achievements in published papers, new ideas and concepts of future research works are suggested. As a result, accuracy as well as productivity in the process of part production can be enhanced by applying robotics and automation in CNC machining operations.

Keywords: robotics, automation, CNC machine tools, automated material handling, automatic tool changin

1. INTRODUCTION

Robotics and automation play a crucial role in computer numerical control (CNC) machining operations, offering numerous advantages in terms of efficiency, precision, and productivity. Robotics and automation have revolutionized CNC machining operations and automated material handling, providing numerous benefits in terms of productivity, efficiency, and cost-effectiveness. Robots are integrated into CNC machining operations to perform various tasks such as part loading and unloading, tool changing, and even complex machining operations [1]. Industrial robots can be programmed to execute precise movements and repetitive tasks with high accuracy and speed, reducing the need for manual labour and improving overall productivity. They can work around the clock without fatigue, ensuring continuous and efficient operation [2]. This aspect enables manufacturers to produce complex components at a faster rate and with higher quality, contributing to the advancement of modern manufacturing industries [3, 4]. Advanced robotic systems equipped with sensors and artificial intelligence capabilities can perform adaptive machining [5]. These systems can monitor and adjust machining parameters in real time based on feedback from sensors, ensuring optimal performance, improved accuracy, and reduced scrap rates [6]. Adaptive machining enables efficient machining of complex parts and the ability to respond to variations in material properties. Automated material-handling systems are employed to streamline the movement of raw materials, workpieces, and finished parts within the CNC machining facility [7]. These systems utilise conveyors, gantries, robotic arms, and other automated mechanisms to transport materials between different workstations, CNC machines, and storage areas [8]. By eliminating manual handling, automated material-handling systems reduce the risk of errors, damage, and injuries, while also increasing throughput and minimizing production delays [9, 10]. Automation in CNC machining operations often involves the use of advanced sensors and identification systems. For example, barcodes, radio-frequency identification (RFID) tags, or QR codes can be attached to workpieces and materials to enable automated tracking throughout the manufacturing process [11]. This facilitates seamless integration with CNC machines and ensures accurate identification and routing of materials, reducing the chances of errors and mix-ups [12]. Here are some of the key advantages of incorporating robotics and automation in CNC machining operations:

- Increased productivity: Automation enables continuous operation of CNC machines, reducing downtime and maximising productivity. Robots can work tirelessly without breaks or fatigue, leading to higher output and increased efficiency [13].
- Enhanced precision and accuracy: Robots are highly precise and repeatable, ensuring consistent and accurate machining operations. They can achieve tight tolerances and eliminate human errors, resulting in improved part quality and reduced scrap or rework [14].
- Improved safety: Automation removes the need for manual intervention in hazardous or physically demanding tasks. This reduces the risk of accidents, injuries, and exposure to harmful environments, making the workplace safer for operators [6].

- Higher production speeds: Robots can perform CNC machining operations at significantly faster speeds compared to human operators. They can execute complex movements and tool changes quickly, resulting in reduced cycle times and increased production rates [15].
- Flexibility and adaptability: Robotic systems can be programmed to perform a wide range of tasks, allowing for greater flexibility in CNC machining operations. They can easily switch between different machining operations or workpieces, enabling efficient batch production or rapid product changeovers.
- 24/7 operation: Automated systems can run continuously, including outside regular working hours, without the need for human supervision. This maximizes machine utilisation and can lead to round-the-clock production, improving overall production capacity [16].

Also, disadvantages of using robotics in CNC machine tools can be presented as:

- High initial investment: Implementing robotics and automation in CNC machining operations requires a significant upfront investment. Costs include the purchase of robotic systems, integration with existing CNC machines, programming, and training. This may be a barrier for smaller businesses with limited budgets [14].
- Complex setup and programming: Setting up and programming robotic systems for CNC machining can be complex and time-consuming. Skilled personnel or specialized expertise may be required to program the robots accurately and optimize their performance. This can increase implementation costs and project timelines [17].
- Integration and programming: To implement robotics and automation in CNC machining operations, the integration of different systems is crucial. This includes programming the robotic systems, developing software interfaces for seamless communication between machines, and integrating control systems for coordinated operations. Additionally, specialised software tools enable offline programming and simulation, reducing downtime and optimizing the production process [18].
- Limited adaptability to small batch sizes or customisation: While automation excels in high-volume production, it may be less suitable for small batch sizes or highly customized products. Adapting automation systems to frequent product changes or small production runs can be challenging and may result in reduced efficiency [19].
- Potential for job displacement: Automation can lead to a reduction in manual labour requirements. As robots replace some human operators in CNC machining operations, there is a potential for job displacement or reduced employment opportunities for certain roles. However, automation can also create new job roles that focus on robot programming, maintenance, and system supervision [20, 21].
- Dependency on power supply and maintenance: Robotic systems require a stable power supply and regular maintenance to operate optimally. Power outages or equipment breakdown can disrupt production and lead to downtime. Maintaining robotic systems also requires skilled technicians and spare parts, which can add to operational costs [22].

Soori et al. [23] suggested virtual machining techniques to evaluate and enhance CNC machining in virtual environments [23-26]. To investigate and enhance performance in the component-production process employing welding procedures, Soori et al. [27] suggested an overview of current developments in friction-stir welding techniques. Soori and Asamel [28]s examined the implementation of virtual machining technology to minimise residual stress and displacement error throughout turbine blade five-axis milling procedures. Soori and Asmael [29] explored applications of virtualized machining techniques to assess and reduce the cutting temperature throughout milling operations of difficult-to-cut objects. Soori et al. [30] indicated an advanced virtual machining approach to improve surface characteristics throughout five-axis milling procedures for turbine blades. Soori and Asmael [31] created virtual milling processes to reduce the displacement error throughout five-axis milling operations of impeller blades. In order to analyse and develop the process of part production in virtual environments, virtual product development is presented by Soori [32]. Soori and Asmael [33] proposed an overview of current advancements from published research to review and enhance the parameter technique for machining-process optimisation. To improve the efficiency of energy consumption, the quality and availability of data across the supply chain, and the accuracy and dependability of component manufacture, Dastres et al. [34] proposed a review of the RFID-based wireless manufacturing systems. Soori et al. [35] explored machine learning and artificial intelligence in CNC machine tools to boost productivity and improve profitability in production processes of components, employing CNC machining operations. To improve the performance of machined components, Soori and Arezoo [36] reviewed the topic of measuring and reducing residual stress in machining operations. To improve surface integrity and decrease residual stress during Inconel 718 grinding operations, Soori and Arezoo [37] proposed the optimum machining parameters employing the Taguchi optimisation method. To increase the life of cutting tools during machining operations, Soori and Arezoo [38] examined different methods of tool wear-prediction algorithms. Soori and Asmael [39] investigated computer-assisted process planning to boost productivity in the part-manufacturing procedure. Dastres and Soori [40] addressed improvements in web-based decision-support systems to provide solutions for data warehouse management using decision-making assistance. Dastres and Soori [41] reviewed applications of artificial neural networks in different sections, such as analysis systems of risk, drone navigation, evaluation of welding, and evaluation of computer simulation quality, to explore the execution of artificial neural networks for improving the effectiveness of products. Dastres and Soori [42] proposed employing communication systems for environmental concerns to minimise the negative effects of technological advancement on natural catastrophes. To enhance network and data online security, Dastres and Soori [43] suggested the secure socket layer. Dastres and Soori [44] studied the developments in web-based decision-support systems for developing the methodology of decision-support systems by evaluating and suggesting the gaps between proposed approaches. To strengthen network-security measures, Dastres and Soori [45] discussed an analysis of recent advancements in network threats. To increase the potential of image-processing systems in several applications, Dastres and Soori [46] evaluated image processing and analysis systems. Dimensional, geometrical, tool deflection, and thermal defects have been modified by Soori and Arezoo [47] to improve the accuracy in five-axis CNC milling processes. Recent developments given in published articles are examined by Soori et al. [48] to assess and improve the impacts of artificial intelligence, machine learning, and deep learning in advanced robotics. Soori and Arezoo [49] developed a virtual machining system application to examine whether cutting parameters affect tool life and cutting

Mohsen Soori, Fooad Karimi Ghaleh Jough, Roza Dastres, Behrooz Arezoo
Robotical Automation in CNC Machine Tools: A Review

DOI 10.2478/ama-2024-0048

temperature during milling operations. Soori and Arezoo [50] studied the impact of coolants on the cutting temperature, roughness of the surface, and tool wear during turning operations with Ti6Al4V alloy. Recent developments from published papers are reviewed by Soori [51] to examine and alter composite materials and structures. Soori et al. [52] examined the Internet of Things (IoT) application for smart factories in Industry 4.0 to increase quality control and optimise part-manufacturing processes. To minimise cutting tool wear during drilling operations, Soori and Arezoo [53] designed a virtual machining system. Soori and Arezoo [54] decreased residual stress and surface roughness to improve the quality of items produced utilising abrasive water jet machining. To enhance accuracy in turbine blade five-axis milling operations, deformation errors are calculated and compensated by Soori [55]. To analyse and enhance accuracy in CNC machining operations and structures, applications of the finite element method in CNC machine tool modification are reviewed by Soori and Arezoo [56]. In order to analyse and optimise energy consumption in industrial robots, different methods of energy usage optimisation are reviewed by Soori et al. [57]. To analyse and modify the application of virtual manufacturing in productivity enhancement of part production, advanced virtual manufacturing systems are reviewed by Soori et al. [58]. Meta-heuristic algorithms for assessing the collapse risk of steel moment frame midrise buildings are presented by Karimi Ghaleh Jough and Şensoy [59] to provide a better risk-management strategy for steel moment frames. The steel moment-resisting frame dependability via interval analysis using the FCM-PSO method is studied by Karimi Ghaleh Jough and Şensoy [60] to enhance accuracy and decrease execution time in the calculation of siesmic fragility curves. Assessment of out-of-plane behaviour of non-structural masonry walls using FE simulations is presented by Karimi Ghaleh Jough and Golhashem [61] to reduce the self-weight axial compression of the walls with modern lightweight masonry units.

The paper reviews and discusses the application of robotics and automation in the modification of CNC machine tools. Robotics and automation are used in CNC machine tools for a variety of purposes, including integrated production lines, automated material handling, automatic tool changing, robotic work cells, adaptive machining, machine tending, and quality inspection. Thus, fresh thoughts and suggestions for the subsequent research tasks are proposed by examining previous successes in published articles. As a consequence, by implementing robots and automation in CNC machining processes, accuracy as well as productivity in the process of component manufacturing may be improved.

2. AUTOMATED MATERIAL HANDLING

The Robotics and automation play a crucial role in CNC machining operations and automated material handling. Robots can be employed to load and unload raw materials and finished parts onto CNC machines. They can pick up raw materials from a storage area, place them in the CNC machine for machining, and then transfer the finished parts for subsequent operations or storage [62]. This automation significantly reduces cycle times and enables lights-out machining, where operations continue without human intervention [63]. This eliminates the need for manual labour, reduces setup time, and ensures consistent and accurate placement of materials [64]. The integration of robotics and automation in CNC machining operations and automated material handling brings numerous benefits, including increased productivi-

ty, improved precision, reduced labour costs, enhanced safety, and better utilisation of resources [65]. It allows manufacturers to streamline their operations, optimise production flow, and respond quickly to changing demands in a competitive manufacturing environment [66]. How these technologies are used in these areas is explored below:

1. CNC machining operations:
- Robotic arms: Industrial robots equipped with articulated arms can perform various tasks in CNC machining operations. Robots can be programmed to handle different types of materials and workpieces, adapting to various machining requirements. They can handle the loading and unloading of workpieces into the CNC machines, change cutting tools, and perform quality inspections [67].
- Automated tool changing: CNC machines can be equipped with automatic tool changers (ATCs) that are controlled by computer programs. This enables the machine to switch between different cutting tools without manual intervention, reducing downtime and increasing efficiency [68].
- Continuous operation: Automated material-handling systems ensure a constant supply of raw materials and removal of finished products, minimising downtime [69].
- Optimised workflow: Robots can perform repetitive tasks with high precision and speed, allowing for a streamlined production process.
- Reduced errors: Automation minimises the chances of human errors, resulting in more precise machining and fewer defects [70].
- Energy efficiency: Optimised processes and reduced idle time in material handling contribute to energy savings, making the manufacturing process more cost-effective.
- Vision systems: Automated vision systems can be integrated with CNC machines to perform tasks such as part inspection, alignment, and measurement. These systems use cameras and image-processing algorithms to ensure accuracy and quality control [71].
- In-process monitoring: Sensors and probes can be incorporated into CNC machines to monitor the cutting process in real time. This allows for adaptive control, where the machine automatically adjusts parameters like cutting speed, feed rate, and tool path to optimise performance and prevent errors [72].

2. Automated material handling:
- Conveyor systems: Automated conveyor systems are commonly used in CNC machining operations for material handling. They transport workpieces, raw materials, and finished parts between different stages of the manufacturing process, reducing manual handling and improving efficiency [73].
- Automated guided vehicles (AGVs): AGVs are autonomous vehicles that can navigate within a manufacturing facility without human intervention. They can transport materials, such as raw stock and finished parts, between CNC machines, storage areas, and inspection stations [74].
- Palletising systems: Automated palletising systems are used to stack and organise workpieces or finished parts on pallets. These systems can handle heavy loads and precisely position the parts for efficient transportation and storage. This facilitates continuous machining by enabling the preparation of the next workpiece while the machine is still in operation [75].
- Robotic material handling: Industrial robots equipped with specialised end-effectors can handle materials, such as loading and unloading workpieces onto CNC machines, palletising

finished parts, or sorting and organising materials in a ware-house or storage area [76].

− Material tracking and management software: These systems can provide insights into production efficiency, predict mainte-nance needs, and optimise overall manufacturing processes.

Cloud robotics for material handling in cognitive industrial In-ternet of things is shown in Fig. 1 [77].
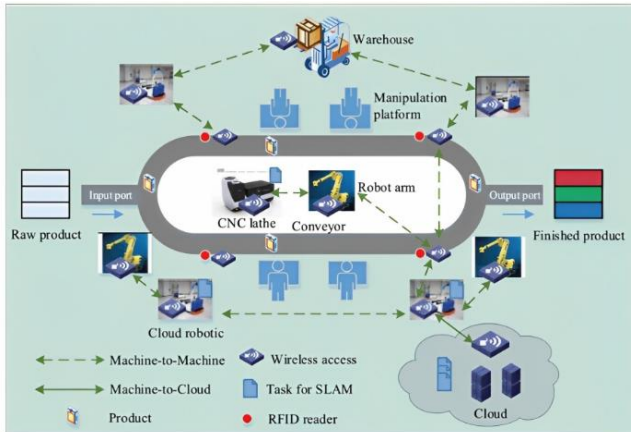
Fig. 1. Cloud robotics for material handling in cognitive industrial Inter-net of things [77]

Overall, robotics and automation have transformed CNC ma-chining operations and have automated material handling, en-hancing productivity, quality, and efficiency. These technologies continue to evolve, with advancements in artificial intelligence, machine learning, and collaborative robotics opening up new possibilities for the future of CNC machining [78]. The specific choice of automated material-handling solutions will depend on the type of CNC machine, the nature of the materials being pro-cessed, and the desired production goals. In the next section, the applications and benefits of automatic tool-changing systems in CNC machining operations will be discussed. The integration of robotic arms in tool changing not only streamlines the process but also extends the capabilities of CNC machines, offering dynamic adaptability and minimising human intervention.

3. AUTOMATIC TOOL CHANGING FROM CLASSICAL CHANGER TO ROBOTICAL CHANGER

CNC machines often require different tools for different opera-tions. Robots are employed to handle tooling tasks in CNC ma-chining centres. Automation in CNC machine tools, particularly in the context of tool changing, involves the use of various technolo-gies and systems to streamline the process of switching tools during machining operations. They can automatically load and unload cutting tools, reducing manual intervention and minimising machine downtime [79]. This results in increased productivity and improved machine utilisation. Robots can be employed to handle tooling tasks in CNC machining centres to enhance accuracy and productivity in machining operations [17, 80]. They can automati-cally load and unload cutting tools, reducing manual intervention and minimising machine downtime [81]. Some CNC machines are equipped with in-machine probing systems that can measure and verify tool dimensions without the need for manual intervention. Using such systems results in increased productivity and im-proved machine utilisation [82]. It allows for increased productivity, reduced downtime, and the ability to perform complex machining

tasks without human intervention [83]. An overview of how auto-mated tool changing works in CNC machining is given below:

− Tool magazine: The CNC machine is equipped with a tool magazine, which is essentially a storage unit for holding vari-ous cutting tools. The tool magazine can be located on the machine itself or as a separate unit adjacent to the machine [84].

− Tool identification: Each cutting tool is uniquely identified using a barcode, RFID, or some other form of identification system. This identification helps the machine recognise and select the appropriate tool for a specific machining operation [85].

− Tool selection: When a particular machining operation re-quires a tool change, the CNC machine's control system sends a command to the tool-changer mechanism to retrieve the required tool from the magazine. The command is usually based on the program being executed and the specific tool needed at that stage [86].

− Tool-changing mechanism: The CNC machine is equipped with a tool-changing mechanism, often referred to as an ATC. The ATC consists of a gripper or a robotic arm that can grasp and manipulate the cutting tools [80]. Linear tool changers are systems where the tools are arranged in a linear fashion, and a mechanism moves along the line to select and change tools. Rotary tool changers are mounted on a rotary carousel, and the carousel rotates to bring the desired tool into position [87].

− Tool retrieval and replacement: The tool-changer mechanism moves to the designated position in the tool magazine and re-trieves the required tool. It then moves to the spindle area where the previous tool is stored [88].

− Tool exchange: The ATC releases the current tool and grasps the new tool using its gripper or robotic arm. The tool-exchange process is usually automated and performed with precision to ensure proper alignment and secure attachment of the new tool [89].

− Tool calibration and verification: After the tool exchange, the CNC machine may perform calibration and verification pro-cesses to ensure the new tool is properly aligned and ready for use. This can include checking the tool length, diameter, and other parameters to ensure accurate machining [82].

− Resuming machining: Once the tool change is completed and verified, the CNC machine resumes the machining operation using the new tool. This process can be repeated multiple times during a machining operation, depending on the com-plexity and requirements of the workpiece [90].

Structure of an ATC of CNC machine tools is shown in Fig. 2 [91].
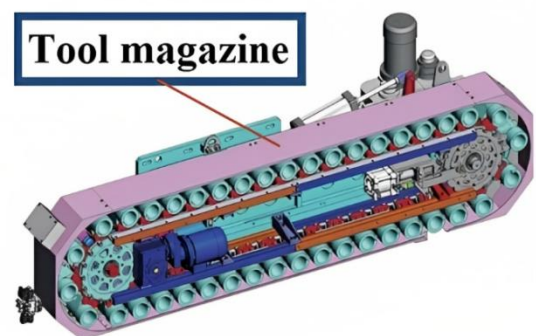
Fig. 2. Structure of an automatic tool changer of computer numerical control machine tools [91]

Mohsen Soori, Fooad Karimi Ghaleh Jough, Roza Dastres, Behrooz Arezoo
Robotical Automation in CNC Machine Tools: A Review

DOI 10.2478/ama-2024-0048

Application of robots in the vertical machining centre tool changer is shown in Fig. 3 [92].
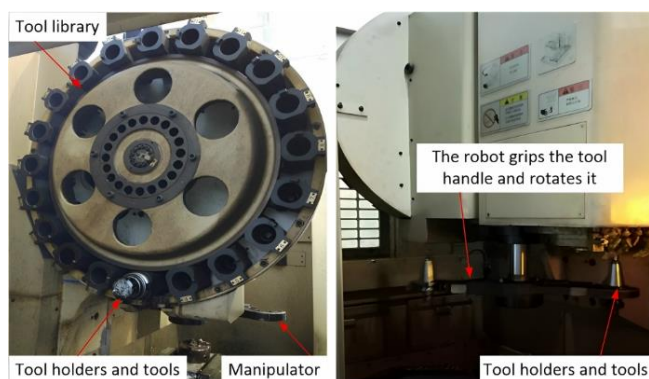
Fig. 3. Application of robots in the vertical machining centre tool changer [92]

The ATC significantly reduces manual intervention and minimises downtime associated with manual tool changes. It enables CNC machines to perform complex machining tasks with multiple tools seamlessly increasing productivity and efficiency in manufacturing processes [93]. Robotic work cells can efficiently manage tool changes by coordinating with the automatic tool-changing systems, ensuring a smooth transition between machining operations. In the next section, applications of robotic work cells in the automation of CNC machine tools are discussed.

4. ROBOTIC WORK CELLS

Robotics and automation play a significant role in CNC machining operations, particularly in the form of robotic work cells. These advanced systems combine CNC machines with industrial robots to enhance the productivity, efficiency, and flexibility in manufacturing processes [94, 95]. Robotic work cells in the context of CNC machine tools refer to integrated systems where robots are deployed to perform various tasks alongside or in collaboration with CNC machines. Robotic work cells combine multiple machines, such as CNC machines, measuring systems, and assembly stations, with robotic arms [96]. These cells enable seamless integration and coordination between different processes, optimising production flow and minimising cycle times [97]. Here are some key aspects and benefits of robotic work cells in CNC machining:

- Increased productivity: Robotic work cells can operate continuously without breaks, leading to increased productivity and reduced cycle times. Robots can perform repetitive tasks with high accuracy and speed, minimising human errors and maximising output [98].
- Flexibility and adaptability: Robots are programmable and can be easily reprogrammed or reconfigured to handle different machining tasks or part variations. This flexibility enables manufacturers to efficiently switch between different product lines or adapt to changing production requirements [99].
- Enhanced precision and consistency: CNC machines provide precise control over machining operations, and when combined with robots, they ensure consistent and repeatable results. This level of accuracy is crucial for industries like aerospace and automotive, where tight tolerances are required [100].

- Improved safety: Robots can handle hazardous or physically demanding tasks, reducing the risk of injuries to human workers. They can operate in enclosed work cells or behind safety barriers, safeguarding operators from potential accidents associated with machining processes [101].
- Offline programming and simulation: Robots and CNC machines can be programmed offline using simulation software [102]. This allows for the optimisation of robotic movements and CNC machining processes without interrupting actual production [103].
- Reduced labour costs: Robotic work cells can replace manual labour for routine machining operations, leading to cost savings in terms of labour expenses. While human operators are still needed for tasks like programming and supervision, the overall labour requirement can be significantly reduced [104].
- Lights-out manufacturing: With robotic work cells, it is possible to achieve lights-out manufacturing, where production can continue unattended even during non-working hours. This can optimise machine utilisation and increase overall production capacity [105].
- Integration with other automation technologies: Robotic work cells can be integrated with other automation technologies, such as conveyor systems, part feeders, vision systems, and quality control devices. This integration streamlines the production process, improves material flow, and enhances overall system efficiency. Moreover, it can facilitate real-time monitoring, scheduling, and optimisation [106].
- Predictive maintenance: Robotics and CNC systems can be equipped with sensors for predictive maintenance. This helps in identifying potential issues before they lead to downtime, reducing unplanned interruptions in production [16, 107].
- Data collection and analysis: Robotics and automation systems in CNC machining often come with advanced data-collection capabilities. By collecting real-time data on machine performance, tool wear, and part quality, manufacturers can analyse and optimise their processes for improved efficiency and predictive maintenance [108].

KUKA milling robot is shown in Fig. 4 [109].
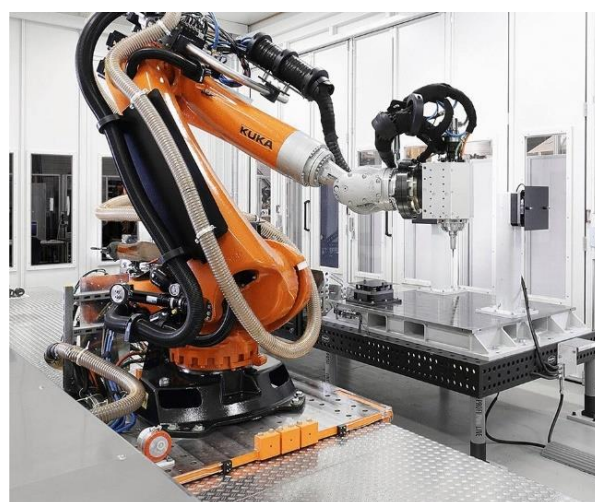
Fig. 4. KUKA milling robot [109]

The implementation of robotic work cells in CNC machine tools is a key component of Industry 4.0, where smart manufacturing technologies converge to create more agile and responsive

production systems [110, 111]. Robotic work cells can be part of flexible manufacturing systems where CNC machines and robots work together to handle a variety of tasks. Also, robots can be employed to load raw materials onto CNC machines and unload finished products, streamlining the material-handling process [112]. While the benefits of robotic work cells in CNC machining operations are substantial, their implementation requires careful planning, programming, and maintenance. It is essential to consider factors like workspace layout, robot programming, safety measures, and proper training for operators to ensure successful integration and operation of robotic work cells [113]. Adaptive machining using robots and automation introduces a dynamic approach to manufacturing, allowing for real-time adjustments in machining processes based on changing conditions, which is discussed in the next section.

5. ADAPTIVE MACHINING USING ROBOTS AND AUTOMATION

Adaptive machining in CNC machining operations refers to the application of robotics and automation to dynamically adjust the machining process based on real-time feedback and data. It involves the integration of sensors, control systems, and intelligent algorithms to optimise machining operations for improved efficiency, accuracy, and productivity [114]. Advanced robotic systems are equipped with sensors and vision systems that enable them to adapt to variations in workpiece dimensions or alignment [115]. Sensor data can be used to adapt machining strategies dynamically, compensating for variations in workpiece geometry, material properties, and tool wear. This capability allows for real-time adjustments in machining parameters, ensuring precise and accurate results, even with slight variations in the workpiece [16]. Some key aspects of adaptive machining in CNC operations are given below:

- Real-time sensing: Adaptive machining relies on sensors such as force sensors, acoustic emission sensors, temperature sensors, and vision systems to collect data during the machining process. These sensors provide feedback on factors like cutting forces, tool wear, workpiece condition, and dimensional accuracy [116].
- Data analysis and interpretation: The collected sensor data is analysed using advanced algorithms and machine learning techniques to extract valuable insights. This analysis helps in identifying patterns, anomalies, and deviations from the desired machining parameters [117].
- Decision-making and control: Based on the data analysis, adaptive machining algorithms make real-time decisions to optimise the machining process. These decisions can include adjusting feed rates, changing tool paths, modifying cutting parameters, or replacing worn-out tools [118].
- Dynamic adjustments: Adaptive machining systems use robotics and automation to implement the necessary adjustments identified through the data analysis. This can involve automatically repositioning the workpiece, changing cutting tools, adjusting spindle speeds, or modifying tool paths to ensure optimal machining conditions [119].
- Optimisation objectives: Adaptive machining aims to achieve various optimisation objectives, including reducing cycle times, improving surface finish quality, minimising tool wear, maximising tool life, maintaining dimensional accuracy, and reducing energy consumption [120].

Benefits of adaptive machining in CNC operations:

- Improved efficiency: By dynamically adjusting machining parameters, adaptive machining optimises the process to reduce cycle times, minimise material waste, and increase productivity [121].
- Enhanced accuracy and quality: Real-time adjustments based on sensor feedback help maintain dimensional accuracy, improve surface finish quality, and reduce errors in CNC machining operations [122].
- Extended tool life: Adaptive machining systems detect tool wear in real time and make necessary adjustments, leading to longer tool life and reduced tooling costs. Also, adaptive machining systems can continuously monitor and adjust cutting parameters based on real-time data, optimising tool paths for improved accuracy and cutting tool life during machining operations.
- Reduced downtime: By monitoring the machining process continuously, adaptive machining systems can detect anomalies and potential issues, allowing for proactive maintenance and reducing unplanned downtime [123].
- Reduced scrap and rework: Adaptive systems can adapt cutting speeds, feeds, and other parameters to optimise machining and minimise scrap [124].
- Flexibility and adaptability: Adaptive machining enables CNC machines to handle variations in workpiece material properties, tool wear, and other factors, making the process more adaptable to changing production requirements [125].

Block diagram of robotic belt grinding trajectory planning steps is shown in the Fig. 5 [126].
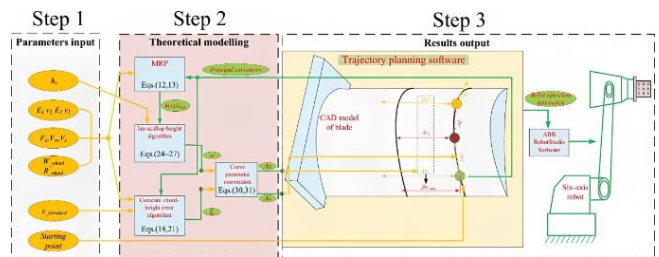
Fig. 5. Block diagram of robotic belt grinding trajectory planning steps [126]

Overall, adaptive machining utilising robotics and automation plays a crucial role in optimising CNC machining operations and improving efficiency, accuracy, and productivity, while reducing costs and downtime [127]. By integrating robots and automation into CNC machine tools, manufacturers can achieve a more adaptive and responsive manufacturing environment, leading to improved overall efficiency, product quality, and competitiveness in the market.

6. MACHINE TENDING

Machine tending refers to the process of loading and unloading workpieces to and from CNC machines. Robotics and automation have revolutionised CNC machining operations, particularly in the area of machine tending. Robots can be used to tend multiple CNC machines simultaneously, optimising production throughput [128]. They can load and unload workpieces, initiate machining processes, and perform inspections or measurements. Automation in machine tending aims to replace or assist human operators

Mohsen Soori, Fooad Karimi Ghaleh Jough, Roza Dastres, Behrooz Arezoo
Robotical Automation in CNC Machine Tools: A Review

DOI 10.2478/ama-2024-0048

in these repetitive and often labour-intensive tasks. This process is critical for maximising the efficiency of CNC machining operations and reducing manual labour [129]. An overview of a machine-tending system is shown in Fig. 6 [130].
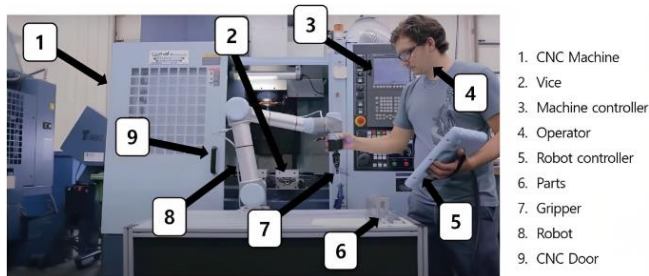
Fig. 6. Overview of a machine-tending system [130]

By integrating robotics and automation into machine-tending tasks, several advantages can be achieved:

- Increased productivity: Automation eliminates the need for human operators to manually load and unload workpieces, allowing the CNC machines to operate continuously without interruptions. This leads to higher productivity and reduced cycle times [131].
- Enhanced safety: CNC machines can be hazardous to operate, especially during the loading and unloading of heavy workpieces. By employing robots, human operators can be kept at a safe distance from the machine, minimising the risk of accidents and injuries [132].
- Improved accuracy and consistency: Robots are capable of precise movements and can consistently position workpieces with high accuracy. This ensures consistent and repeatable machining results, reducing errors and scrap rates.
- Flexibility and adaptability: Robots can be programmed to handle various types of workpieces and can quickly switch between different tasks. This enables manufacturers to respond to changing production needs, such as varying product designs or batch sizes, without significant retooling or reprogramming [133].
- Extended machine uptime: Automated machine tending allows CNC machines to operate continuously, even outside regular working hours. This maximises machine uptime and overall production capacity, leading to increased efficiency and reduced idle time.
- Reduction in labour costs: By automating machine-tending tasks, manufacturers can reduce their dependency on manual labour, resulting in cost savings associated with labour wages, training, and employee benefits [134].
- Integration with other processes: Robotic machine tending can be seamlessly integrated with other automation processes, such as material handling, quality inspection, and postprocessing. This holistic automation approach further optimises the entire production workflow [135].

To enhance the efficiency of the machine-tending system in advanced CNC machining operations, the application of the digital twin is studied. The flowchart of a developed machine-tending system using the digital twin system is shown in Fig. 7 [130].

To implement robotics and automation in machine tending, various technologies are employed, including industrial robots, vision systems, sensors, and advanced programming techniques. Additionally, collaborative robots (cobots) are gaining popularity, as they can work safely alongside human operators, further increasing the flexibility and versatility of machine-tending operations [136]. Overall, machine tending by automation in CNC machine tools plays a crucial role in enhancing the efficiency, safety, and precision in manufacturing processes. Robotics and automation have significantly transformed CNC machining operations, making them more efficient, safe, and adaptable to evolving manufacturing demands [137]. In smart CNC machining operations, robotics, artificial intelligence, and machine learning are applied to enhance the capabilities of automated machine-tending systems.
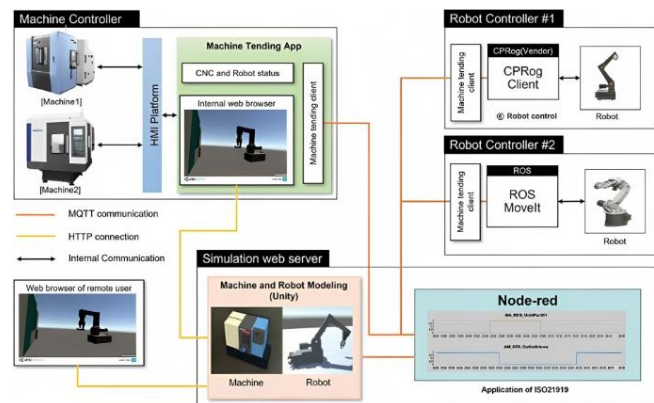
Fig. 7. The flowchart of a developed machine-tending system using the digital twin system [130]

7. QUALITY INSPECTION

Robotics and automation have significantly impacted CNC machining operations, particularly in the area of quality inspection. Automated systems can be integrated with CNC machines to inspect and measure machined parts using sensors, vision systems, or coordinate measuring machines (CMMs). This ensures that parts meet specified tolerances, reduces human error, and improves overall quality control [138]. Implementing automation in CNC machine tools for quality inspection not only improves the efficiency and accuracy of the inspection process but also contributes to reducing the likelihood of defects and increases overall productivity [139]. Some ways in which robotics and automation are used in CNC machining quality inspection are given below:

- Automated measurement systems: Robots are often integrated with automated measurement systems to perform precise and accurate measurements of machined parts. These systems use advanced sensors, such as laser scanners or CMMs, to capture dimensional data. The robots can move the sensors along programmed paths to inspect critical features of the machined parts [140].
- Vision systems: Vision systems, including cameras and image-processing algorithms, are commonly employed in CNC machining quality inspection. Robots equipped with vision systems can capture images of machined parts and analyse them to detect defects, surface-finish irregularities, or dimensional deviations. They can compare the captured images with reference models to ensure that the parts meet the required specifications [141].
- Automated sorting and packaging: Once parts are inspected and meet the quality criteria, automated systems can be employed for sorting and packaging, reducing the risk of human error and ensuring consistency [142].

- In-process monitoring: Robotics and automation enable real-time monitoring of CNC machining operations. Sensors and probes can be integrated into the machining process to collect data on parameters like tool wear, cutting forces, temperature, and vibration. These data are then analysed by automated systems to detect anomalies and ensure that the machining process is within the desired parameters, ultimately enhancing the quality of the machined parts [143].

- Non-destructive testing (NDT): Robotic systems can be used to perform NDT on machined parts. For example, automated ultrasonic or eddy current testing systems can be employed to inspect the integrity of critical components. The robots can position the testing equipment precisely and perform scans according to pre-defined paths, allowing for efficient and reliable quality inspections [144].

- Defect identification and sorting: Robots equipped with machine vision and robotic arms can identify defective parts and sort them accordingly. Once defects are detected during the quality-inspection process, the robots can remove or separate the faulty parts from the production line, ensuring that only high-quality components are delivered [145].

- Data analysis and feedback loop: Automation systems can collect and analyse vast amounts of data generated during CNC machining operations. By using machine-learning algorithms, patterns and trends can be identified, enabling predictive maintenance, process optimisation, and continuous improvement in quality inspection [146].

- Statistical process control: Automation can be applied to implement statistical process-control methods, monitoring key process parameters and ensuring that the machining process operates within specified tolerances.

Probing tool of robotic arms for the quality control is shown in Fig. 8 [147].

Overall, the integration of robotics and automation in CNC machining quality inspection brings several benefits, including increased accuracy, efficiency, and reliability. These technologies allow for faster inspections, reduced human error, and enhanced process control, ultimately leading to improved product quality and higher customer satisfaction. It is important to carefully design and integrate these automation solutions based on the specific requirements of the CNC machining operations and the desired quality standards.

Fig. 8. Probing tool of robotic arms for quality control [147]

8. DATA MONITORING AND ANALYSIS

Robotics and automation have significantly transformed the CNC machining industry by enabling enhanced data monitoring and analysis capabilities. These advancements have resulted in improved productivity, accuracy, and efficiency in machining operations [148]. Automation systems can collect real-time data from CNC machines, such as cutting parameters, machine performance, and tool wear. These data can be analysed to optimise processes, predict maintenance requirements, and improve overall efficiency [149]. Continuous improvement based on data-driven insights is key to achieving higher levels of efficiency and productivity. How robotics and automation have impacted data monitoring and analysis in CNC machining is given below:

- Real-time data collection: Automated systems integrated with CNC machines can collect real-time data during machining operations. These data include information such as cutting speeds, tool wear, temperatures, vibrations, and other relevant parameters. Robots and sensors can be used to capture these data accurately and consistently [150].

- Data integration and connectivity: Automation allows for seamless integration and connectivity between CNC machines and data-monitoring systems. The collected data can be transmitted to centralised databases or cloud platforms for storage and analysis. This connectivity enables real-time monitoring, remote access, and analysis of machining data from anywhere, facilitating timely decision-making [16].

- Condition monitoring and predictive maintenance: By analysing the collected data, advanced algorithms and machine-learning techniques can identify patterns and anomalies related to the machine condition. This enables predictive maintenance, where potential issues can be detected early and maintenance actions can be scheduled proactively. Also, remote monitoring and control of CNC machines can be implemented using data monitoring and analysis. Machine-learning algorithms can be used to predict tool wear, optimise cutting parameters, and improve overall machining efficiency. This approach minimises unplanned downtime and optimises machine availability [148].

- Performance optimisation: Data monitoring and analysis can help identify inefficiencies in machining processes. By examining the collected data, manufacturers can analyse factors such as tool paths, cutting parameters, and material characteristics to optimise machining performance. This leads to improved cycle times, reduced scrap rates, and enhanced overall productivity [151].

- Quality control and process improvement: Data analysis enables manufacturers to monitor and control product quality in real time. By comparing machining data with pre-defined quality parameters, any deviations or defects can be quickly identified, allowing for immediate corrective actions. Continuous data analysis also provides insights for process improvement, enabling manufacturers to refine their machining strategies and achieve higher quality standards [152].

- Data-driven decision-making: The availability of accurate and timely data empowers manufacturers to make informed decisions. Through data analysis, manufacturers can identify bottlenecks, optimise workflows, allocate resources effectively, and identify opportunities for cost reduction and process optimisation. This data-driven decision-making approach enhances overall operational efficiency [153].

- Traceability and compliance: Automated data monitoring and analysis provide a comprehensive record of machining operations, including the parameters used, measurements taken, and quality checks performed. This traceability is valuable for regulatory compliance, quality audits, and product validation [154].

Applications of digital twin and big data analysis in cloud-based manufacturing systems are shown in Fig. 9 [155].
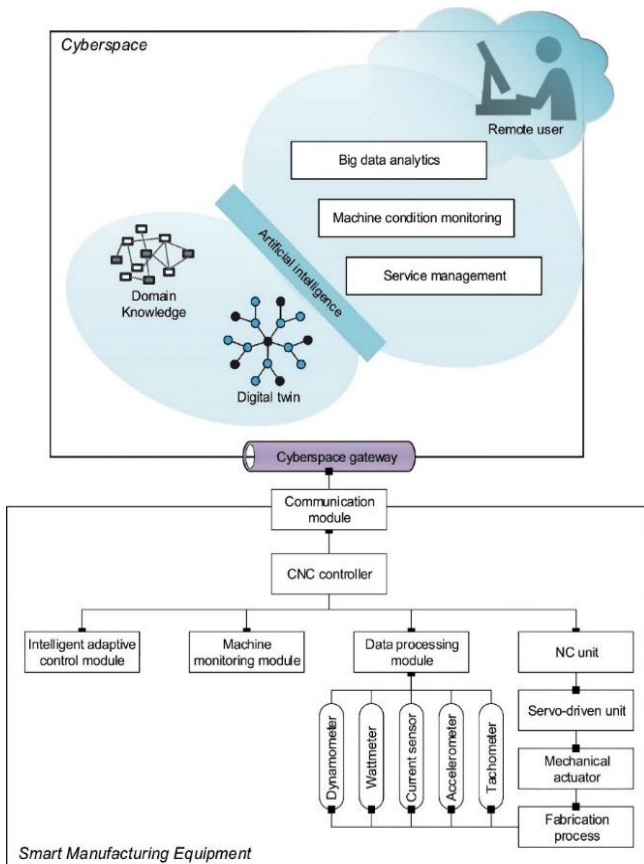
Fig. 9. Applications of digital twin and big data analysis in cloud-based manufacturing systems [155]

In summary, robotics and automation have revolutionised data monitoring and analysis in CNC machining operations [156]. Real-time data collection, connectivity, predictive maintenance, performance optimisation, quality control, data-driven decision-making, and traceability are some of the significant benefits offered by these advancements. The obtained data from CNC machines can be integrated with manufacturing execution systems to provide a comprehensive view of the entire manufacturing process. As a result, these capabilities lead to improved productivity, enhanced quality, and increased efficiency in CNC machining processes [14]. However, implementing a comprehensive data monitoring and analysis system for automation in CNC machine tools requires a multi-disciplinary approach involving technology integration, data science, and a thorough understanding of the machining process.

9. COLLABORATIVE ROBOTICS

Robotics and automation have revolutionized CNC machining operations, making them more efficient, precise, and safe. Collaborative robotics, specifically, has played a significant role in enhancing the capabilities of CNC machines and improving human–robot interactions [157]. Collaborative robots, or cobots, can work alongside human operators in CNC machining operations. These robots are designed to be safe to work with and can assist in tasks such as part handling, deburring, or cleaning, thereby enhancing productivity and relieving workers from repetitive or hazardous tasks [158]. They are equipped with sensors and advanced control systems that allow them to detect and respond to the presence of humans, ensuring a safe working environment. This collaboration enables a more flexible and adaptive manufacturing process where humans and robots can work together on different tasks, with the robot handling repetitive or strenuous tasks, while humans focus on more complex and cognitive aspects [159]. This integration leads to improved efficiency, reduced costs, and a safer and more collaborative working environment [160]. Some key aspects of robotics and automation in CNC machining operations, with a focus on collaborative robotics, are given below:

- Automated material handling: Collaborative robots (cobots) are used to automate material-handling tasks in CNC machining operations. They can pick up raw materials, place them in the machine, and remove finished parts. Cobots are equipped with sensors and vision systems that allow them to detect and grasp objects safely, enabling efficient material handling without the need for physical barriers or safety cages [160].
- Machine tending: Cobots are commonly employed for machine-tending tasks in CNC machining. They can load and unload workpieces, set up the machine, and initiate the machining process. Cobots work alongside human operators, taking over repetitive and physically demanding tasks, while humans focus on more complex operations, such as programming and quality control [161].
- Safety features: Traditional industrial robots are often separated from human workers by physical barriers for safety reasons. Collaborative robots are designed with safety features that enable safe interaction with human workers. They are equipped with force-sensing technologies that allow them to detect human presence and apply force accordingly [162]. This ensures that if a human comes into contact with the robot, it will stop or slow down to prevent injuries. Additionally, cobots have rounded edges and lightweight construction to minimise the risk of harm during accidental collisions [163].
- Programming and flexibility: Collaborative robots in CNC machining operations are programmed using intuitive interfaces. This simplifies the programming process, making it accessible to non-experts. Operators can teach cobots tasks by physically guiding their movements or by using graphical programming interfaces. This flexibility allows quick reprogramming and reconfiguration of the robot for different tasks, making them highly adaptable to changing production requirements [164].
- Enhanced precision and quality: Robotics and automation improve the precision and quality of CNC machining operations. Collaborative robots can perform tasks with high repeatability, ensuring consistent results. They can execute complex movements and follow precise paths, resulting in improved machining accuracy and reduced errors. Additionally, cobots can integrate with measurement systems to perform in-process inspections, enhancing quality control throughout the machining process [165].
- Easy integration: Cobots are designed to be easily integrated into existing workflows and systems. They can be programmed to work in tandem with CNC machine tools, handling tasks such as material loading and unloading, tool changes, and part inspection.

- Increased productivity: By automating repetitive and time-consuming tasks, collaborative robots increase productivity in CNC machining operations. Cobots can work around the clock, reducing machine idle time and maximising production efficiency. They can perform tasks with high speed and accuracy, resulting in shorter cycle times and increased output [166].
- Human–machine collaboration: Cobots allow for a closer collaboration between human workers and machines. While the cobot takes care of routine and physically demanding tasks, human workers can focus on more complex activities that require decision-making, problem-solving, and creativity.
- Maintenance and diagnostics: Collaborative robots can be programmed to perform routine maintenance tasks on CNC machines, such as cleaning, lubricating, or even diagnosing simple issues. This proactive maintenance approach helps prevent unplanned downtime and ensures the longevity of CNC machine tools [167].
- Adaptive machining: Cobots can be integrated into the CNC machining processes to adapt to changes in the production environment. They can dynamically adjust their movements and tasks based on real-time feedback, improving overall system flexibility [168].
- Workforce augmentation: Collaborative robotics in CNC machining operations do not replace human workers but rather augment their capabilities. By taking over mundane and physically demanding tasks, cobots free up human operators to focus on higher-level activities that require creativity, problem-solving, and decision-making. This leads to a more skilled and engaged workforce [169].

Conceptual design of a collaborative robot for drilling modelled after human operation is shown in Fig. 10 [170].
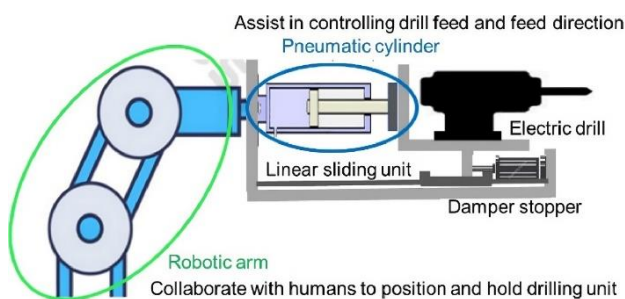
Fig. 10. Conceptual design of collaborative robot for drilling modelled after human operation [170]

In summary, collaborative robotics has transformed CNC machining operations by automating material handling, machine tending, and other tasks. Collaborative robotics enhances automation in CNC machine tools by combining the precision of CNC machining with the flexibility and adaptability of robots. This integration leads to improved efficiency, reduced costs, and a safer and more collaborative working environment [171].

10. PRODUCTION LINE INTEGRATION USING ROBOTICS AND AUTOMATION

Robotics and automation play a crucial role in CNC machining operations and production line integration. They offer numerous benefits such as increased productivity, improved precision, enhanced safety, and reduced labour costs [172]. Automation systems enable the integration of multiple CNC machines into a cohesive production line [173]. Robots can transport workpieces between different machining stations, optimising the workflow and minimising idle time [174, 175]. Automated systems can also communicate with other manufacturing systems, such as enterprise resource planning (ERP) software, for seamless production management. Some key aspects of robotics and automation in CNC machining are given below:

- Robotic material handling: Robots can be employed for the automated loading and unloading of raw materials and finished parts in CNC machines. They can handle heavy loads, operate in a precise and repeatable manner, and eliminate the need for manual intervention [176].
- Machine tending: Robots can be utilised for the continuous operation and supervision of CNC machines. They can perform tasks like tool changes, part measurement, and coolant application. This minimises downtime and maximises machine utilisation [177].
- Palletized automation: Palletized automation systems involve the use of robots to move workpieces between different CNC machines, inspection stations, and other manufacturing processes. This allows for seamless integration and optimisation of the entire production line [178].
- Vision systems: Vision-guided robotics enable robots to locate workpieces, align them accurately, and perform tasks with high precision. They use cameras and advanced algorithms to analyse the environment, ensuring precise positioning and improved quality control [179].
- In-process inspection: Automated inspection systems can be integrated into CNC machining operations, allowing robots to measure dimensions, check tolerances, and detect defects in real time. This reduces the need for manual inspection and improves quality control [180].
- Collaborative robots (cobots): Cobots are designed to work alongside human operators, enhancing productivity and safety in CNC machining operations. They can assist with tasks such as part loading, deburring, and quality inspection, while ensuring human–robot collaboration [181].
- Data analytics and integration: Robotics and automation systems generate vast amounts of data that can be leveraged for process optimisation. By integrating CNC machines and automation systems with data analytics tools, manufacturers can gain insights into machine performance, predictive maintenance, and production efficiency [180].
- Programming and simulation: Offline programming tools can be used to simulate and optimise the robot's movements before actual implementation. This can involve using a common programming language or developing a communication interface between the robot controller and the CNC machine [182].
- Flexibility and scalability: Robotic systems offer the advantage of flexibility and scalability. They can be easily reprogrammed to adapt to different parts, product variations, and production volumes. This allows manufacturers to respond quickly to changing market demands [183].

The structure of a modern smart manufacturing factory using robotics and automation is shown in Fig. 11 [180].
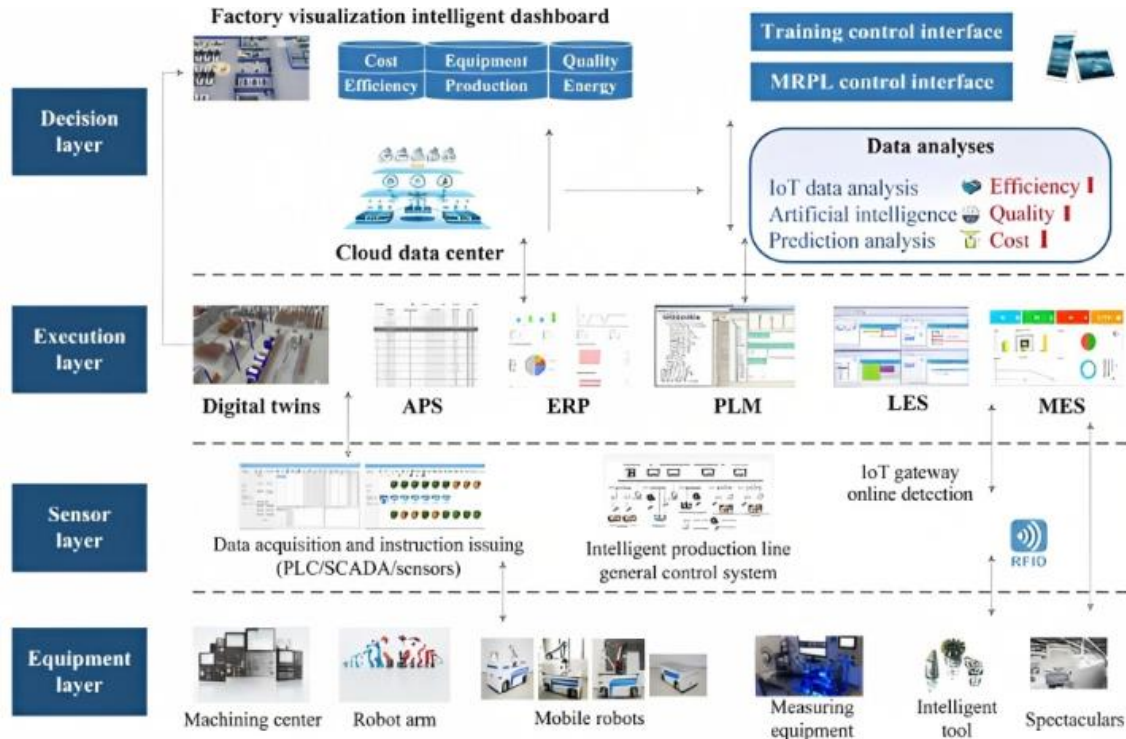
Fig. 11. Structure of a modern smart manufacturing factory using robotics and automation [180]

Overall, the integration of robotics and automation in CNC machining operations and production lines leads to increased efficiency, improved quality, reduced cycle times, and enhanced competitiveness in the manufacturing industry [184]. However, it is crucial to carefully plan and integrate these systems to ensure a smooth and efficient operation in terms of productivity enhancement of CNC machining operations.

11. CONCLUSION

Robotics and automation play a significant role in CNC machining operations, revolutionising the manufacturing industry. The use of robotics and automation in CNC machining operations brings greater efficiency, accuracy, and productivity while reducing costs, human error, and downtime. It enables manufacturers to meet increasing demands, improve quality control, and stay competitive in the industry. Robots are used to load and unload workpieces from CNC machines. They can handle heavy and bulky materials with precision, reducing the need for manual labour and increasing efficiency. Robotic machine tending allows for continuous operation, as robots can work 24/7 without fatigue. CNC machines often require different tools for various machining operations. Automation systems, such as tool changers, enable the machine to automatically switch between tools based on the programmed instructions. This eliminates the need for manual tool changes, saving time and reducing human error. Robots equipped with sensors and vision systems can perform accurate measurements and inspections of machined parts. They can compare the dimensions of the workpiece against the CAD model or predefined specifications, ensuring the quality and consistency of the output. Robots can be programmed to automatically change tools based on the machining requirements. This eliminates the need for manual intervention and reduces downtime during tool changes.

Automation systems integrated with CNC machines can adapt the machining process in real time based on feedback from sensors and monitoring systems. This allows for adjustments to compensate for tool wear, material variations, or environmental factors, resulting in improved precision and reduced scrap. Cobots are designed to work alongside human operators in CNC machining operations. They can assist in tasks such as part loading, deburring, or finishing, enhancing productivity and safety. Cobots have built-in safety features, such as force sensing and collision detection, allowing them to operate in close proximity to humans. Automation in CNC machining operations enables the collection of vast amounts of data, including process parameters, tool wear, and machine performance. These data can be analysed using advanced analytics techniques to identify trends, optimise processes, and predict maintenance requirements, leading to improved efficiency and reduced downtime.

The integration of several CNC machines into an integrated production line is made possible by automation systems. By moving objects between several machining stations, robots may streamline the operation and save downtime. Additionally, automated systems can interface with other manufacturing systems, such as ERP software, to provide smooth production management. Automation in CNC machining allows for the collection of vast amounts of data related to machining parameters, tool performance, and production metrics. These data can be analysed using artificial intelligence and machine-learning algorithms to identify patterns, optimise processes, and improve overall efficiency.

Future research in robotics and automation in CNC machining operations is likely to focus on several key areas. Here are some potential avenues of investigation:

– Advanced machine learning and AI algorithms: Researchers can explore the application of advanced machine learning and artificial intelligence algorithms to improve CNC machining operations. This includes developing algorithms for predictive

maintenance, tool wear prediction, optimising machining parameters, and adaptive control systems.

- Collaborative robotics (cobots): Cobots are designed to work alongside human operators, enhancing their capabilities and improving safety. Future research can investigate the development of more sophisticated cobots that can perform intricate tasks in CNC machining operations while maintaining safe and efficient collaboration with human workers.
- Intelligent process monitoring and control: Research can focus on developing intelligent monitoring and control systems that can analyse real-time data from CNC machines, sensors, and other sources to make informed decisions. This can involve the integration of machine-learning algorithms to detect anomalies, optimise cutting parameters, and ensure consistent product quality.
- Multi-axis machining and complex geometry: As the demand for intricate and complex parts increases, future research can explore advancements in multi-axis machining. This includes developing algorithms and strategies for efficient machining of complex geometries, such as freeform surfaces and non-uniform rational basis splines (NURBS).
- Automated tool path planning: Optimising tool paths is crucial for efficient machining. Future research can focus on automated tool path planning algorithms that minimise cycle time, reduce tool wear, and optimise chip evacuation. This may involve considering factors like material properties, machine dynamics, and geometric constraints.
- Human–machine interfaces (HMIs): User interfaces play a critical role in CNC machining operations. Future research can investigate intuitive and user-friendly HMIs that enable operators to interact with machines more effectively. This includes exploring the use of augmented reality (AR) and virtual reality (VR) technologies to enhance training, programming, and monitoring processes.
- Integration of IoT and Industry 4.0 technologies: The integration of the IoT and Industry 4.0 technologies can enable real-time data collection, analysis, and remote monitoring of CNC machines. Future research can focus on developing scalable and secure architectures to enable seamless connectivity and data exchange between machines, sensors, and other systems.
- Sustainability and energy efficiency: As sustainability becomes increasingly important, future research can explore ways to make CNC machining operations more environmentally friendly. This includes investigating energy-efficient machining strategies, optimising material usage, and minimising waste generation through advanced process monitoring and control.
- Energy efficiency and sustainability: With growing concerns about energy consumption and environmental impact, future research can focus on developing energy-efficient machining strategies and optimising the use of resources in CNC operations. This can involve investigating novel machining techniques, tool materials, or cooling strategies to minimise energy consumption and reduce waste.
- Cybersecurity and safety: As CNC machines become more connected and integrated into manufacturing networks, ensuring cybersecurity and safety becomes crucial. Future research can concentrate on developing robust cybersecurity measures to protect CNC machines from potential threats and implementing safety protocols to prevent accidents or unauthorised access.

- Machine learning for defect detection: Applying machine-learning techniques to detect defects or anomalies in machined parts can help improve quality control. Future research can focus on developing algorithms that can automatically analyse sensor data, such as vibration, temperature, or acoustic signals, to detect and classify defects in real time.
- Adaptive control systems: Research can be directed towards the development of adaptive control systems that can dynamically adjust machining parameters based on real-time feedback. These systems can optimise cutting parameters, tool wear compensation, and feed-rate control, leading to improved machining accuracy and reduced production time.
- Sensor integration and data analytics: Exploring advanced sensing technologies and integrating them into CNC machining processes can provide valuable data for optimisation. Future research can investigate the integration of various sensors, such as force/torque sensors, vision systems, and 3D scanners, and utilise data analytics techniques to extract valuable insights for process improvement.
- Intelligent process planning: Research can be conducted to develop intelligent algorithms that can optimise the process-planning phase in CNC machining. These algorithms can consider various factors such as tool selection, toolpath optimisation, and fixture design, aiming to improve the efficiency, accuracy, and cost-effectiveness.
- HMIs: Enhancing the interaction between humans and CNC machines can lead to improved productivity and ease of use. Future research can explore the development of intuitive user interfaces, AR or VR systems, and haptic feedback to facilitate efficient programming, monitoring, and control of CNC machines.

These research areas have the potential to drive significant advancements in robotics and automation in CNC machining operations, improving productivity, flexibility, and overall manufacturing efficiency.

REFERENCES

1. Bloss R. Machine tools become much more than just a lathe or milling machine. Assembly Automation. 2007;27(1):9-11.
2. Ribeiro J, Lima R, Eckhardt T, Paiva S. Robotic process automation and artificial intelligence in industry 4.0–a literature review. Procedia Computer Science. 2021;181:51-8.
3. Bârsan A. A Brief Review of Robotic Machining. Acta Universitatis Cibiniensis Technical Series. 2019;71(1):9-13.
4. Yuwen S, Jinjie J, Jinting X, Mansen C, Jinbo N. Path, feedrate and trajectory planning for free-form surface machining: A state-of-the-art review. Chinese Journal of Aeronautics. 2022;35(8):12-29.
5. Bartoš M, Bulej V, Bohušík M, Stanček J, Ivanov V, Macek P. An overview of robot applications in automotive industry. Transportation Research Procedia. 2021;55:837-44.
6. Martinova LI, Kozak NV, Kovalev IA, Ljubimov AB. Creation of CNC system's components for monitoring machine tool health. The International Journal of Advanced Manufacturing Technology. 2021;117(7-8):2341-8.
7. Nasir V, Sassani F. A review on deep learning in machining and tool monitoring: methods, opportunities, and challenges. The International Journal of Advanced Manufacturing Technology. 2021;115(9-10):2683-709.
8. Liu C, Zheng P, Xu X. Digitalisation and servitisation of machine tools in the era of Industry 4.0: a review. International journal of production research. 2021:1-33.
9. Nguyen V, Johnson J, Melkote S. Active vibration suppression in robotic milling using optimal control. International Journal of Machine Tools and Manufacture. 2020;152:103541.

Mohsen Soori, Fooad Karimi Ghaleh Jough, Roza Dastres, Behrooz Arezoo
Robotical Automation in CNC Machine Tools: A Review

DOI 10.2478/ama-2024-0048

10. Gienke O, Pan Z, Yuan L, Lepper T, Van Duin S. Mode coupling chatter prediction and avoidance in robotic machining process. The International Journal of Advanced Manufacturing Technology. 2019;104:2103-16.

11. Hao D, Wang W, Liu Z, Yun C. Experimental study of stability prediction for high-speed robotic milling of aluminum. Journal of Vibration and Control. 2020;26(7-8):387-98.

12. Javaid M, Haleem A, Singh RP, Suman R. Substantial capabilities of robotics in enhancing industry 4.0 implementation. Cognitive Robotics. 2021;1:58-75.

13. Onstein IF, Semeniuta O, Bjerkeng M, editors. Deburring using robot manipulators: A review. 2020 3rd International Symposium on Small-scale Intelligent Manufacturing Systems; 2020: IEEE.

14. Zerun Z, Xiaowei T, Chen C, Fangyu P, Rong Y, Lin Z, et al. High precision and efficiency robotic milling of complex parts: Challenges, approaches and trends. Chinese Journal of Aeronautics. 2022;35(2):22-46.

15. Zheng H, Lin J, editors. A deep learning approach for high speed machining tool wear monitoring. 2019 3rd international conference on robotics and automation sciences (ICRAS); 2019: IEEE.

16. Luo W, Hu T, Ye Y, Zhang C, Wei Y. A hybrid predictive maintenance approach for CNC machine tool driven by Digital Twin. Robotics and Computer-Integrated Manufacturing. 2020;65:101974.

17. Zhu D, Feng X, Xu X, Yang Z, Li W, Yan S, et al. Robotic grinding of complex components: a step towards efficient and intelligent machining–challenges, solutions, and applications. Robotics and Computer-Integrated Manufacturing. 2020;65:101908.

18. Iglesias Sánchez I, Ares JE, González Gaya C, Rosales Prieto V. A new approach to the consideration and analysis of critical factors in robotic machining. Applied Sciences. 2020;10(24):8885.

19. Johansen K, Ashourpour M, Rao S. Positioning sustainable automation in production of customized products. Procedia Manufacturing. 2021;55:358-64.

20. Evjemo LD, Gjerstad T, Grøtli EI, Sziebig G. Trends in smart manufacturing: Role of humans and industrial robots in smart factories. Current Robotics Reports. 2020;1:35-41.

21. Chen Q, Zhang C, Hu T, Zhou Y, Ni H, Xue X. Posture optimization in robotic machining based on comprehensive deformation index considering spindle weight and cutting force. Robotics and Computer-Integrated Manufacturing. 2022;74:102290.

22. Pantazis D, Pease SG, Goodall P, West A, Conway P. A design of experiments Cyber–Physical System for energy modelling and optimisation in end-milling machining. Robotics and Computer-Integrated Manufacturing. 2023;80:102469.

23. Soori M, Arezoo B, Habibi M. Accuracy analysis of tool deflection error modelling in prediction of milled surfaces by a virtual machining system. International Journal of Computer Applications in Technology. 2017;55(4):308-21.

24. Soori M, Arezoo B, Habibi M. Virtual machining considering dimensional, geometrical and tool deflection errors in three-axis CNC milling machines. Journal of Manufacturing Systems. 2014;33(4):498-507.

25. Soori M, Arezoo B, Habibi M. Dimensional and geometrical errors of three-axis CNC milling machines in a virtual machining system. Computer-Aided Design. 2013;45(11):1306-13.

26. Soori M, Arezoo B, Habibi M. Tool deflection error of three-axis computer numerical control milling machines, monitoring and minimizing by a virtual machining system. Journal of Manufacturing Science and Engineering. 2016;138(8):081005.

27. Soori M, Asmael M, Solyalı D. Recent Development in Friction Stir Welding Process: A Review. SAE International Journal of Materials and Manufacturing. 2020(5):18.

28. Soori M, Asmael M. Virtual Minimization of Residual Stress and Deflection Error in Five-Axis Milling of Turbine Blades. Strojniski Vestnik/Journal of Mechanical Engineering. 2021;67(5):235-44.

29. Soori M, Asmael M. Cutting temperatures in milling operations of difficult-to-cut materials. Journal of New Technology and Materials. 2021;11(1):47-56.

30. Soori M, Asmael M, Khan A, Farouk N. Minimization of surface roughness in 5-axis milling of turbine blades. Mechanics Based Design of Structures and Machines. 2021;51(9):1-18.

31. Soori M, Asmael M. MINIMIZATION OF DEFLECTION ERROR IN FIVE AXIS MILLING OF IMPELLER BLADES. Facta Universitatis, series: Mechanical Engineering. 2021;21(2):175-90.

32. Soori M. Virtual product development: GRIN Verlag; 2019.

33. Soori M, Asmael M. A Review of the Recent Development in Machining Parameter Optimization. Jordan Journal of Mechanical & Industrial Engineering. 2022;16(2):205-23.

34. Dastres R, Soori M, Asmael M. RADIO FREQUENCY IDENTIFICATION (RFID) BASED WIRELESS MANUFACTURING SYSTEMS, A REVIEW. Independent Journal of Management & Production. 2022;13(1):258-90.

35. Soori M, Arezoo B, Dastres R. Machine Learning and Artificial Intelligence in CNC Machine Tools, A Review. Sustainable Manufacturing and Service Economics. 2023:100009.

36. Soori M, Arezoo B. A Review in Machining-Induced Residual Stress. Journal of New Technology and Materials. 2022;12(1):64-83.

37. Soori M, Arezoo B. Minimization of Surface Roughness and Residual Stress in Grinding Operations of Inconel 718. Journal of Materials Engineering and Performance. 2022:1-10.

38. Soori M, Arezoo B. Cutting Tool Wear Prediction in Machining Operations, A Review. Journal of New Technology and Materials. 2022;12(2):15-26.

39. Soori M, Asmael M. Classification of research and applications of the computer aided process planning in manufacturing systems. Independent Journal of Management & Production. 2021;12(5):1250-81.

40. Dastres R, Soori M. Advances in web-based decision support systems. International Journal of Engineering and Future Technology. 2021;19(1):1-15.

41. Dastres R, Soori M. Artificial Neural Network Systems. International Journal of Imaging and Robotics (IJIR). 2021;21(2):13-25.

42. Dastres R, Soori M. The Role of Information and Communication Technology (ICT) in Environmental Protection. International Journal of Tomography and Simulation. 2021;35(1):24-37.

43. Dastres R, Soori M. Secure Socket Layer in the Network and Web Security. International Journal of Computer and Information Engineering. 2020;14(10):330-3.

44. Dastres R, Soori M. Advances in Web-Based Decision Support Systems. International Journal of Engineering and Future Technology. 2021.

45. Dastres R, Soori M. A review in recent development of network threats and security measures. International Journal of Information Sciences and Computer Engineering. 2021.

46. Dastres R, Soori M. Advanced image processing systems. International Journal of Imagining and Robotics. 2021;21(1):27-44.

47. Soori M, Arezoo B. Dimensional, geometrical, thermal and tool deflection errors compensation in 5-Axis CNC milling operations. Australian Journal of Mechanical Engineering. 2023:1-15.

48. Soori M, Arezoo B, Dastres R. Artificial Intelligence, Machine Learning and Deep Learning in Advanced Robotics, A Review. Cognitive Robotics. 2023;3:54-70.

49. Soori M, Arezoo B. Effect of cutting parameters on tool life and cutting temperature in milling of AISI 1038 carbon steel. Journal of New Technology and Materials. 2023.

50. Soori M, Arezoo B. The effects of coolant on the cutting temperature, surface roughness and tool wear in turning operations of Ti6Al4V alloy. Mechanics Based Design of Structures and Machines. 2023:1-23.

51. Soori M. Advanced Composite Materials and Structures. Journal of Materials and Engineering Structures. 2023.

52. Soori M, Arezoo B, Dastres R. Internet of things for smart factories in industry 4.0, a review. Internet of Things and Cyber-Physical Systems. 2023.

53. Soori M, Arezoo B. Cutting tool wear minimization in drilling operations of titanium alloy Ti-6Al-4V. Proceedings of the Institution of Mechanical Engineers, Part J: Journal of Engineering Tribology. 2023:13506501231158259.

54. Soori M, Arezoo B. Minimization of surface roughness and residual stress in abrasive water jet cutting of titanium alloy Ti6Al4V. Proceedings of the Institution of Mechanical Engineers, Part E: Journal of Process Mechanical Engineering. 2023:09544089231157972.

55. Soori M. Deformation error compensation in 5-Axis milling operations of turbine blades. Journal of the Brazilian Society of Mechanical Sciences and Engineering. 2023;45(6):289.

56. Soori M, Arezoo B. Modification of CNC Machine Tool Operations and Structures Using Finite Element Methods, A Review. Jordan Journal of Mechanical and Industrial Engineering. 2023.

57. Soori M, Arezoo B, Dastres R. Optimization of Energy Consumption in Industrial Robots, A Review. Cognitive Robotics. 2023.

58. Soori M, Arezoo B, Dastres R. Advanced Virtual Manufacturing Systems, A Review. Journal of Advanced Manufacturing Science and Technology. 2023.

59. Jough FKG, Şensoy S. Prediction of seismic collapse risk of steel moment frame mid-rise structures by meta-heuristic algorithms. Earthquake Engineering and Engineering Vibration. 2016;15: 743-57.

60. Karimi Ghaleh Jough F, Şensoy S. Steel moment-resisting frame reliability via the interval analysis by FCM-PSO approach considering various uncertainties. Journal of Earthquake Engineering. 2020;24(1):109-28.

61. Karimi Ghaleh Jough F, Golhashem M. Assessment of out-of-plane behavior of non-structural masonry walls using FE simulations. Bulletin of Earthquake Engineering. 2020;18(14):6405-27.

62. Taher GA, Yousuf Howlader MAR, Touqir FA. Automation of material handling with bucket elevator and belt conveyor. International Journal of Scientific and Research Publications. 2014;4(3): 1-13.

63. Deal WF, Jones CE. Integrating CAD/CAM in automation and materials handling. Technology and Engineering Teacher. 2012;71(6):13.

64. Maheswari C, Priyanka E, Thangavel S, Parameswari P. Development of unmanned guided vehicle for material handling automation for industry 4.0. Int J Recent Technol Eng. 2018;7(4):428-32.

65. Mahdavi I, Shirazi B, Sahebjamnia N. Development of a simulation-based optimisation for controlling operation allocation and material handling equipment selection in FMS. International Journal of Production Research. 2011;49(23):6981-7005.

66. Lee S, Kim Y, Kahng H, Lee S-K, Chung S, Cheong T, et al. Intelligent traffic control for autonomous vehicle systems based on machine learning. Expert Systems with Applications. 2020;144:113074.

67. Kunduru AR. Cloud BPM Application (Appian) Robotic Process Automation Capabilities. Asian Journal of Research in Computer Science. 2023;16(3):267-80.

68. Muñoz-Benavent P, Solanes JE, Gracia L, Tornero J. Robust auto tool change for industrial robots using visual servoing. International Journal of Systems Science. 2019;50(2):432-49.

69. Soori M, Arezoo B, Dastres R. Artificial Neural Networks in Supply Chain Management, A Review. Journal of Economy and Technology. 2023;1:179-96.

70. Javaid M, Haleem A, Singh RP, Rab S, Suman R. Exploring impact and features of machine vision for progressive industry 4.0 culture. Sensors International. 2022;3:100132.

71. Jordaan GD, Van Nieuwenhuizen RJ, editors. Machine Vision in an Automated Component-Handling System. 2006 IEEE International Conference on Mechatronics; 2006: IEEE.

72. Mudiyanselage SE, Nguyen PHD, Rajabi MS, Akhavian R. Automated workers' ergonomic risk assessment in manual material handling using sEMG wearable sensors and machine learning. Electronics. 2021;10(20):2558.

73. Guzzi J, Abbate G, Paolillo A, Giusti A, editors. Interacting with a conveyor belt in virtual reality using pointing gestures. 2022 17th

74. Rahman HF, Nielsen I. Scheduling automated transport vehicles for material distribution systems. Applied Soft Computing. 2019;82:105552.

75. Tyagi AK, Fernandez TF, Mishra S, Kumari S, editors. Intelligent automation systems at the core of industry 4.0. International conference on intelligent systems design and applications; 2020: Springer.

76. Brecher C, Schröter B, Almeida C, editors. Development and programming of portable robot systems for material handling tasks. Proceedings of the CIRP International Conference on Reconfigurable Manufacturing; 2005: Citeseer.

77. Wan J, Tang S, Hua Q, Li D, Liu C, Lloret J. Context-aware cloud robotics for material handling in cognitive industrial Internet of Things. IEEE Internet of Things Journal. 2017;5(4):2272-81.

78. Xie C, Allen TT. Simulation and experimental design methods for job shop scheduling with material handling: a survey. The International Journal of Advanced Manufacturing Technology. 2015;80(1-4):233-43.

79. Obreja C, Stan G, Andrioaia D, Funaru M, editors. Design of an automatic tool changer system for milling machining centers. Applied Mechanics and Materials; 2013: Trans Tech Publ.

80. Verma K, Belokar R, Verma VK, Ntalianis K. Track-based analysis for profile generation on globoidal cam in automatic tool changer of CNC machining center. Assembly Automation. 2019;39(2):369-79.

81. Wu X, Liu Y, Zhou X, Mou A. Automatic identification of tool wear based on convolutional neural network in face milling process. Sensors. 2019;19(18):3817.

82. Lianzhong Z, Li W, editors. Machining center automatic ATC analysis and research. 2010 3rd International Conference on Information Management, Innovation Management and Industrial Engineering; 2010: IEEE.

83. Leng J, Liu Q, Ye S, Jing J, Wang Y, Zhang C, et al. Digital twin-driven rapid reconfiguration of the automated manufacturing system via an open architecture model. Robotics and Computer-Integrated Manufacturing. 2020;63:101895.

84. Dereli T, Filiz İH. Allocating optimal index positions on tool magazines using genetic algorithms. Robotics and Autonomous Systems. 2000;33(2-3):155-67.

85. Gong QS, Luo M, Ren AH, Chang ZB. Research on dynamic characteristics of globoidal cam employed in automatic tool changer (ATC) of machine center. Applied Mechanics and Materials. 2014;456:133-6.

86. Chen S-L, Su C-F, Cheng Y-T. A novel framework for diagnosing automatic tool changer and tool life based on cloud computing. Advances in Mechanical Engineering. 2016;8(3):1687814016637319.

87. Oliveira TLL, Zitoune R, Ancelotti Jr AC, da Cunha Jr SS. Smart machining: Monitoring of CFRP milling using AE and IR. Composite Structures. 2020;249:112611.

88. Van HP, Thuy DN. Influence of relative humidity and air temperature on the stopping position of the automatic tool changer in a CNC machine when using a pneumatic cylinder. International Journal of Modern Physics B. 2021;35(14n16):2140013.

89. Lu XH, Han PZ, Wu WY, Jie W, editors. Reliability evaluation of circular tool magazine and automatic tool changer. Advanced Materials Research; 2013: Trans Tech Publ.

90. Nakamoto K, Shirase K, Wakamatsu H, Tsumaya A, Arai E. Automatic production planning system to achieve flexible direct machining. JSME International Journal Series C Mechanical Systems, Machine Elements and Manufacturing. 2004;47(1):136-43.

91. Tian H, Yang Z, Li G, Chen C. Study on failure warning of tool magazine and automatic tool changer based on tool-pulling force. Journal of Mechanical Science and Technology. 2019;33:4371-81.

92. Zhou L, Li F, Wang Y, Wang L, Wang G. A new empirical standby power and auxiliary power model of CNC machine tools. The International Journal of Advanced Manufacturing Technology. 2022;120(5-6):3995-4010.

Mohsen Soori, Fooad Karimi Ghaleh Jough, Roza Dastres, Behrooz Arezoo
Robotical Automation in CNC Machine Tools: A Review

DOI 10.2478/ama-2024-0048

93. Chen C, Tian H, Zhang J, Shi X, Chen L, Bao J, et al. Study on failure warning of tool magazine and automatic tool changer. Journal of Vibroengineering. 2016;18(2):883-99.

94. Barbosa M, Silva F, Pimentel C, Gouveia RM. A novel concept of CNC machining center automatic feeder. Procedia Manufacturing. 2018;17:952-9.

95. Živanović S, Slavković N, editors. Programming of machine tools and robots for machining using STEP-NC in the era of Industry 4.0. Proceedings of the 15th International Conference on Accomplishments in Mechanical and Industrial Engineering DEMI 2021; 2021: University of Banjaluka, Faculty of Mechanical Engineering.

96. Leali F, Pellicciari M, Pini F, Vergnano A, Berselli G, editors. A calibration method for the integrated design of finishing robotic workcells in the aerospace industry. Robotics in Smart Manufacturing: International Workshop, WRSM 2013, Co-located with FAIM 2013, Porto, Portugal, June 26-28, 2013 Proceedings; 2013: Springer.

97. Chen Y, Dong F. Robot machining: recent development and future research issues. The International Journal of Advanced Manufacturing Technology. 2013;66:1489-97.

98. Gultekin H, Akturk MS, Karasan OE. Bicriteria robotic operation allocation in a flexible manufacturing cell. Computers & operations research. 2010;37(4):779-89.

99. Vaher K, Kangru T, Otto T, Riives J. THE MOBILITY OF ROBOT-ISED WORK CELLS IN MANUFACTURING. Annals of DAAAM & Proceedings. 2019;30.

100. Schneider U, Drust M, Ansaloni M, Lehmann C, Pellicciari M, Leali F, et al. Improving robotic machining accuracy through experimental error investigation and modular compensation. The International Journal of Advanced Manufacturing Technology. 2016;85:3-15.

101. Buerkle A, Eaton W, Lohse N, Bamber T, Ferreira P. EEG based arm movement intention recognition towards enhanced safety in symbiotic Human-Robot Collaboration. Robotics and Computer-Integrated Manufacturing. 2021;70:102137.

102. Soori M, Arezoo B, Dastres R. Virtual manufacturing in industry 4.0: A review. Data Science and Management. 2023.

103. Bedaka AK, Vidal J, Lin C-Y. Automatic robot path integration using three-dimensional vision and offline programming. The International Journal of Advanced Manufacturing Technology. 2019;102:1935-50.

104. Devine K, Reifschneider L, editors. Agile robotic work cells for teaching manufacturing engineering. Proceedings of ASEE; 2009.

105. Lee NK. Total automation: The possibility of lights-out manufacturing in the near future. Missouri S&T's Peer to Peer. 2018;2(1):4.

106. Søndergaard A, Feringa J, Stan F, Maier D. Robotic abrasive wire cutting of polymerized styrene formwork systems for cost-effective realization of topology-optimized concrete structures. Construction Robotics. 2018;2(1-4):81-92.

107. Zonta T, Da Costa CA, da Rosa Righi R, de Lima MJ, da Trindade ES, Li GP. Predictive maintenance in the Industry 4.0: A systematic literature review. Computers & Industrial Engineering. 2020;150:106889.

108. Liu Y, Candell R, Kashef M, Montgomery K, editors. A collaborative work cell testbed for industrial wireless communications—the baseline design. 2019 IEEE 28th International Symposium on Industrial Electronics (ISIE); 2019: IEEE.

109. KUKA milling robot [Available from: https://www.kuka.com/en-my/products/process-technologies/milling.

110. Soori M, Arezoo B, Dastres R. Internet of things for smart factories in industry 4.0, a review. Internet of Things and Cyber-Physical Systems. 2023;3:192-204.

111. Soori M, Arezoo B, Dastres R. Digital Twin for Smart Manufacturing, A Review. Sustainable Manufacturing and Service Economics. 2023;2:100017.

112. Ashima R, Haleem A, Bahl S, Javaid M, Mahla SK, Singh S. Automation and manufacturing of smart materials in Additive Manufacturing technologies using Internet of Things towards the adoption of Industry 4.0. Materials Today: Proceedings. 2021;45:5081-8.

113. Oyekan J, Farnsworth M, Hutabarat W, Miller D, Tiwari A. Applying a 6 DoF robotic arm and digital twin to automate fan-blade reconditioning for aerospace maintenance, repair, and overhaul. Sensors. 2020;20(16):4637.

114. Lotti N, Xiloyannis M, Durandau G, Galofaro E, Sanguineti V, Masia L, et al. Adaptive model-based myoelectric control for a soft wearable arm exosuit: A new generation of wearable robot control. IEEE Robotics & Automation Magazine. 2020;27(1):43-53.

115. Carlucho I, De Paula M, Acosta GG. An adaptive deep reinforcement learning approach for MIMO PID control of mobile robots. ISA transactions. 2020;102:280-94.

116. Wang L, Orban P, Cunningham A, Lang S. Remote real-time CNC machining for web-based manufacturing. Robotics and Computer-Integrated Manufacturing. 2004;20(6):563-71.

117. Reel PS, Reel S, Pearson E, Trucco E, Jefferson E. Using machine learning approaches for multi-omics data analysis: A review. Biotechnology Advances. 2021;49:107739.

118. Enthrakandi Narasimhan G, Bettyjane J. Implementation and study of a novel approach to control adaptive cooperative robot using fuzzy rules. International Journal of Information Technology. 2021;13:2287-94.

119. Sheridan TB. Adaptive automation, level of automation, allocation authority, supervisory control, and adaptive control: Distinctions and modes of adaptation. IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans. 2011;41(4):662-7.

120. Jezernik S, Colombo G, Morari M. Automatic gait-pattern adaptation algorithms for rehabilitation with a 4-DOF robotic orthosis. IEEE Transactions on Robotics and Automation. 2004;20(3):574-82.

121. Zou H, Chen C-L, Li M, Yang J, Zhou Y, Xie L, et al. Adversarial learning-enabled automatic WiFi indoor radio map construction and adaptation with mobile robot. IEEE Internet of Things Journal. 2020;7(8):6946-54.

122. Weingarten JD, Lopes GA, Buehler M, Groff RE, Koditschek DE, editors. Automated gait adaptation for legged robots. IEEE International Conference on Robotics and Automation, 2004 Proceedings ICRA'04 2004; 2004: IEEE.

123. Neythalath N, Søndergaard A, Bærentzen JA. Adaptive robotic manufacturing using higher order knowledge systems. Automation in Construction. 2021;127:103702.

124. Sahil D, Sarabjeet SS, editors. REDUCTION OF SCRAP AND INSPECTION EFFORT: AN APPROACH INCORPORATING INDUSTRY 4.0. Electrophysical machining in modern industry; 2021.

125. Xiao R, Xu Y, Hou Z, Chen C, Chen S. An adaptive feature extraction algorithm for multiple typical seam tracking based on vision sensor in robotic arc welding. Sensors and Actuators A: Physical. 2019;297:111533.

126. Lv Y, Peng Z, Qu C, Zhu D. An adaptive trajectory planning algorithm for robotic belt grinding of blade leading and trailing edges based on material removal profile model. Robotics and Computer-Integrated Manufacturing. 2020;66:101987.

127. Perzylo A, Rickert M, Kahl B, Somani N, Lehmann C, Kuss A, et al. SMErobotics: Smart robots for flexible manufacturing. IEEE Robotics & Automation Magazine. 2019;26(1):78-90.

128. Jia F, Tzintzun J, Ahmad R, editors. An improved robot path planning algorithm for a novel self-adapting intelligent machine tending robotic system. Industrial and Robotic Systems: LASIRS 2019; 2020: Springer.

129. Annem V, Rajendran P, Thakar S, Gupta SK, editors. Towards remote teleoperation of a semi-autonomous mobile manipulator system in machine tending tasks. International Manufacturing Science and Engineering Conference; 2019: American Society of Mechanical Engineers.

130. Park Y-K, Park J, Kim S, Lee D, Lee S, Um J. Development of Digital twin for Plug-and-Produce of a Machine tending system through ISO 21919 interface. International Journal of Computer Integrated Manufacturing. 2023:1-16.

131. Steele JW, Wysk RA, Ferreira JC. A resource-oriented tolerance representation scheme for the planning of robotic machine tending operations in automated manufacturing systems. The International Journal of Advanced Manufacturing Technology. 2008;38:741-56.

132. Bahrin MAK, Othman MF, Azli NHN, Talib MF. Industry 4.0: A review on industrial automation and robotic. Jurnal teknologi. 2016;78(6-13).

133. Jia F, Jebelli A, Ma Y, Ahmad R. An Intelligent Manufacturing Approach Based on a Novel Deep Learning Method for Automatic Machine and Working Status Recognition. Applied Sciences. 2022;12(11):5697.

134. Schneider C, Klos M, Bdiwi M, Putz M. Machine-To-Machine (M2M) Communication of Robotic Platform in Machine Tending Applications.

135. Landscheidt S, Kans M, Winroth M. Opportunities for robotic automation in wood product industries: the supplier and system integrators' perspective. Procedia Manufacturing. 2017;11:233-40.

136. Rooks B. Robots make a show at the UK automation and machine tool exhibitions. Industrial Robot: An International Journal. 2002.

137. Chen Q, Heydari B, Moghaddam M. Leveraging task modularity in reinforcement learning for adaptable industry 4.0 automation. Journal of Mechanical Design. 2021;143(7).

138. Giles DM, Sinyuk A, Sorokin MG, Schafer JS, Smirnov A, Slutsker I, et al. Advancements in the Aerosol Robotic Network (AERONET) Version 3 database–automated near-real-time quality control algorithm with improved cloud screening for Sun photometer aerosol optical depth (AOD) measurements. Atmospheric Measurement Techniques. 2019;12(1):169-209.

139. Plaza EG, López PN, González EB. Efficiency of vibration signal feature extraction for surface finish monitoring in CNC machining. Journal of Manufacturing Processes. 2019;44:145-57.

140. Dev Anand M, Selveraj T, Kumanan S, Ajith Bosco Raj T. Robotics in online inspection and quality control using moment algorithm. Advances in Production Engineering & Management. 2012;7(1): 27-38.

141. Montironi M, Castellini P, Stroppa L, Paone N. Adaptive autonomous positioning of a robot vision system: Application to quality control on production lines. Robotics and Computer-Integrated Manufacturing. 2014;30(5):489-98.

142. Tripathi S, Shukla S, Attrey S, Agrawal A, Bhadoria VS. Smart industrial packaging and sorting system. Strategic system assurance and business analytics. 2020:245-54.

143. Navon R. Process and quality control with a video camera, for a floor-tilling robot. Automation in construction. 2000;10(1):113-25.

144. Gupta M, Khan MA, Butola R, Singari RM. Advances in applications of Non-Destructive Testing (NDT): A review. Advances in Materials and Processing Technologies. 2022;8(2):2286-307.

145. Azamfirei V, Granlund A, Lagrosen Y. Multi-layer quality inspection system framework for industry 4.0. International journal of automation technology. 2021;15(5):641-50.

146. Brito T, Queiroz J, Piardi L, Fernandes LA, Lima J, Leitão P. A machine learning approach for collaborative robot smart manufacturing inspection for quality control systems. Procedia Manufacturing. 2020;51:11-8.

147. Sun Y, Lu L, Wu F, Xiao S, Sha J, Zhang L, editors. Error Analysis of a Coordinate Measuring Machine with a 6-DOF Industrial Robot Holding the Probe. Actuators; 2023: MDPI.

148. Saez M, Maturana FP, Barton K, Tilbury DM. Real-time manufacturing machine and system performance monitoring using internet of things. IEEE Transactions on Automation Science and Engineering. 2018;15(4):1735-48.

149. My CA. The role of big data analytics and AI in smart manufacturing: An overview. Research in Intelligent and Computing in Engineering: Select Proceedings of RICE 2020. 2021:911-21.

150. Chen S-L, Jen Y. Data fusion neural network for tool condition monitoring in CNC milling machining. International journal of machine tools and manufacture. 2000;40(3):381-400.

151. Ridwan F, Xu X. Advanced CNC system with in-process feed-rate optimisation. Robotics and Computer-Integrated Manufacturing. 2013;29(3):12-20.

152. Li H, Li R, Zhang J, Zhang P. Development of a pipeline inspection robot for the standard oil pipeline of China national petroleum corporation. Applied Sciences. 2020;10(8):2853.

153. Bhatia P, Liu Y, Nagaraj S, Achanta V, Pulaparthi B, Diaz-Elsayed N, editors. Data-Driven Multi-Criteria Decision-Making for Smart and Sustainable Machining. ASME International Mechanical Engineering Congress and Exposition; 2021: American Society of Mechanical Engineers.

154. Raval MB, Joshi H. Categorical framework for implementation of industry 4.0 techniques in medium-scale bearing manufacturing industries. Materials Today: Proceedings. 2022;65:3531-7.

155. Lu Y, Xu X. Cloud-based manufacturing equipment and big data analytics to enable on-demand manufacturing services. Robotics and Computer-Integrated Manufacturing. 2019;57:92-102.

156. Duro JA, Padget JA, Bowen CR, Kim HA, Nassehi A. Multi-sensor data fusion framework for CNC machining monitoring. Mechanical systems and signal processing. 2016;66:505-20.

157. Pieskä S, Kaarela J, Mäkelä J, editors. Simulation and programming experiences of collaborative robots for small-scale manufacturing. 2018 2nd International Symposium on Small-scale Intelligent Manufacturing Systems (SIMS); 2018: IEEE.

158. Ronzoni M, Accorsi R, Botti L, Manzini R. A support-design framework for Cooperative Robots systems in labor-intensive manufacturing processes. Journal of Manufacturing Systems. 2021;61:646-57.

159. Lima F, De Carvalho CN, Acardi MB, Dos Santos EG, De Miranda GB, Maia RF, et al. Digital manufacturing tools in the simulation of collaborative robots: Towards industry 4.0. Brazilian Journal of Operations & Production Management. 2019;16(2):261-80.

160. Hashemi-Petroodi SE, Thevenin S, Kovalev S, Dolgui A. Operations management issues in design and control of hybrid human-robot collaborative manufacturing systems: a survey. Annual Reviews in Control. 2020;49:264-76.

161. Norman AR, Schönberg A, Gorlach IA, Schmitt R. Validation of iGPS as an external measurement system for cooperative robot positioning. The International Journal of Advanced Manufacturing Technology. 2013;64:427-46.

162. Glatt M, Sinnwell C, Yi L, Donohoe S, Ravani B, Aurich JC. Modeling and implementation of a digital twin of material flows based on physics simulation. Journal of Manufacturing Systems. 2021;58:231-45.

163. Ammar M, Haleem A, Javaid M, Walia R, Bahl S. Improving material quality management and manufacturing organizations system through Industry 4.0 technologies. Materials Today: Proceedings. 2021;45:5089-96.

164. George P, Cheng C-T, Pang TY, Neville K. Task Complexity and the Skills Dilemma in the Programming and Control of Collaborative Robots for Manufacturing. Applied Sciences. 2023;13(7):4635.

165. Perez-Ubeda R, Gutierrez S, Zotovic R, Lluch-Cerezo J. Study of the application of a collaborative robot for machining tasks. Procedia Manufacturing. 2019;41:867-74.

166. Michalik P, Hatala M, Dobransky J, Macej J, Petrus M, Tirpak P, et al., editors. Design and Evaluation of Production of a Robotic Angle Arm for Collaborative Robot Using the WorkNC CAM Application. 5th EAI International Conference on Management of Manufacturing Systems; 2022: Springer.

167. Liu Y, Guo L, Gao H, You Z, Ye Y, Zhang B. Machine vision based condition monitoring and fault diagnosis of machine tools using information from machined surface texture: A review. Mechanical Systems and Signal Processing. 2022;164:108068.

168. Kim H, Lim D-E, Lee S. Deep learning-based dynamic scheduling for semiconductor manufacturing with high uncertainty of automated material handling system capability. IEEE Transactions on Semiconductor Manufacturing. 2020;33(1):13-22.

169. Wang KB, Dailami F, Matthews J. Towards collaborative robotic polishing of mould and die sets. Procedia Manufacturing. 2019;38:1499-507.

170. Miyake Y, Kondo Y. A study on new machining method applied to a collaborative robot for drilling. Robotics and Computer-Integrated Manufacturing. 2022;78:102409.

171. Borboni A, Reddy KVV, Elamvazuthi I, AL-Quraishi MS, Natarajan E, Azhar Ali SS. The Expanding Role of Artificial Intelligence in Collaborative Robots for Industrial Applications: A Systematic Review of Recent Works. Machines. 2023;11(1):111.

172. El Makrini I, Elprama SA, Van den Bergh J, Vanderborght B, Knevels A-J, Jewell CI, et al. Working with walt: How a cobot was developed and inserted on an auto assembly line. IEEE Robotics & Automation Magazine. 2018;25(2):51-8.

173. Touzani H, Hadj-Abdelkader H, Séguy N, Bouchafa S. Multi-robot task sequencing & automatic path planning for cycle time optimization: Application for car production line. IEEE Robotics and Automation Letters. 2021;6(2):1335-42.

174. Goel R, Gupta P. Robotics and industry 4.0. A Roadmap to Industry 40: Smart Production, Sharp Business and Sustainable Development. 2020:157-69.

175. Tan M, Chen J, Radhakrishnan R, editors. Design of Control System of Automated Production Line Based on PLC and Robot. Tenth International Conference on Applications and Techniques in Cyber Intelligence (ICATCI 2022) Volume 2; 2023: Springer.

176. Rahman HF, Janardhanan MN, Nielsen P. An integrated approach for line balancing and AGV scheduling towards smart assembly systems. Assembly Automation. 2020;40(2):219-34.

177. Pedersen MR, Nalpantidis L, Andersen RS, Schou C, Bøgh S, Krüger V, et al. Robot skills for manufacturing: From concept to industrial deployment. Robotics and Computer-Integrated Manufacturing. 2016;37:282-91.

178. Susanti S, Sutopo W, Ngadiman N, editors. Equipment Replacement Analysis from Manual Line to Automatic Line in Palletizing Activities: A Case Study. IOP Conference Series: Materials Science and Engineering; 2021: IOP Publishing.

179. Mohammed A, Schmidt B, Wang L. Active collision avoidance for human–robot collaboration driven by vision sensors. International Journal of Computer Integrated Manufacturing. 2017;30(9):970-80.

180. Wang S, Jiang L, Meng J, Xie Y, Ding H. Training for smart manufacturing using a mobile robot-based production line. Frontiers of Mechanical Engineering. 2021;16:249-70.

181. Malik AA, Brem A. Digital twins for collaborative robots: A case study in human-robot interaction. Robotics and Computer-Integrated Manufacturing. 2021;68:102092.

182. Kukartsev V, Boyko A, Mikhalev A, Tynchenko V, Rukosueva A, Korpacheva L, editors. Simulation-dynamic model of working time costs calculation for performance of operations on CNC machines. Journal of Physics: Conference Series; 2020: IOP Publishing.

183. Scholz S, Mueller T, Plasch M, Limbeck H, Adamietz R, Iseringhausen T, et al. A modular flexible scalable and reconfigurable system for manufacturing of microsystems based on additive manufacturing and e-printing. Robotics and Computer-Integrated Manufacturing. 2016;40:14-23.

184. Kousi N, Gkournelos C, Aivaliotis S, Lotsaris K, Bavelos AC, Baris P, et al. Digital twin for designing and reconfiguring human–robot collaborative assembly lines. Applied Sciences. 2021;11(10):4620.

Mohsen Soori: https://orcid.org/0000-0002-4358-7513

Fooad Karimi Ghaleh Jough: https://orcid.org/0000-0003-0697-516X

Roza Dastres: https://orcid.org/0000-0002-0278-5003

Behrooz Arezoo: https://orcid.org/0000-0002-3054-4696

VISUAL SIMULTANEOUS LOCALISATION AND MAPPING METHODOLOGIES

Zoulikha BOUHAMATOU*⊙**, Foudil ABDESSEMED***⊙

**Faculty of Technology, Department of Electronics, University de Batna 2 - Mostefa Ben Boulaïd
53, Route de Constantine. Fésdis, Batna 05078, Algeria*

z.bouhamatou@univ-batna2.dz, f.abdessemed@univ-batna2.dz

Abstract: Simultaneous localisation and mapping (SLAM) is a process by which robots build maps of their environment and simultaneously determine their location and orientation in the environment. In recent years, SLAM research has advanced quickly. Researchers are currently working on developing reliable and accurate visual SLAM algorithms dealing with dynamic environments. The steps involved in developing a SLAM system are described in this article. We explore the most-recent methods used in SLAM systems, including probabilistic methods, visual methods, and deep learning (DL) methods. We also discuss the fundamental techniques utilised in SLAM fields.

Key words*: simultaneous localisation and mapping, SLAM, visual SLAM, deep-learning SLAM*

1. INTRODUCTION

Simultaneous localisation and mapping (SLAM) has been a focus of active study in contemporary robotics for the past few years. In this challenge, a mobile robot locates itself in an unfamiliar location and continuously generates a map of that environment. SLAM can be applied to both indoor and outdoor settings. The technique can be applied to a wide variety of fields, including underwater or aerial planning, and is not just limited to land-based mobile robots. Navigating, locating and mapping are the core technologies meant for use by intelligent, autonomous mobile robots. The goal is to first create a map of an unknown environment; then, information pertaining to the robot's motion and that of the unknown environment are determined so as to track the position of the robot in the environment.

The goal of the probabilistic SLAM problem is to find the position of the robot x_t at time k, which moves in an unknown environment, from the set of all observed landmarks m, the set of observations $z_{0:t}$, the set of commands given to the robot $u_{0:t}$ and the initial state x_0. The robot moves in an erratic manner, making it harder and harder to pinpoint where it is right now in terms of global coordinates. The robot's noise sensor allows it to detect its surroundings while it is moving. After creating the map, the goal is to be able to gauge the vehicle's position.

According to pose graph optimisation in robotics, the state of the vehicle can be indicated. There are two techniques relevant to SLAM. If the present and previous postures of the robot are taken into consideration, the full SLAM technique can determine the entire trajectory of the robot. Based on the total sensor data, it estimates the entire set of poses. The online SLAM technique is carried out if we take into account the current pose and disregard environmental features (mapping) by observing the environmental features with a sensor and applying the command vector to the robot, often based on the most recent sensor data. The rule of Bayes can present the incremental nature of the problem.

Proprioceptive, exteroceptive or a combination of both sensors is used to determine the location and mapping of a robot. In SLAM systems, well-known sensors, including GPS, SONAR, LIDAR, IR, inertial measurement units (IMU), and cameras, have been used. When a camera serves as the single external sensor, the SLAM system is known as visual SLAM or V-SLAM. Visual SLAM is primarily divided into the monocular, RGB-D and stereo SLAM techniques based on the type of camera used. (1) Monocular SLAM is focused on using just one camera. (2) The RGB-D SLAM sensor, or RGB-D camera, is made up of the monocular camera and the infrared sensor combinations. When used in RGB-D cameras, they can produce colourful images with depth and real-time 3D data. It is based on structured light. These cameras capture real-time 3D data. (3) SLAM stereo vision refers to the employment of multiple cameras, two or more lenses and a separate image sensor. The visual sensors have visual odometry on their own. It is precise, robust, and easy to implement.

The camera is the most-popular sensor for acquiring visual information. However, it has several drawbacks, such as its poor optical resolution and sensitivity, which are particularly apparent in dim and complicated lighting conditions. Several imaging technologies, including LIDAR, have been created to overcome these drawbacks. However, because cameras are so closely modelled after the human visual acquisition system (eye), they provide significant benefits in terms of their ability to record colour and texture information as well as their ease of interpretation and understanding by a human observer.

The type of map needed and the environment will influence the sensor selection. To accurately estimate the robot's pose and model the scene spatially, one can put a variety of sensors on the robot's body and combine the collected data.

The task of visual localisation depends on three principal concepts: VO (1)(2), structure from motion (SFM) (3), [4] and SLAM, where VO depends on locating the ego-motion or 3D motion of a robot by relying on the input from the camera's 'image.' It is primarily focused on reconstructing the camera's path. The SFM is based on the recovery of the relative poses of a camera and the

three-dimensional (3D) structure from a set of (2D) images of a camera or video. SLAM consists of using these two pieces of information at once, estimating the trajectory of the camera while simultaneously reconstructing the environment.

Visual simultaneous localisation and mapping (VSLAM) is used to enhance surgical performance in the medical field since a large number of individuals face surgical difficulties each year. Fifty percent of these issues can be avoided with proper surgical training and evaluation. Current research combines many deep learning (DL) approaches. Automating surgical reviews, keeping an eye on surgical procedures and assisting surgeons in making decisions during operations all depend on the recognition of surgical tools and workflows to improve surgical performance. Various neural networks (NNs) have been developed in this sector to conduct tool and workflow recognition as well as to extract visual information from surgical videos (5)(6).

This research has also found applications in the agriculture domain. It is necessary to create innovative approaches that can boost production while reducing the demand for human labour. Automated and intelligent agricultural systems are crucial to addressing issues including the lack of manpower, improving worker safety and cutting production costs by preserving energy, money and time. Precision agriculture may be characterised as a strategy that enables the producer to make better decisions per unit area of land and per unit of time. Nowadays, a greater number of fruits and vegetables are cultivated in greenhouses, and it is just as crucial to monitor indoor cultivars as it is for crops produced outdoors (7). Image sensors are becoming more and more common, and they are being utilised in greenhouses to gather data for purposes like plant-monitoring techniques. A technique for identifying and categorising bacterial spot infections in tomato crops using camera pictures was developed by Borges et al. (8). In the study by Liu et al. (9), the authors take pictures of cucumbers within a greenhouse using a handheld camera, and then they apply DL to recognise the objects. This is yet another example using camera image analysis. Methods for calculating the animal condition score have also been used after digital picture processing (10)(11). Real-time site monitoring is a current difficulty in indoor precision farming and animal management. The most-common and time-consuming tasks in on-farm operations were found to be gathering data for tracking crop growth or animal conditions (12). Thus, novel remote-sensing techniques based on self-governing robots may prove to be a valuable resource for indoor agriculture and dairy farm administration in the future.

The objective of this paper is to present the evolution of SLAM since its inception, the technique used at each stage and their contributions to tackling various difficult applied research problems. The second section presents an overview of visual SLAM, its architecture and its different parts. The third section presents the probabilistic methods of SLAM. The fourth section is devoted to the SLAM based on vision. The fifth section introduces deep-learning-based approaches. The sixth section raises problems and challenges. A conclusion is drawn in section seven.

2. VISUAL SLAM OVERVIEW

VSLAM is an emerging embedded vision technology and is found very effective. The architecture of visual SLAM consists of three principal tasks: initialisation, localisation and mapping. The first phase of initialisation is to create a 3D initial map, made

possible by the extraction of feature points and then determining their 3D world locations from the depth image. The phases of tracking and mapping are applied simultaneously; tracking estimates the path of the camera by matching features and refining them by tracking the local map. Localisation computes the novel 3D map points. Fig. 1 presents the architecture of visual SLAM. To improve its performance, two modules have been added: relocalisation and global map optimisation. Sometimes, the tracking process fails due to several constraints, including rapid camera motion, disturbances, scenes without texture or a dynamic environment. To solve these problems, the task of relocalisation is necessary to compute the camera pose. While the camera is moving, a previously recognised image is captured, from which the loop-closing steps are designed.

The latter compares the current landmarks with the previous keyframes. The cumulative estimation error is generated at the map level. To get rid of this error, global map optimisation is usually done. This process is done to refine the map, taking into account the consistency of the entire map information.

Visual SLAM requires feature points from the environment, but it also requires static landmarks to provide an accurate approximation. In addition, a classic SLAM and a current SLAM make up the V-SLAM domain. The classic V-SLAM technique supposedly depends on the surroundings. With a moving camera, the surroundings are actually thought of as static. During the VO procedure, a number of dynamic feature points are considered in the real world. To find dynamic feature points and discard them from the V-SLAM estimate process, modern visual SLAM integrates the architecture of V-SLAM with object detection. This method of pose estimation and mapping lowers the amount of pose estimation and mapping error by accounting for the overall movement of dynamic objects in the scene. Without using any previous object models, the environment is examined to gather all relevant data, including dynamic, geometric and contextual information.

In recent works on the current V-SLAM, moving objects in a dynamic environment are estimated and then represented in a spatiotemporal map. The estimation of the cumulative error of localisation and mapping is decreased by the improvement in feature point selection. Differentiating between static and dynamic objects is one of the most crucial aspects of the V-SLAM method. As a result, scientists have created cutting-edge algorithms based on DL, computer vision and artificial intelligence. The three stages of the authorised discrimination process are as follows: detection of prospective dynamic objects based on categories of dynamic objects that have been established. The second stage, segmentation, is optional. The third stage, optical flow estimation-based motion detection of possibly dynamic moving objects, comes next.

The ability to obtain information about the location and shape of objects is a benefit of localisation and object detection. There are overlapping concepts in this work that need to be clarified.

Classification/recognition—Finding the identity of the object in an image is necessary for this activity. In another way, assign it to a category from a list of pre-established categories. The localisation process seeks to pinpoint the object's position and create a bounding box around it.

All objects in the image are identified and classified during the object-detection process. Each object has a bounding box around it and is assigned a category.

By constructing a pixel mask for each object in the image, the segmentation approach enables a deeper understanding of the scene. Semantic segmentation and instance segmentation are the two basic categories into which it is divided. Semantic segmenta-

tion-based classification represents all pixels that fall under a certain classification. Semantic segmentation groups related items and assigns them to a single class, rather than classifying pixels. Each pixel in the image is classified into a class using instance segmentation, and each class is then assigned to a different instance of the item. Fig. 2 presents different types of segmentation.

As a result of the relative movement between the observer and the scene, optical flow expresses the shape of the apparent movement of objects, texture and edges in a visual scene. The distribution of apparent velocities in the brightness level of the image created by moving objects is another way to describe it.
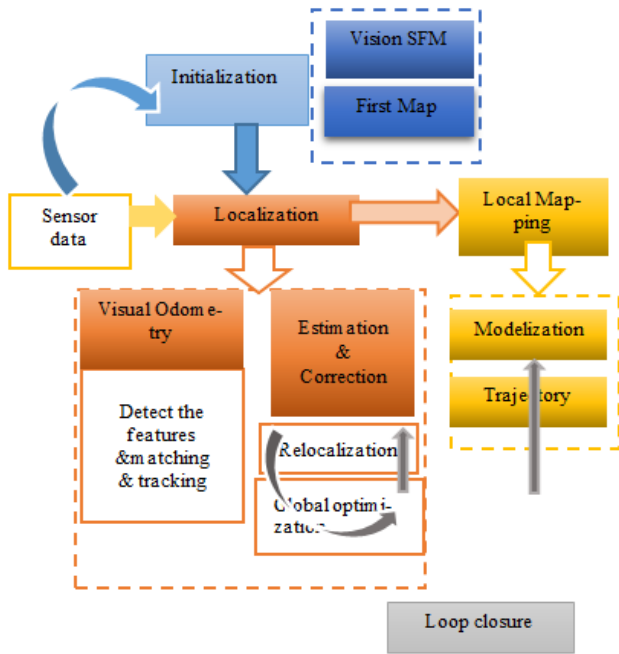
Fig. 1. SLAM Architecture, SFM, structure from motion; SLAM, simultaneous localization and mapping
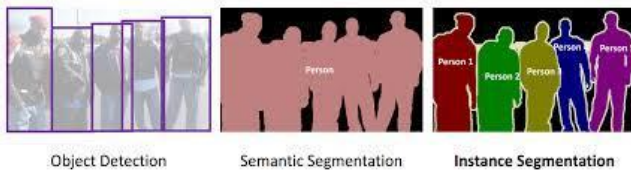
Fig. 2. Different types of segmentation

SLAM algorithms have evolved over time in response to the advancement of sensors, objectives and the pursuit of answers to specific issues in many research areas. Three phases can be identified in the evolutionary process. The initial phase was based on 1980 probability methodologies, which included filter-based techniques and optimisation-based strategies. The filtering techniques fit into iterative workflows that are appropriate for online SLAM. The full SLAM problem is addressed by the optimisation techniques, which group batch processing approaches. Classical sensors like Lidar, GPS and other sensors are the main focus of this phase. The second phase, based on computer vision and camera vision, was introduced in 2003. This technique is called vision-based SLAM. Its research advanced quickly and was able to resolve the SLAM issue and reconstruct 3D maps. The most recent phase, perception, began in 2014. The goal is to use learn-

ing to determine the system's correctness and robustness. It is based on DL, an algorithm that uses a convolutional neural network (CNN) to recognise objects in an image. Fig. 3 presents the phases of evolution.
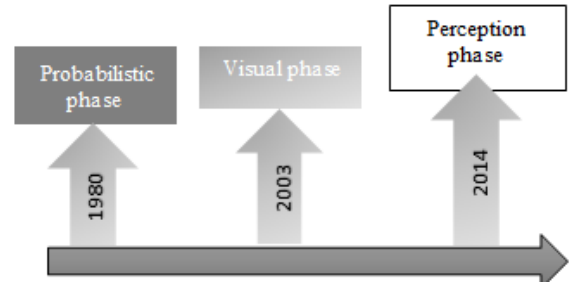
Fig. 3. SLAM evolution

3. SLAM PROBABILISTIC METHODS

3.1. Techniques based on filters

The methods based on filters are derived from Bayesian filtering. It works as an iterative process with two phases: prediction and correction. To forecast vehicle states and maps, the prediction phase first uses the evolution model and control inputs u_k. The second stage aims to correct the state that had been previously projected by comparing the map's current observation with the sensor data's current observation. The observation model combines mapping and observation. To estimate the location of the vehicle and the map, these two phases iterate and then gradually integrate sensor data.

Four main paradigms, on which various models have been created, form the foundation of SLAM. 'EKF' stands for extended Kalman filter. It is the oldest robotic system in terms of history. However, due to its restricted mathematical capabilities, it has lost some of its appeal. The second is the unscented Kalman filter that is known as 'UKF'. It was created to solve EKF issues with extremely non-linear systems. Information filtering ('IF') is the third strategy. It is the Kalman filter's (KF) inverse form. The fourth approach makes use of particle filters, which are non-parametric statistical filtering methods. They are widely used as online SLAM techniques and can address the issue of data association.

3.1.1. EKF

The first branch derivative of the KF created by Kalman (13) is the EKF. To handle linear systems, the KF was created. SLAM also employs it extremely rarely, despite its high convergence. On the other hand, The EKF can linearize non-linear systems using the first-order Taylor expansion(14).

EKF SLAM's first publications appeared in Refs (16)(17). Their method is based on estimating the movement of robot locations and a set of environmental characteristics using a single-state vector. A covariance matrix generates their estimation uncertainty as well as the correlations between the robot condition and the estimation of attributes. Using the EKF, the system's state vector and covariance matrix are updated (13)(18). When new features are noticed, more examples are added to the vector of states, and the system's covariance matrix grows significantly in size.

The EKF-SLAM methods have been extended by many authors to accommodate the issue. The calculation of the Jacobian and the linear approximation of non-linear models are two significant issues with EKF-SLAM. It can result in a filter inconsistency issue. A SLAM technique based on the central difference Kalman filter (CDKF) has been suggested by Zhang et al. (19) as a solution to this issue. For the purpose of approximating the non-linear models, the authors created the Sterling's polynomial interpolation method. It is based on trying to solve the SLAM issue in the probabilistic state space. The adaptive KF employing the AKF (20) has the benefits to include real-time processing. AKF has the ability to modify KF's parameters and improve the filtering. Additionally, the method can enhance the mapping and localisation accuracy by overcoming the challenge of information mismatch.

The adaptive EKF is a method that was proposed by Tian et al. (21) to enhance the conventional EKF. It is based on both maximisation of expectation creation (EM) and the maximum likelihood estimation (MLE). Its goal is to enable repeated approximations of the statistical noise and its covariance matrices by the standard EKF. As a result, EKF offers the capability to modify and improve the values created by MLE and EM production. Although it uses unbiased estimation to estimate the noise recursive statistics and produces high-quality results, one potential issue is that non-positive matrices of statistical covariance of process and measurement noise are defined. To lessen this issue, innovation covariance estimation (ICE) is added.

An autonomous wheeled mobile robot's SLAM problem is resolved using the suggested method (AEKF-SLAM). The disadvantage of this method is that it has a larger computational cost than EKF.

3.1.2. UKF

Julier and Uhlmann (22) introduced the unbiased KF, sometimes known as the UKF in the literature. The gradient calculation of the system equations is not explicitly used by the UKF algorithm, in contrast to EKF. The method relies on sampling particles, or a collection of points dubbed 'sigma', which are weighted around the expected value using a probability function and then passed to the non-linear function to recalculate the estimation. They enable accurate evaluation of the state vector distribution's mean and standard deviation up to the second-order approximation. Hence, to obtain the equivalent set of modified points, the sigma points are replaced. The UKF-SLAM was developed to address issues with the EKF, such as inconsistent performance. Traditional UKF-SLAM models the system as being precisely known and the perturbations as Gaussian noises with well-known statistics. Asymmetry in the actual applications could result from all these presumptions. These defects must be avoided to boost estimate precision; therefore, the Robust SLAM (RSLAM) was developed by Havangi (23). The H∞ square root UKF, which is applicable to non-linear systems with non-Gaussian disturbances, forms the basis of RSLAM. This technique has the benefit of not requiring knowledge of noise distributions or that they must be Gaussian, which makes it more adaptable and less constrained in practical applications. Additionally, because the resulting covariance matrices will continue to be semi-positive definite, RSLAM has steadily increased the numerical stability when compared to the UKF-SLAM technique. An adaptive neuro-fuzzy inference system is used to tweak the RSLAM parameters, producing better

performances. RSLAM thus outperforms other UKF-SLAMs, according to the Monte Carlo simulation.

An unscented adaptive Kalman filter (AUKF), employed on the SLAM problem both in the simulation data and in the actual application, was presented by Bahraini et al. [27, 28]. It is suggested that the scale parameter, which is based on the maximum likelihood function at each time step, be adjusted. The algorithm's results show that it reduces error estimation and increases navigational accuracy.

The covariance positive defined the positive loss prevents UKF from operating, and its strong correction amount reduces the SLAM algorithm's efficiency to improve the performance of UKF. Tang et al. (26) proposed an improved Schmidt Orthogonal Unscented Kalman Filter (ISOUKF). The approach is based on a two-step modification of the UKF algorithm. The Schmidt Orthogonal transform (SOT) sampling method is used in the first step to select sample points, and the SOT sampling approach is employed to lower the computational amount of UKF to some level. The notion of a strong tracking algorithm (27) then employs an adaptive fading factor, and the prediction covariance matrix uses the fading factor effect to boot system state tracking capacity. The ISOUKF algorithm was enhanced, and the SLAM technique was made more effective in the second stage using the square root. This technique lowers processing costs while presenting a high degree of precision in tracking robots for SLAM.

3.1.3. Information filter

IF is a KF variant that is the inverse of the KF, as shown by Maybeck (28). This filter adds the vector and the informational matrices directly, presenting an inverse information matrix of the covariance matrix, which contains the state error. The extended information filter (EIF), a non-linear version of the IF, is computationally comparable to the EKF with one key distinction: the EIF has an inverse covariance matrix.

The SLAM issue was also addressed by using the candidate's EIF techniques (29). The sparse extended information filter (SEIF) technique, which was introduced by Thrun et al. (30), has been suggested as another extension of IF. The IF can be upgraded to exactly sparse extended information (ESEIF) (31), which is more consistent locally than SEIF, by leveraging parsed data for less complexity. He et al. (32) proposed the iterative sparse extended information filter (ISEIF). By solving the measurement update equations iteratively and adaptively, this approach seeks to minimise linearisation errors. The consistency and accuracy of SEIF have increased because the scaling advantage is still present. However, IF has various uses as given in Refs. [36(34). It is not popular in SLAM.

3.1.4. Particle filter

This was proposed by Del Moral (35), under the name 'Particle Filters' also called 'bootstrap filter' (36), and 'Sequential Monte-Carlo (SMC)'' (37). It is a filter that allows for finding solutions to a problem of localisation. In the observation, it does not need the limitations of the Gaussian noise, and it can adapt to any distribution. It is based on a set of generated points called 'particles'. Each of these particles represents the probable state of the system. The weight coefficients (weights) on each particle are

a measure of the degree of confidence one may have in the latter to effectively represent the state. Their principle is as follows: samples of the state are taken with a set of particles according to their probability density. The particles are evaluated according to the equation of the state of the system; this is the prediction step, and then the weights are adjusted according to the observations; this is the correction step. The most probable particles are kept, the rest are removed and new particles are generated. Many versions of particulate filters have been proposed in the literature, such as sampling importance resampling (SIR) (38) and regularised particle filter (RPF) (39).

The first to make particulate filters adaptable to the SLAM problem was Blackwell (40), which is known as Rao–Blackwellisation. Doucet and Murphy (41) and Kevin Murphy (42) observed that the probability between the landmark sites is conditionally independent when the robot's route is known. Rao–Blackwellised (RB) decomposition was therefore introduced and carried out in a manner that added to the broad framework of PF for solving the SLAM issue. Based on this concept, Montemerlo et al. (43)(44) suggested the FastSLAM once more, this time utilising a few low-dimensional EKF to estimate the landmark locations and the Rao–Blackwellised particle filter (RBPF) to estimate the robot path. Stated differently, FastSLAM employs a hybrid technique that combines the PF with EKF, allowing the robot to attain more precision. The procedure in this method is predicated on the robot's prior posture prediction. Additionally, the method presumes that the landmarks are not conditionally dependent on one another while the robot's position is known. Furthermore, the robot localisation problem and the challenge of gathering estimated landmarks, both of which depend on estimating the robot's pose are separated from the SLAM by the method.

The computing complexity of FastSLAM, denoted as (M log N), is contingent upon the quantity of landmarks (M) and particles (N), both of which may have fixed values. Since every particle prescribes landmarks in a distinct way, FastSLAM performs several data associations, making the data association incredibly error-resistant. FastSLAM is easy to use and has a significant advantage in data association over EKF-based SLAM techniques, but, in some situations, the chosen samples are frequently ineffective. It is not necessary to linearise the robot's motion and measurement models. Its use in non-linear and non-Gaussian systems is more effective and convenient. The primary benefit of the FastSLAM approach is that particles carry out their own data associations, whereas the KF-based SLAM technique bases its system design on a single data association assumption for the whole filter. Furthermore, compared to KF-based approaches, the use of particle filters for robot trajectory sampling results in lower memory use and processing costs. However, because FastSLAM must perform an independent data association, it is vulnerable to divergence, and its computing cost increases significantly in noisy situations due to sparse maps. However, the limited feature position dependencies in FastSLAM instantiations lead to sluggish convergence. Moreover, the method's poor universal consistency renders it unsuitable for long-term navigation in expansive situations.

A better version of this technique, known as FastSLAM 2.0, was later published by Michael et al. (45). According to them, the proposed distribution of this approach depends on the actual measurement of the mobile robot as well as the previously estimated pose. Along with the improvements of FastSLAM 1.0, FastSLAM 2.0 also has an improved proposal distribution that results in a more consistent computing cost. The derivation of the Jacobian matrices and the linear approximations of the non-linear functions are two significant potential shortcomings of FastSLAM. It takes work to calculate the Jacobian matrix, and the estimate accuracy degenerates when the posterior covariance is not accurately approximated. To solve these problems, a novel method named Unscented FastSLAM (UFastSLAM) (46) was proposed to address linearisation-related issues in the FastSLAM framework. It is based on the use of scale unscented transformation. The linearisation procedure involving Jacobian computations is eliminated without the buildup of linearisation mistakes. This approach offers resilience in the mapping and localisation processes. However, the UFastSLAM often reduces particle diversity throughout the particle resampling process, and importance sampling is prohibited owing to covariance positive definite loss, resulting in accuracy degradation.

Variations of FP have appeared, such as distributed particle DP-SLAM approaches (47) and DP-SLAM 2.0 (48). These approaches proposed a data-storage structure based on the use of a minimal ancestry data tree. It makes quick updates by leading the PF while the number of iterations of the latter is reduced. In 2015, a new improved version of FastSLAM2.0 called six degrees of freedom (6-Dof) low dimensionality SLAM (L-SLAM) was developed by Zikos and Petridis (49). L-SLAM is based on using a particle filter of lower dimensionality than FastSLAM, for a small number of particles. L-SLAM achieved better accuracy than FastSLAM1.0 and FastSLAM2.0, and its speed surpasses FastSLAM2.0 by a factor of 3. L-SLAM is suitable for solving problems with high dimensions that have high computational complexity. To update the particles of the L-SLAM approach using a linear KF, in contrast we use an EKF to update the FastSLAM algorithm. To build a map by RBPF and ensure overall consistency, Nei et al. (50) presented an efficient system of RBPF that is an improved Lidar SLAM system by adding loop detection and correlation called LCPF-SLAM. The suggested LCPF SLAM enhances the consistency of the RBPF SLAM to be usable in comparatively wider scenarios. It also has enhanced loop identification and a new metric known as the usable ratio for determining the relevant information gained from laser readings. Still, the approach performs slowly since additional criteria are used to determine if a loop is reliable.

Resampling fixes the major flaw of the particle filter, the degeneracy of the weights, but after several iterations, particle diversity in particle concentration is completely absent; it is the problem of particle depletion. Hua and Cheng (51) proposed an adaptive fading unscented KF method (UFastSLAM) to solve the problem of particle degradation using the resampling method. It uses the UT transformation to eliminate the Jacobian matrix from the FastSLAM approach and improves the assessment of the position estimation. They replaced the KF with an unscented KF, which is suitable for non-linear systems. It also has other advantages in avoiding the accumulation of errors during linearity and has a better effect on pose estimation. In the UFastSLAM algorithm, the particles of PF are produced from the distribution of system state variables that do not depend on the posterior probability. It builds a proposed distribution function to edit and adjust the parameters adaptively and make the function of distribution closer to the system's posterior probability distribution. It is effective in improving the problem of particle degradation. An improved transformed unscented FastSLAM (ITUFastSLAM) with the adaptive genetic resampling ITUFastSLAM was introduced by Lin et al. (52). An improved importance sampling using the UKF was transformed to improve the performance of FastSLAM. A new fuzzy noise estima-

tor is used for the improvement, which allows adjusting the state and observation noises online according to the residual and associated covariance and results from it, attenuating the flaws resulting from the imprecision of the model. They replaced the step of conventional resampling with adaptive genetic resampling. Tang and Chen (53) developed an improved adaptive unscented FastSLAM (IAUFastSLAM) with genetic resampling, to ameliorate the low tracking accuracy. This algorithm uses QR and SVD decomposition to deal with the positive definite loss of covariance in the UKF and to give the system the ability to track. They used an adaptive factor consisting of the orthogonal principle of residual vectors to predict the covariance matrix, and the function of Huber cost is generated by the modified covariance matrix to effectively eliminate the error of the measurement model. To increase the particle diversity, they used an improved genetic approach (GA) and used the suppressed sample impoverishment effectively to complete the resampling for UFastSLAM.

3.2. Techniques based on optimisation

SLAM's improvement-oriented approaches branch out into two disciplines. The first subsection is based on finding a match between the novel observations and the map derived from the sensor data. The second subsection seeks to obtain a coherent whole by refining the car's position (and subtracting the past) and the map by looking at the constraints. When it comes to optimisation, we can classify these algorithms into two main branches: the SLAM graph and bundle adjustment (BA).

3.2.1. The graph SLAM

This is an algorithm that solves the SLAM problem owing to non-linear parsimonious optimisation. Lu and Milios (54) proposed this algorithm as the first work in robotics to solve the problem of SLAM, based on the graphical representation of the Bayesian SLAM shown in Fig. 4 (55). The graphic has been translated into a matrix that describes and combines the relationships between features and robot positions. It can easily be constructed for use to optimise the framework. The graph SLAM is based on two types of nodes: motion nodes and measurement nodes. Motion nodes connect two consecutive robot locations x_{t-1} and x_t. The measurement nodes connect the poses x_t, to the landmarks m_i. The graph edges present a non-linear constraint that represents the negative logarithmic likelihood of both the measurement and movement patterns. One of the greatest disadvantages of this method is the problem of the non-linear least squares produced by the sum of all the constraints. Many implementations are used to develop Graph-SLAM TORO (56), TreeMap (57), HOGMan (58), ISAM2 (59), g2o (60), GTSAM (61), DCS (62), SacViSLAM (63) and SSA (64).

Zhao et al. (66) proposed a method named LinearSLAM to solve the problem of large scale in SLAM based on a submap joining approach. The local sub-map is constructed using the local information to find the solution to a small-scale SLAM. The advantages of combining sub-maps include solving linear least squares and establishing non-linear coordinate transformations. This approach does not require initial values and iterations since there are closed form solutions to linear least squares problems. The algorithm can be used in pose-graph SLAM, D-SLAM, fea-

ture-based SLAM and in both 2D and 3D scenarios. Holder et al. (67) presented an algorithm that builds a map from radar detections by applying the iterative closest point (ICP) algorithm with the goal of matching successive scans given from a single radar sensor. Youyang et al. (68) proposed a G-pose graph optimisation algorithm that is an algorithm without having to handle the complex Bayes factor graph. In their proposed method, they transform the absolute pose estimation problem into a relative pose estimation problem. The main advantage of the G-pose graph optimisation method is its robustness to outliers. In fact, they added a loop closure metric to handle outliers. Fan et al. (69) presented the CPL-SLAM algorithm, which is efficient and certifiably correct. It uses complex numbers to solve SLAM based on a planar graph. Sun et al. (70) proposed an active integrated method by using the method of a Cartographer to build and do efficient frontier detection. Pierzchała et al. (71) used the Graph-SLAM algorithm to generate localised forest maps. With the aim of mapping, they collected the 3D data using a specially designed mobile platform composed of several sensors.
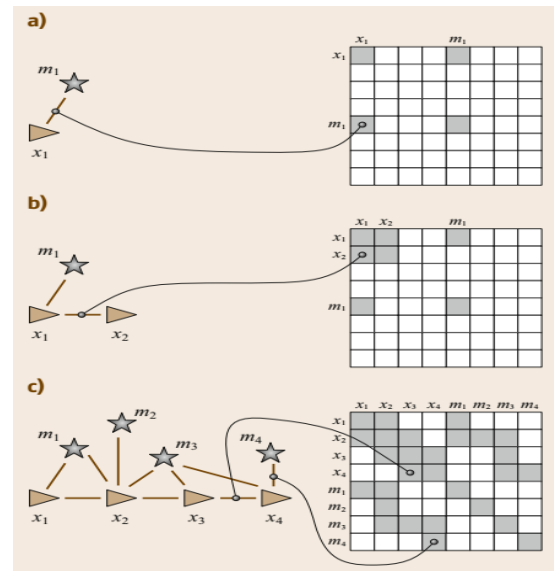
Fig. 4. (a–c) Schematic diagram of building a graph. The (a) diagram graph shows the observation 1s landmark m1, the (b) shows the constraints in the matrix form and presents robot motion from ×1 to ×2. The (c) shows several steps later (65)

3.2.2. Bundle adjustment

It is a vision technology that aims to refine a visual reconstruction of the three-dimensional structure and parameters of the camera (pose and calibrations). The symbol 'bundles' refers to rays of light leaving each 3D feature and converging on each camera centre. Then they are optimally 'adjusted' concerning both the feature and the positions of the camera. The main idea is optimisation, usually based on the objective function (ML) Levenberg-Marquardt algorithm (72). The optimisation of the best parameters (camera and landmark positions) is achieved by reducing some cost functions that determine the fitting error and finding the optimal solution concerning both structure and camera variations. To perform optimisations, many approaches have been proposed (73)(74).

In many works, the BA method is used in visual state-estimation problems SLAM and Visual-Inertial Odometry. The

objective of the BA approach is to estimate the 6-DOF camera path and 3D map (3D point cloud) according to the tracks of the input feature. One of the disadvantages of this algorithm is that it is computationally heavy and cumbersome because, in the process of optimisation, it takes all the variables at once. Second, 3D structure estimation needed a baseline sufficiently inherent in BA; the algorithm of SLAM will struggle in slow–motion periods or pure rotational motion. Melbouci et al. (75) proposed a method based on combining depth measurements and monocular visual information in a cost function fully presented in pixels. The idea is to consider sparse depth information as an extra constraint in BA. Frost et al. (76) presented a technique for integrating scale data from object classes into monocular visual SLAM based on BA. In Schops et al. (77), one finds a description of a methodology named fast-direct BA formulation that can be applied in a real-time dense RGB-D SLAM approach. This results in the use of rich information in global optimisation, which obtains paths with great precision. Zhao et al. (78) proposed a novel, rigorous and efficient method called good graph. They used a BA-based V-SLAM back-

end to improve their cost efficiency. Their objective is to define graphs with small sizes to be improved in the phase of local BA using preservation conditions. Wang et al. (79) introduced the SLAM framework based on saliency and the backbone given by ORB-SLAM3 (80). They developed the salient BA based on the saliency map value that can make the salient feature fully play its value. A new SLAM system is proposed by Gonzalez et al. (81). It uses the semantic segmentation of objects and structures in the scene. The authors modified the classical BA formulation using geometrical priors to constrain each cluster, which allows for improving both camera localisation and reconstruction and enables a better understanding of the scene. Tanaka et al. (82) proposed a learning-based BA based on a graph network. It replaces the standard Levenberg–Marquardt approach of BA with an algorithm based on learning. The advantage here is that it runs very fast and can be applied instead of conventional optimisation-based BA. Tab. 1 summarises the probabilistic methods by specifying the type of algorithm adopted for each method and the year they appeared.

Tab. 1. Probabilistic methods strengths and problems

Methods	Type	Algorithm	Year	Comment
SLAM- probabilistic methods				
Filters- Techniques	Extended Kalman	KF (13)	1986	Strength: -Efficient convergence -Adapt to uncertainty -The mean must be known. Problem: -Gaussian restrictions -The issue with high-dimensional maps.
		EKF (15)	1990	-Association of data in large environments -First order Tylor expansion. -Vulnerable to linearisation errors.
		CDKF (19)	2009	Strength: -Approximate the non-linear model-solve the SLAM issue in the probabilistic state space -Reduction in the ambiguity for data association. Problem: -Calculate the mean and covariance.
		AKF (20)	2016	Strength: -Gain adjustment in real time-accurate-robust mapping. -Strong estimation for AKF-Unbiased estimation for AEKF.
		AEKF (21)	2020	Problem: -High computational cost-Association data problem -Gaussian noise.
	Unscented Kalman	UKF (22)	2000	Strength: -Dealing with non-linearities-Coping with uncertainty -Expansion of the second order Taylor. Problem: -It is necessary to know the mean and covariance -Assumes that the system is exactly understood and that disturbances are stationary Gaussian noises with known statistics. The covariance positive defines loss and its calculation amount is large.
		AUKF [27(25)	2019	Strength: -Find the appropriate value for the scaling parameter and improve the estimate accuracy-Accurate.
		RSLAM (23)	2016	Strength: -Applied to non-linear systems with non-Gaussian noise -It is more flexible and adaptative -Has fewer limitations in real application.
		ISOUKF (26)	2022	Strength: -High precision-reduces computational cost -Accuracy and efficiency.
	Information	IF (28)	1979	Strength: -Straightforward and easy to execute-Handle maps with high dimension. Problem: -Challenges when integrating maps-Issue with connection of data.
		EIF (29)		
		SEIF (30)	2004	Strength: -Representation of graphical grids. -Sparsification-constant computational cost. Problem: -Inadequate representation-Iterative and slow.
		ESEIF (31)	2007	Strength: -More consistent with leveraging parsed data.
		ISEIF (32)	2015	Strength: -Measurement update equations iteratively and adaptively.
	Particle filter	PF (35)	1996	Strength: -Handles non-linearities -Handles with non-Gaussian noises. Problem: -Big complexity-Data Association.
		Rao-Blackwell PF (40)	2000	Strength: -Cost of calculation in logarithms-linearisation is not necessary-Accuracy. Problem: -Data association has a high cost-The landmarks' information is limited. -Higer dimensional map.
		FastSLAM (43)	2002 2007	Strength: -Higher accuracy-Path and landmark estimation 2007. -Does not necessary to linearise the robot's motion and measurement models.
		FastSLAM 2.0 (45)	2003	-Its use in non-linear and non-Gaussian systems -FastSLAM2.0 more consistent computing cost-FastSLAM2.0 linearises the non-linear model. Problem: -It must perform an independent data association -It is vulnerable to divergence. -It is computing cost increases significantly in noisy situations due to sparse maps -Universal consistency renders it unsuitable for long-term navigation in expansive situations. The derivation of the Jacobian matrices and the linear approximations of the non-linear functions.

		DP-SLAM (47)(48)	2003	Strength: -Data storage. -It makes quick updates.
		L-SLAM (49)	2015	Strength: -Uses small number of particles.-Better accuracy than FastSLAM. -Speed-Solve problems with high dimensions that have high complexity computational.
		LCPF (50)	2020	Strength: -Improved loop detection .-Detects the useful information obtained from laser readings. - Improve the consistency. Problem: -The approach performs slowly since additional criteria are used to determine if a loop is reliable.
		ITUFastSLAM (52)	2019	Strength: -Adjusts the state and observation noises online.
		UFastSLAM (51)	2020	Strength: -Solves the problem of particle degradation. -Uses the UT transformation to eliminate the Jacobian matrix.-Improves the assessment of the position estimation.-Avoids the accumulation of errors. Problem: -Reduces particle diversity throughout the particle resampling process. -The importance sampling is prohibited owing to covariance positive definite loss.
		IAUFastSLAM (53)	2021	Strength: -Ameliorates the low tracking accuracy. -Deals with the positive definite loss of covariance in UKF. -Predicts the covariance matrix. - Eliminates the error of the measurement model effectively.
Optimization technique	The Graph SLAM	TreeMap (57)	2006	Strength: -incremental optimisation approach-Update O.log N/ time. Problem: -Only provides a mean estimate
		TORO (56)	2008	Strength: -Optimisation strategy based on SGD. - Robust under the poor first predictions. -Assumes that constraints have covariance matrices that are generally spherical. Problem: -ecovers fast from big mistakes but has delayed minimum convergence. -Only provides a mean estimate.
		HOGman (58)	2010	Strength:- Incremental optimization approach via hierarchical pose graphs and lazy optimization. Problem: -Requires pose-graphs with full rank constraints.
		g^{2}o (60)	2011	Strength: -Flexible and readily adaptable SLAM optimization framework -It includes many optimisation methods and error routines. -External plugins are supported.
		iSAM2 (63)	2012	Strength: -General incremental non-linear optimisation with variable elimination. -Sparsity is preserved by variable re-ordering. -Relinearization of specified variables on demand.
		GTSAM (61)	2012	Strength: -Flexible optimisation framework for SLAM and SFM structure derived from motion. -Direct and iterative optimization approaches are used. -SAM, iSAM, and iSAM2 are all supported. -BA for Visual SLAM and SFM is implemented.
		SSA (64)	2012	Strength: -Optimises both robot positions and proximity sensor data. -Estimates the smoothness of a surface. -Assumes the presence of a range sensor (e.g., laser scanner, Kinect, or similar).
		DCS (62)	2013	Strength: -Outliers are dealt with by optimising with a strong cost function included into g2o.
		SacViSLAM (63)	2011	Strength: -For on-the-fly processing, it combines local bundle correction with sparse global optimization.
		LinearSLAM (66)	2018	Strength: -Solves the problem of large scale.
		G-pose graph (68)	2020	Strength: -handling the complex Bayes factor graph. -It is robustness to outliers. -Adds a loop closure metric to handle outliers.
		CLP-SLAM (69)	2020	Strength: -It uses complex numbers to solve SLAM based on a planar graph
	BA		1999–2023	-The primary concept is optimization. -Based on the Levenberg –Marquardt algorithm's objective function (ML). -The optimal parameters (camera and landmark locations) are optimised by lowering several cost functions that affect the fitting error. -Finds the best option in terms of structure and camera variations. -It is important to note that many SLAM methods developed after 2014 do not exclusively fall into the category of DL-based SLAM. Pose graph optimisation with BA remains a mainstream back-end algorithm.

AKF, adaptive Kalman filter, AEKF, Extended Adaptive Kalman filter, UKF, unscented Kalman filter, AUKF,(24)(23)(23) adaptive unscented Kalman filter; BA, bundle adjustment; (25)(24)(24)CDKF(19)(18)(18), central difference Kalman filter; DCS, *(62)(61)(61)*dynamic covariance scaling; DL, deep learning; (47)(46)(46)EIF, extended information filter; EKF, extended Kalman filter; ESEIF(31)(30)(30), exactly sparse extended information; IAUFastSLAM, improved adaptive unscented FastSLAM; IF(28)(27)(27), information filter; iSAM, incremental Smoothing And Mapping; ISEIF, iterative sparse extended information filter; ISOUKF, improved Schmidt Orthogonal Unscented Kalman Filter; ITUFastSLAM, improved transformed unscented FastSLAM; (52)(51)(51)KF, Kalman filter; L-SLAM, low dimensionality SLAM; (23)(22)(22)RSLAM, robust SLAM; *(63)(62)(62)*SAM, smoothing and mapping; SEIF, sparse extended information filter; SFM, structure from motion; SGD, stochastic gradient descent(13)(13)(13); SLAM, simultaneous localisation and mapping; SSA, sparse surface adjustment; UFastSLAM, Unscented FastSLAM, distributed particle DP-SLAM. iterative closest point (ICP), G-pose graph optimisation, CPL-SALM Correct Planar Graph-Based SLAM. LCPF: A Particle Filter Lidar SLAM, HOGman, Hierarchical optimization on manifolds, GTSAM, Georgia Tech Smoothing and Mapping, DCS, Dynamic Covariance Scaling.

4. VISUAL SLAM AND RGB-D-SLAM

4.1. Classical methods

In theory, the method of visual localisation uses the theory of geometry mainly to estimate motion; it is based on the extraction of geometric constraints from images. It is based on elegant well-established principles and is extensively studied. The VO algorithms can be classified according to the type of image used: stereoscopic or monocular VO. Their processing techniques are based on feature direct and indirect methods, which are 'appearance-based' and 'feature-based', respectively.

4.1.1. Feature-based methods

The first approach 'feature-based' or the indirect method is based on two steps: detecting and tracking a set of salient features of the image, such as corners and lines, and following them in the following images. The calculation of the Euclidean distances of each element, the points between frames and the displacement and the velocity vectors by using detectors such as: Feature From Accelerated Segment Test (FAST) (83), Speeded Up Robust Features (SURF) (84), Binary Robust Independent Elementary Features (BRIEF) (85), Oriented Fast and Rotated BRIEF (ORB) (86), Harris and Stephens (87) detected the corners. The features are used to estimate the camera's state and reconstruct the environment. This technique is able to deal with large motions from frame to frame due to the distinctiveness of the features and is ideal to optimise the motion of the camera and the geometric structure; BA is suitable for its use. Camera tracking depends on the geometric feature error by reducing the Euclidean distances between the two corresponding sets of geometric primitives in 2D or 3D. The geometric errors are classified into three types: 2D point-to-point error, 3D point-to-point error and 3D point-to-plan error (88). Several techniques have been developed for this approach:

MonoSLAM The first monocular V-SLAM was developed in 2007 (89)(90). They were based on estimating simultaneously the movement of the camera in 6-DoF and the 3D positions of the characteristic points of an unknown environment by applying an EKF and representing them as a vector of state in EKF. The disadvantage of this approach is that the computational cost increases proportionally with the size of the environment. The algorithm of parallel tracking and mapping (PTAM) has been proposed to solve this problem (91)(92). The PTAM algorithm divided both tracking and mapping into different threads on the CPU, which are run in parallel, and therefore the computational cost is not affected. PTAM is the first algorithm that integrates a BA optimisation process into real-time V-SLAM algorithms with freed-up computing capacity.

RGB-D-SLAM was proposed by Endres et al. (93). The approach created for SLAM is based on RGB cameras. This system enables it to handle challenging data in common indoor scenarios and is fast to work online.

ORB-SLAM was designed by Mur-Artal et al. (94). It is an extension of the main ideas of PTAM algorithms: location recognition (95), scale-sensitive loop closure (96) and use of co-visibility information for large-scale operations (63), with some improvements and novelties. Indeed, ORB-SLAM is a feature-based single SLAM system that operates in real–time; the third parallel phase is added to detect the loop closure [105(98). All these additions make the system efficient and reliable.

ORB-SLAM2 was proposed by Mur-Artal and Tardos (99). It is an extension of the ORB-SLAM algorithm. It is suitable for monocular, RGB-D and stereo cameras and allows the reuse of maps, relocalisations and loop closing. In RGB-D results, the use of BA presents more precision than the methods of ICP or photometric and depth error minimisation. In stereo SLAM, they used near and far stereo points and monocular observations; the results depicted a high accuracy compared to the direct method. It allows reusing the map with mapping disabled by using the lightweight localisation mode.

OpenVSLAM was proposed by Sumikura et al. (100). It is a visual SLAM framework; it corresponds to a monocular, stereo and RGBD visual SLAM system, which contains a basic SLAM algorithm. These modules allow creating local and global maps and store and load them.

UcoSLAM was developed by Muñoz-Salinas and Medina-Carnicer (101). It is a monocular V-SLAM system fusing natural and artificial features to have strong long-term tracking. This gives the system an advantage; it can initialise both markers and key points. It makes the real scale of the maps accessible as long as a marker is available. It can solve problems caused by repetitive environments, false relocalisations and loop-closures by using the markers. It is distinguished from ORB-SLAM2 in that it can load and store the generated maps. The main idea to combine the plane and edge features was proposed by Sun et al. (102), named plane-edge-Slam. This methodology estimates robust motion, which depends on constraint analysis and an adaptive weighting algorithm.

4.1.2. Appearance-based methods

The second technique, 'appearance-based' or the direct method, estimates camera movements directly using pixel-intensity changes, usually photometric errors. The pixel selection can be all pixels (dense) or a sparse selection (sparse). The direct method eliminates feature extraction time at a cost that is much greater for optimisation problems than the feature-based method.

DTAM: The first direct method is called 'Dense Tracking and Mapping' and was published by Newcombe et al. (103). It is a method for tracking and reconstructing images from live cameras. To monitor the dense camera, it records the full image with the intention of creating a dense 3D surface model and using it right away. This approach offers keyframe tracking based on the reduction of photometric errors but does not include the closure-detection procedure or global optimisation.

LSD-SLAM: Engel et al. (104) created the large-scale semi-dense (LSD) SLAM. It uses the monocular camera VO technique. To estimate a semi-dense inverted depth map of the current frame, the primary idea is to use dense image alignment to track camera movement. The semi-dense VO was extended to the LSD-SLAM by Engel et al. (105). The recent advancements in this technique are based on a scale-aware image alignment algorithm to increase the similarity transform $\xi \in sim(3)$ between two keyframes. It is a monocular SLAM system that seeks to preserve and track the global map of the environment. The authors propose a new direct tracking method that allows for detecting and explaining scale drift. They developed a probabilistic method for the fusion of noisy depth estimation with tracking. In 2015, Engel et al. (106),(107) used LSD-SLAM with stereo cameras and omnidirectional cameras.

SVO 'semi-direct visual odometry': It is a reliable semi-direct monocular VO algorithm, proposed by Forster et al. (108). They used a probability mapping method that explicitly models external observations to estimate the 3D points, which results in fewer outliers and more accurate points.

DSO 'direct sparse odometry' was created by Engel et al. (109). It aims to combine a model that minimises optical error (full direct probabilistic) and optimisation for all model parameters represented by the intrinsic camera, extrinsic camera and inverse depth value. It is based on continuous photometric error optimisation over a window of recent frames, accounting for the model of a photometrically calibrated image. Gao et al. (110) proposed the LDSO, which is a development of the DSO that adds closing loop detection and pose-graph optimisation. They used a conventional feature called bag-of-words (BoWs) to inject the feature points into the loop closure (95). Another extension of DSO, called dynamic-DSO is proposed by Sheng et al. (111). It is a semantic direct VO of monocular vision using DL in the process of semantic-image segmentation. They applied CNNs to the original RGB image to extract the pixel-level semantic information of dynamic objects.

KinectFusion was introduced as a real-time mapping system in complex conditions and changing lighting by using a moving depth-camera called 'hand-held Kinect' and commodity graphics hardware (112). The obtained current sensor position tracks the live depth frame relative to the global model by applying an iterative nearest point (ICP) algorithm.

RGB-DTAM developed by Concha and Civera (113) introduced a direct RGB-D SLAM system with the ability to close the loop and reuse the map. With advanced technology, the approach allows accuracy and durability at a low cost. The inclusion of multiple RGB visibility limitations in thread tracking and mapping is the technique's key innovation. Extending the RGB-D sensor range, using high-parallel setups, and adding distant locations to the map all improve estimation accuracy.

ID-RGBDO proposed by Fontán et al. (114) aims to achieve great accuracy in calculating the direct speed of RGB-D with minimal losses. Therefore, they introduced new, efficient information to determine the most informative measurements in BA and position-tracking optimisations.

4.1.3. Semi-direct

Another highly popular approach is called semi-direct; it combines the benefits of the two methods mentioned above as well as the success aspects of the feature-based process, such as tracking numerous features, parallel tracking and mapping, with the accuracy and speed of direct methods.

CPA-SLAM was developed by Ma et al. (115). This technique combines frame-to-keyframe and frame-to-plane data. It is co-optimised with alignment constraints between keyframes for global consistency. This technique creates a global model that enables position estimation. A world map is made by segmenting the RGB-D picture planes using the 'agglomerative hierarchical clustering' method and an information association rule. The CPA-SLAM technique provides a photometric residual, a point-to-point residual and a plane-to-plane residual, which use the EM frame to minimise jointly to estimate the camera position.

BundleFusion proposed by Haque et al. (116) is a global pose-optimisation framework, the parallelisable sparse-then-dense. It is a method that accomplishes robust tracking while performing online real-time 3D reconstruction. Additionally, by improving the path globally for each frame retrieved, the loop-closure problem is solved.

KDP-SLAM 'keyframe-based dense planar SLAM' was proposed by Hsiao et al. (117). To estimate odometry, they used a fast dense approach. The depth values from small baseline images are combined in a local map to build dense 3D structures and extract planes. Then they used the method of incremental smoothing and mapping (iSAM) to optimise the positions of keyframes and landmark planes.

FSD-SLAM' fast semi-direct SLAM' was created by Dong et al. (118). This method's goal is to combine the feature point approach with a direct way to estimate and enhance the system's accuracy in a setting with few visual elements and little texture. Based on the sub-graph, a reliable feature point-extraction technique was selected. They suggested a reliable technique based on apparent shape-weighted fusion to determine the camera's position. The incremental dynamic covariance scaling (DCS) approach reduces the inaccuracy in calculating the camera location. They suggested a face element model based on the improved camera position to obtain a flawless 3D point cloud map as well as estimate and integrate the point cloud pose. Tab. 2 summarises all these techniques in the order present in the text.

4.2. Visual-inertial odometry (VIO) methods

The combination of an IMU and a VO system is the foundation of the VIO technique. The fundamental concept is to combine visual data with inertial measures to produce a more accurate and effective measurement. IMU is characterised by strength in certain situations, such as speed motion, textureless and lighting changes. Therefore, IMUs are used because they provide reliable information that we can use instead of visual information, or they add information in typical cases.

VIO systems may be classified into two primary streams: loosely coupled and tightly coupled techniques, based on directly or indirectly fused readings from sensors. In loosely coupled techniques, pictures and IMU measurements are processed by two estimators that estimate relative motion independently. The final result is obtained by fusing the estimates from the two estimators. Tightly coupled techniques combine raw data from the camera and IMU directly into one estimator to find optimum estimates. Tightly connected techniques are often more accurate and resilient than weakly coupled approaches.

ROVIO is presented by Bloesch et al. (119) as a monocular VIO method. It uses the errors of pixel intensity from image patches, which gives accurate and robust tracking. After detection, the multi-level correction feature tracking is based on a basic EKF by directly using the errors of intensity.

MSCKF-VIO stands for multi-state constraint KF used in stereo VIO without GPU. It was proposed by Sun et al. (120) as an approach that proved its accuracy, efficiency and durability compared to other algorithms. This method uses the multi-state KF, which was developed by Mourikis and Roumeliotis (121). It is used in stereo VIO without a GPU.

OKVIS 'Open keyframe-based visual inertial SLAM' is provided by Leutenegger et al. (122). It is a tightly coupled framework presented as a combination of both inertial measurements and image key points. The goal is to form keyframes in the problem of non-linear optimisation that uses linearity and marginalisation.

Maplab was developed by Schneider et al. (123). It is a platform written in the C++ language for visual-inertial mapping. It is a

system ready for planning and visual localisation and offers researchers a set of tools for multi-session mapping that allow map

merging, loop closure and inertial batch optimisation.

Tab. 2. Comparison of visual SLAM methods

V-SLAM methods	Name	Year	Camera Model	Back-End	Mapping	Relocalisation	Loop-closure
Feature-Based	Mono-SLAM (89), (90)	2007	Monocular	Filter-based	Sparse	No	No
	PTAM (91)	2007	Monocular	Optimisation	Sparse	No	No
	RGB-D-SLAM (93)	2012	RGB	Optimisation	Dense	No	Yes
	ORB-SLAM (94)	2015	All types	Optimisation	Sparse	Yes	Yes
	ORB-SLAM2 (99)	2017	All types	Optimisation	Sparse	Yes	Yes
	OpenVSLAM (100)	2019	All types	Optimisation	Sparse	Yes	Yes
	UcoSLAM (101)	2019	All types	Optimisation	Sparse	Yes	Yes
Appearance-Based	DTAM (103)	2011	Monocular	-	Dense	No	No
	LSD-SLAM (104)	2014	Monocular	Optimisation	Semi-Dense	Yes	Yes
	SVO (108)	2014	Monocular	-	Sparse	No	No
	DSO (109)	2017	Monocular	-	Semi-Dense	No	No
	LDSO (110)	2018	Monocular	Optimisation	Semi-Dense	Yes	Yes
	Dynamic-DSO(111)	2020	Monocular	Optimisation	Semi-Dense	No	No
	KinectFusion (112)	2011	RGB-D	-	Dense	No	No
	RGB-DTAM (113)	2017	RGB-D	-	Semi-Dense	No	No
	ID-RGBDO (114)	2020	RGB-D	-	-	No	No
Semi-Direct	CPA-SLAM (115)	2016	RGB-D	Optimisation	Dense	No	Yes
	KDP-SLAM (117)	2017	RGB-D	Optimisation	Dense	No	Yes
	BundleFusion (116)	2022	RGB-D	Optimisation	Dense	Yes	Yes
	FSD-SLAM (118)	2022	All type	Optimisation	-	No	Yes

DSO, direct sparse odometry; FSD-SLAM, Fast Semi-Direct SLAM; LSD, large-scale semi-dense; ORB, oriented fast and rotated BRIEF; PTAM, parallel tracking and mapping; SLAM, simultaneous localization and mapping, Mono-SLAM, monocular Visual SLAM, RGB-D-SLAM, Red-Green-Blue-Depth-SLAM, DTAM, Dense Tracking and Mapping, LSD-SLAM, large-scale semi-dense (LSD) SLAM,SVO semi-direct visual odometry, DSO, direct sparse odometry, LDSO: Direct Sparse Odometry with Loop Closure ,KDP-SLAM,'keyframe-based dense planar SLAM, ID-RGBDO, Information-Driven Direct RGB-D Odometry, FSD-SLAM, fast semi-direct SLAM, CPA-SLAM, Consistent Plane-Model Alignment, KDP, Keyframe-based dense planar SLAM.

ICE-BA stands for incremental, consistent and efficient bundle adjustment developed by Liu et al. (124). It gives a solution accurately and efficiently compared to traditional solutions. It used a larger number of measurements to achieve higher robustness and accuracy. It is based on solving the global consistency problem to ensure the minimisation of the reprojection function and inertial constraint function during loop closure.

SVOGTSAM was proposed by Forster et al. (125) as a program that aims to develop a new theory for the pre-integration stage. It deals with the multiple structures of the rotation group. It operates on the generative scaling model in addition to the nature of the rotation noise and determines the expression for the maximum post-state estimator. It integrates the IMU model into an inertial pipeline under the unified factor graphics framework. Therefore, it is allowed to use the method of incremental-smoothing and the use of a structureless model for visual measurement, which increases computation speed by avoiding optimization via 3D points.

VI-DSO 'direct sparse visual-inertial odometry' is an extension of DSO that uses inertial information developed by Von Stumberg et al. (126). The objective of this algorithm is to find the position of the camera and sparse scene geometry by reducing an energy function that combines the photometric and IMU measurement errors. They used the 'dynamic marginalization' approach in order to achieve marginalisation adaptively.

VINS-Mono 'A monocular visual inertial system' is presented by Qin et al. (127). It is a robust and versatile approach based on a low-cost IMU and a single camera to determine the 6 degrees of

freedom state of the system. The main contributions of this approach presented are a high-precision VIO measurement obtained by integrating IMU measurements and feature observations using a tightly correlated non-linear optimisation-based method. Integrating the module of loop detection with a tightly coupled formula that allows relocalisations with minimal computational cost to achieve global consistency, they optimised the pose graph for four degrees of freedom. This algorithm has been further developed in many research papers (128)(129).

PL-VIO, which is an acronym for point-line-visual inertial odometry, was proposed by He et al. (130). It is a strongly connected point-and-line-based monocular VIO system. Compared to dot features, lines provide more information about the environment's geometric structure. To determine the representative pressure of a 3D spatial line and the ease of calculation, Plucker coordinates and an orthogonal representation of the line are both employed. States are optimised by lowering a cost function, owing to the tightly and effectively integrated information between IMU and optical sensors.

Trifo-VIO (Trifo visual inertial odometry) proposed by Zheng et al. (131) utilised points and lines in a stereo VIO system with tightly coupled filtering. They create a novel technique for closing loops based on light filtering developed as EKF updates, which correctly repositions the sliding window now in use and keeps the filter active to detect loops. They make use of IMU data from the Trifo ironsides sensor, stereo camera data and the Trifo Ironsides dataset.

Co-Planar parametrisation for stereo-SLAM and VIO pipeline

was proposed by Li et al. (132). By creating efficient and reliable parameters for co-planar points and lines that make use of particular geometrical restrictions, this method intends to increase the camera positioning's efficiency and accuracy. The pipeline comprises extracting 2D points and lines, forecasting planar areas with random-sample consensus (RANSAC) and outlier filtering. Two steps are used in the detection of RGB images: robust outlier filtering and the deployment of a NN for planar segmentation. They employ the smaller and more sparsely distributed Hessian matrix, which optimises BA, to determine new parameters for points and coplanar lines to unify the parameters.

Mesh-VIO (133). They devised a method for building a 3D mesh progressively by limiting its extent to the time horizon of VIO optimisation in order to get a representation of the topology of the environment. The 3D mesh offers a richer and lighter model that seeks to identify and enforce structural regularities in the optimisation problem.

ORB-SLAM3 (80). It is a comprehensive system that uses monocular, RGB-D and stereo cameras to perform visual SLAM, visual-inertial SLAM and multi-map SLAM. It is a visual-inertial SLAM system that utilises the maximum-a-posteriori (MAP) estimation method. In both large and small environments, indoors and outdoors, it has a reliable real-time result. Additionally, it has an accuracy that is 2 to 10 times better than earlier techniques. It is a multi-map system built on a cutting-edge method for position identification with better recall. The outcomes demonstrate the precision and reliability of the ORB-SLAM3 system.

HybVIO (134) is a hybrid approach that combines optimisation-based SLAM and filter-based inertial optical measurement (VIO) to estimate ego-motion. The contributions of this strategy include the development of a probabilistic inertial visual odometry (PIVO) methodology that can be used for monocular or stereo applications; the modelisation of the IMU bias in PIVO using the Ornstein–Uhlenbeck random walk approach and using improved and derived mechanisms for aberration detection, stability detection and feature path selection that take advantage of the special characteristics of the probabilistic framework. In real-time, this

technology offers exceptional performance. Tab. 3 summarises the VIO methods in the order present in the text.

5. SLAM DL METHODS

A branch of machine learning called DL is based on artificial NNs. It has more than two layers built on algorithms that can be trained to process nonlinear data (Fig. 5). The learning field is characterised by supervised methods and unsupervised ways of learning. CNNs, recurrent neural networks (RNNs) and other designs are used in DL for a variety of tasks. The training process in the supervised learning method requires supervision and labelled data. Its objective is to train the model so that, given fresh data, it can forecast the outcome. An unsupervised learning method uses unlabelled data without the need for supervision during training. Their main objective is to find hidden patterns and useful insights from the unknown dataset.
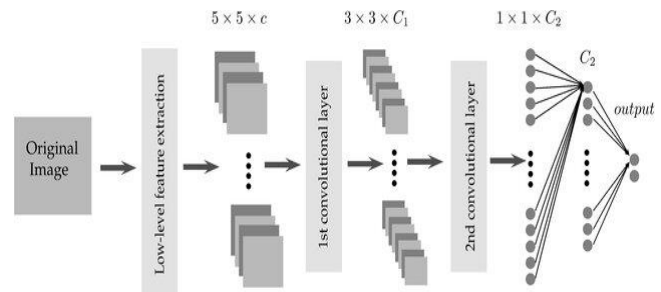
Fig. 5. DL architecture

Numerous papers filed by researchers in the SLAM field have combined DL with visual SLAM to address issues and create algorithms. This section outlines the applications of DL across several SLAM components.

Tab. 3. Comparison of VIO methods

Name	Year	Back-End-Approach	Camera Type	Fusion Type	Mapping	Loop closing	Relocalization
OKVIS (122)	2014	Optimisation-base	Monocular	Tightly coupled	Sparse	No	No
ROVIO (119)	2015	Filtering based	Monocular	Tightly coupled	Sparse	No	No
MSCKF-VIO (120)	2018	Filtering based	Monocular/stereo	Tightly coupled	Sparse	No	No
SVOGTSAM (125)	2017	Optimisation-base	Monocular	Tightly coupled	-	No	No
Maplab (123)	2018	Filtering based	Monocular	Tightly coupled	Dense	Yes	No
ICE-BA (124)	2018	Optimisation-base	-	-	-	Yes	No
VI-DSO (126)	2018	Optimisation-base	Monocular	Tightly coupled	Sparse	No	No
VINS-Mono (127)	2018	Optimisation-base	Monocular	Tightly coupled	Sparse	Yes	Yes
PL-VIO (130)	2018	Optimisation-base	Monocular	Tightly coupled	-	No	No
Trif-VIO (131)	2018	Filtering based	Stereo	Tightly coupled	-	Yes	No
Co-Planar (132)	2020	Optimisation-base	Stereo	Tightly coupled	Dense	No	No
Mesh-VIO (133)	2019	Optimisation-base	Stereo	Tightly coupled	Dense	No	No
ORB-SLAM3 (80)	2021	Optimisation-base	All	-	Sparse	Yes	Yes
HybVIO (134)	2022	Optimisation-base	Monocular/stereo	Loosely coupled	Sparse	Yes	No

ROVIO, Robust visual inertial odometry; ICE-BA, innovation covariance estimation-bundle adjustment; Mesh-VIO, MSCKF-VIO, multi-state constraint Kalman filter-visual-inertial odometry; OKVIS, open Keyframe-based visual inertial SLAM; ORB, oriented fast and rotated BRIEF; SLAM, simultaneous localisation and mapping; Trifo-VIO, Trifo visual inertial odometry; VIO, visual–inertial odometry; VI-DSO, direct sparse visual-inertial odometry; PL-VIO, point-line-visual inertial odometry; VINS-Mono, A Robust and Versatile Monocular Visual-Inertial State Estimator; HybVIO, hybrid visual–inertial odometry.

5.1. Initialisation

One of the most crucial deficiencies is the inability to estimate scale and determine depth during the initialisation phase of monocular visual SLAM. The issue of depth has been addressed in a number of works, some of which include optical flow, ego-motion, scale ambiguity and drift based on DL. We mention some of the most important works : ADAADepth (135), CNN-SLAM (136), Code-SLAM (137), DeepVO (138), UnDeepVO (139), D3VO (140), GeoNet (141), L-VO (142) and Un-L of depth (143).

The algorithm for creating the depth and disparity map in stereovision consists of four steps: feature extraction, feature matching across pictures, computation of disparity and disparity refining and post-processing. Researchers concentrated on estimating these stages using DL techniques. The primary goal of DL in stereo matching is to substitute learned features for manually specified characteristics: (144)(145). Mayer et al. (146) used DL to estimate the disparity, scene flow and optical flow. The works that have been done in Refs (147)(148) used DL to estimate the depth for disparity maps. Song et al. (149) suggested a technique end-to-end that allows to predict the edge map and disparity.

5.2. The front-end enhanced by DL

The two key components of this step are feature point extraction from the successive photographs and VO. The feature points enable the estimation of the camera movement alone without taking into account the entire map. The three-dimensional pose translation and three-dimensional pose rotation make up the estimation of the 6-DoF motion state. This motion estimation is applied to feature points throughout the tracking task using a RANSAC-based matching procedure. To estimate the homogeneous transformation between frames, as well as to ascertain the camera's present location and the environmental characteristics, matching is used.

Shao et al. (150) created a faster region-based convolutional neural network (Faster-R-CNN)-based semantic filter to address the issue of outliers in RANSAC-based F-matrix calculations. The semantic filter's training phase relies on semantic patches created by inliers, which enables various picture regions to define various semantic labels. The approach improves and increases the precision of F matrix calculations.

Zhang et al. (151) focused on the use of visual semantic information in the problem of camera localisation. They suggested a coarse-to-fine strategy in the visual localisation method and created a visual semantic database based on a deep-learning algorithm.

The approach based on integrating DL and machine learning with 2D-SLAM grid maps was proposed by Lin et al. (152) to estimate 2D object segmentation, feature extraction and pattern identification. DL is used by Wang et al. (153) to complete monocular VO in a comprehensive manner. The stances are calculated based on the actual scene. Deep neural networks (DNNs) understand the intricate dynamic motion of image sequences to do sequence-to-sequence posture estimation. By combining the RGB-D SLAM with optical flow-based feature tracking, Li et al. (154) improved the SLAM algorithm. To achieve the function of object detection, they combined 101 layers of deep residual networks (ResNet) with region-based fully convolutional networks (R-FCN).

V-SLAM-CNN (155): Current systems combine DL to automate surgical instrument and workflow identification in order to decrease surgical problems and ensure correct performance. In this study, visual SLAM and Mask R-CNN are combined. They employ V-SLAM for object detection, drawing on geometry data for area recommendations and CNN for object recognition, classifying images using semantic data, and combining these techniques into a single end-to-end training assignment. They are based on visual characteristics and spatiotemporal data gathered from video. By substituting a region proposal module (RPM) for the region proposal network (RPN) in mask R-CNN, bounding boxes are placed precisely, and the need for annotations is decreased. DVS-SLAM: A visual semantic map in a dynamic situation is called a dynamic visual semantic SLAM (156). They employed SSD-MobileNetV2 lightweight DL to obtain the 2D data.

Some studies use supervised or unsupervised learning techniques to estimate the absolute 6-DoF posture or the relative transformation matrix when implementing an end-to-end VO system. The deep convolutional generative adversarial networks (GANs)-based unsupervised learning framework GANVO (157) predicts 6-DoF posture camera motion and a monocular depth map of the scene from unlabelled RGB image sequences. A supervised monocular VO system is called DL_Hybrid (158). It is based on recovering camera trajectory and estimating 6-Dof posture frame-by-frame. First, they concentrate on the DL_Hybrid VO system overview. The dense optical flow map between picture frame pairs is then estimated using a DL NN called DenseFlowNetwork, and the dense depth map per-frame is extracted using a different DL NN called DenseDepthNetwork. Finally, the true monocular scale-estimation methodology is applied frame-by-frame as we describe the hybrid 2d–2d and 3d–2d posture-estimation approach paired with optical flow map and depth map.

Liang et al.'s (159) successful direct sparse VO approach is called SalientDSO. It blends DSO with semantic data in the form of visual saliency. SalientDSO is based on a deep-learning visual saliency and scene-analysis method that selects a feature for accurate and reliable VO. Their contributions help to present a framework of indoor VO in which the features are selected based on a visual saliency map. The authors suggested a method for filtering the saliency map based on scene parsing. A DL technique called GCNv2, which is an extension of the 'Geometric Correspondence Network', depends on a network created by Tang et al. (160) to identify the salient features and descriptors. A binary descriptor vector serves as the ORB feature in GCNv2. In addition to having more computational efficiency than GCN, GCNv2 also maintains accuracy levels comparable to GCN, which results in observable advancements in movement estimation. They used feature vector binaries in the training phase, which significantly accelerated matching.

SuperPoint (161), a self-supervised system for training to detect and describe interest points for the issues of a large number of multiple-view geometries, is used to identify and describe points of interest. A complete CNN is the SuperPoint. It operates on full-size images, producing the interest point detection with fixed-length descriptions in a single forward pass Kwang. et al.'s 'Learned Invariant Feature Transform' (LIFT) was proposed (162). The detector, orientation estimator and descriptor are the three CNNs-based components that make up this system.

SIVO(semantically informed visual odometry and mapping) (163) is founded on a system that chooses which feature to use for V-SLAM. It incorporates NN uncertainty and semantic segmen-

tation into the feature-selection procedure. Since the new feature is added with feature entropy classification from the Bayesian NN, the approach finds the spots offering the biggest Shannon entropy drop between the entropy of the current state and the entropy of the shared state.

To extract the binary visual feature descriptors with triplet loss, even distribution loss, correlation loss and quantisation loss, Gu et al. (164) created the DBLD-SLAM 'deep binary local descriptor'. They create a CNN model with four fundamental loss functions to extract binary visual feature descriptors from picture patches: adaptive scale loss, even distribution loss, quantisation loss and correlation loss.

Based on this learned deep binary feature descriptor, which has the same structure as the ORB descriptor, a monocular SLAM system called DBLD-SLAM is built, with the ORB descriptor replaced by conventional ORB-SLAM. They also train Bag of Words to recognise loop closures visually.

The foundation of ORBDeepOdometry (165) is a method for integrating DL with pipeline methodology to tackle the monocular VO problem. It models sequential data by stacking multiple deep LSTMs, feature extraction ORB and dimensionality reduction based on CNN.

It was suggested to use the 'Criss-Cross Network' (CCNet) (166) to obtain contextual information for the entire image. It is developed by utilising the criss-cross recurrent attention module to get the best results in benchmarks dependent on segmentation, such as Cityscapes, ADE20K and COCO. A general framework called MonoGRNet (167) is used to learn how to detect monocular 3D objects based on geometric reasoning, the observable 2D projection and the depth dimension that is not being seen. This method splits the job into four smaller tasks – 2D object identification, instance-level depth estimation, projection 3D centre estimation and local corner regression – and uses the network to perform each of them simultaneously.

Tab. 4 summarises the front-end methods enhanced by DL in the order they appeared in the text.

5.3. Back-end enhanced by deep learning

This step aims to enhance this estimation through tasks involving localisation, optimisation and loop closure.

5.3.1. Optimisation

The global optimisation process aims to maintain the geometric consistency of the full map. It has been made for localisation and mapping tasks.

To estimate the motion, the sequence-to-sequence learning algorithm VINet (168) was developed. It is supported by optical and inertial sensors. For the VIO, it is an end-to-end system that is completely trainable. The authors suggested a method for training the architecture's parameters as well as a design for recurrent networks.

A brand-new frame-to-frame estimation technique called Deep_VO (169) makes use of CNN to forecast camera motion. The best visual feature and the best estimator for visual ego-motion estimation are both learned using the CNN architecture.

Tab. 4. Comparison of front-end methods

Name	Year	Architecture/method	Main contribution
LIFT (162)	2016	-CNNs	-Learning invariant features
Faster-R-CNN (150)	2020	-CNN -Semantic filter	-It solves the outlier problem in F-Matrix computations based RANSAC.
A 3D Semantic Visual SLAM (156)	2021	- Mask R-CNN/MySQL - It creates a semantic database based on the information contained in the object.	-It is beneficial for localisation. -The accuracy and efficiency of the localisation.
V-SLAM-CNN (155)	2022	- Mask R-CNN. - It combines the greatest features of both worlds, such as (1) object detection using vSLAM and (2) CNN for identifying objects. -Spatio-temporal information.	-Concentrating on geometric data for suggested regions. -Concentrating on semantic data for picture classification and merging them into a single, collaborative, end-to-end training procedure.
DVS-SLAM (156)	2021	-Multi-view geometry and region growing algorithm. -SSD-MobileNetV2 lightweight DL. -Colour bumpy supervoxel clustering algorithm.	-Creating a visual semantic map. Removing dynamic feature points will improve localisation accuracy. -Get 2D data. -Achieve the extraction of 3D target information.
GANVO (157)	2019	-unsupervised learning framework. -GANs	-Predicts 6-DoF pose camera motion and camera depth.
DL_Hybrid (158)	2021	-Hybrid 2D–2D and 3D–2D localisation theory. -DNN 'DenseDepthNetwork' -DFN 'DenseFlowNetwor'	-One-frame-at-a-time estimation of a six-degree-of-freedom pose and camera trajectory recovery are possible. -Accurate key points extracted from each frame even in harsh scene conditions, and the system performs well even in situations where motion is restricted to the camera, such as when it is rotating or stationary. -Large-scale displacement motion of the camera is also a possibility.
SalientDSO (159)	2019	-High semantic information -CNNs+VO	-Drive feature selection for visual saliency. -Offers a technique for saliency map filtering depending on scene parsing.

GCNv2 (160)	2019	-Geometric Correspondence Network -Incorporates feature vector binarisation into training.	-Offering remarkable gains in motion estimation compared to similar DL-based feature extraction algorithms, while dramatically lowering inference time. -The matching is substantially accelerated.
Superpoint (161)	2018	-FCN -Homographic adaptation -Multi-scale -Multi-homograph approach.	-Self-supervised interest features. -A deep SLAM frontend. -Designed for real time.
SIVO (163)	2019	-BNN	-Allow for long-term localisation
DBLD-SLAM (170)	2021	-CNN	-Using four important loss functions, extract binary visual feature descriptors from picture patches. -Train Bag of Words to recognise loop closures visually.
ORBDeepOdometry (165)	2019	-CNN	-For modelling the sequential data, the authors propose using an ORB-based feature extractor, CNN-based dimensionality reduction, and stacking several deep LSTMs.
CCNet (166)	2019	-Mask R-CNN+ResNet-101	-Acquiring such contextual information in a more efficient and effective manner.
MonoGRNet (167)	2021	-End-to-End network +Joint geometric loss	-Detecting 3D objects in monocular pictures.

Faster-R-CNN, faster region-based convolutional neural network; DNNs, Deep neural networks ; R-FCN, region-based fully convolutional networks ; CCNet, criss-cross network; CNN, convolutional neural network; DL, deep learning; DNN, deep neural network; DVS, dynamic visual semantic SLAM ; Faster-R-CNN, faster region-based convolutional neural network; RPM, region proposal module; RPN, region proposal network; FCN, fully convolutional network; GANs, generative adversarial networks; LIFT, learned invariant feature transform; LSTM, long-short term memory; ORB, oriented fast and rotated BRIEF; RANSAC, random-sample consensus; SIVO, semantically informed visual odometry; SLAM, simultaneous localisation and mapping; GANVO, generative adversarial networks visual odometry; DL_Hybrid, deep learning Hybrid; GCNv2 ,Geometric Correspondence Network; DBLD-SLAM, deep binary local descriptor; CCNet ,Criss-Cross Network; MonoGRNet, monocular geometric reasoning netwoek, SalientDSO, Salient Direct Sparse Odometry

A NN that is cognizant of geometry is SFM-Net (171), a DL technique that is self-supervised and works with videos to gauge motion. Scenes, object depth, camera motion and 3D object translations and rotations are the categories used to categorise frame-to-frame pixel motion. The program makes predictions about object motion, depth and masks.

The Konda approach (172) was used to predict the direction and velocity changes from visual input using an end-to-end DL architecture. Based on learning rules and a single computational model, the extraction of depth from visual motion and information from odometry are both possible.

DeepVO (173) is a monocular VO that makes use of a cutting-edge end-to-end architecture and a deep recurrent convolutional neural network. The primary goal is to directly predict portions from raw RGB images. To restore the absolute scale, no prereq-uisite information or criteria are required. The RCNN architecture enables the DL-based VO technique to be generalised to entirely new contexts by using the representation of the geometric fea-tures learned via the CNN. Deep recurrent neural networks (DRNNs) are used to automatically learn the complex motion dynamics of image sequences. Tab. 5 summarises optimisation methods by DL in the order they appear in the text.

5.3.2. Relocalisation

When tracking is unsuccessful, the task of relocalisation seeks to increase the accuracy of the camera posture. In this part, we outline some DL-based research projects that try to solve this issue.

VidLoc (174): It is a recurrent model that tries to reduce pose estimate error and accomplish 6-Dof localisation of video. The authors created a spatio-temporal model for global localisation and utilised CNN to predict the scene coordinates. A technique for calculating the instantaneous covariances of position estimations of the input RGB-D picture was implemented into their network.

YOLO (175): It is a method that enhances relocalisation through the use of semantic data. It presents the object 'YOLO' as an array and classifies it using a DL NN using high-level features. This array makes it possible to reject weak candidates and short-en the computation time for the relocalisation tasks.

The research conducted on indoor relocalisation entitled Dual-Stream-CNN (176). It seeks to offer a dual stream CNN-based indoor relocalisation system that takes both colour and depth images as inputs. The suggested technique effectively illustrated the system's robustness in difficult circumstances like large-scale, dynamic, fast-moving and nighttime settings.

Outlier-aware neural tree (177): It is a brand-new outlier-aware neural tree that links decision trees and DL techniques. It uses only stable and secure regions of the surroundings to estab-lish point correspondences for an accurate estimation of camera position. The approach also has decision trees' broad framework characteristics. It is built around three main sections: a hierar-chical space section over the indoor scene to create a decision tree; a deep classification network used to better comprehend the 3D scene and an outlier rejection module used to filter dynamic points during the hierarchical routing process.

SIR-NET (178): The CNN is used by the authors to build a framework for relocalisation. It can be trained end-to-end and is unaffected by the environment. Using the backpropagation of relocalisation faults to both processes enhances retrieval and matching to have the best accuracy in relocation. By selecting pixels based on uncertainty, they can accelerate the unit-matching inference without compromising the accuracy of relocation.

LSTMFCN (long-short term memory fully convolutional net-work) (179): The research was designed to compare two DL-based algorithms to address the issue of single-picture relocalisa-tion. The first uses a DNN end-to-end to directly understand the relationship between an image's position and its mapping. The LSTMFCN algorithm is the second. The LSTMFCN method is distinguished by a much larger receiving range, which avoids the problem of aperture and makes it robust to partial blockages and

moving objects. It is composed of a fully convolutional network (FCN) that performs feature extraction and a long-short term memory (LSTM) that is a pooling layer to group information across the image.

xyzNet (180) is a light CNN. It is a hybrid technique; to relocalise the camera pose from a single RGB image, the researchers merged the geometric method with the machine learning method. The most precise camera position calculation is provided by the geometric information about 2D–3D correspondences, which also eliminates uncertain predictions. Tab. 6 summarises relocalisation methods by DL in the order appearing in the text.

Tab. 5. The optimisation methods of DL SLAM

Name	Year	Architecture/method	Main contribution
Konda Approach (172)	2015	-End-to-End + DL	-Using VO, predict velocity and direction.
Deep_VO (169)	2016	-CNN	-Estimate scale and motion robustly
SFM-Net [194]	2017	-Self-supervised GNN	-A DNN that predicts pixel-wise depth from a single frame as well as camera motion, object motion and object masks from a pair of frames.
VINet (168)	2017	-Sequence-to-sequence + RCNNs	-Offer a unique recurrent network design and training approach to optimise model parameter training
DeepVO (173)	2017	End-to-End + RCNNs	-Presents an RCNN architecture that allows the DL-based VO technique to be generalised to whole new settings by using the CNN's geometric feature representation.

VINet, Visual-Inertial Odometry; DBLD-SLAM , Binary Local Descriptor SLAM; CNN, convolutional neural network; DL, deep learning; DNN, deep neural network; SLAM, simultaneous localisation and mapping; Deep_VO, Deep- visual odometry; SfM-Net: Learning of Structure and Motion.

Tab. 6. The relocalisation methods of DL SLAM

Name	Year	Architecture/method	Main contribution
VidLoc (174)	2017	-CNN	-Attempts to decrease pose estimation error and achieve 6-D of video localisation.
Dual-Stream-CNN (176)	2018	-CNN	-Improves the relocalisation accuracy. -Investigates depth image encoding techniques and proposes a fresh approach termed minimised normal.
LSTMFCN (179)	2018	-FCN	-Avoids the problem of aperture and makes it robust to partial blockages and moving objects. -Suggest refining as a way for improving training model performance.
xyzNet [197]	2018	-Light CNN (xyzNet)	-Geometric information concerning 2D-3D correspondences enables the elimination of unclear predictions and the creation of more precise camera poses. -The accuracy and the performance of our solution on diverse datasets as well as the power to solve difficulties involving dynamic scenario.
SIR-NET (178)	2019	-CNN	-This system simultaneously optimises retrieval and matching tasks to maximise relocalisation accuracy.
Outlier-aware Neural tree (177)	2021	-DL + decision tree approaches.	-Relocalisation in dynamic indoor environments. It achieves robust neural routing through space partitions.
YOLO (175)	2022	-YOLO - Semantic data.	-Rejects unqualified candidates. -Shortens the computation time for the relocalisation tasks.

LSTMFCN, long-short term memory fully convolutional network; SIR-Net : Scene-Independent End-to-End Trainable Visual Relocalize; CNN, convolutional neural network; DL, deep learning; FCN, fully convolutional network; LSTMFCN, long-short term memory fully convolutional network; SLAM, simultaneous localisation and mapping.

5.4. Loop-closure detection

An essential function of the SLAM system is the loop-closure process, which lowers the drift that has collected over time. There are a number of stable, efficient and light-weight DL loop-closure techniques. Traditional feature-point extraction algorithms are used in loop-closing detection methods. The majority of algorithms made use of hand-crafted features and bags of visual words (BoVW).

Wu et al. (181) presented the loop-closure detection for visual SLAM derived from the SuperPoint Network. The SuperPoint NN, which is intended to concurrently recognise points of interest and their associated descriptors, was utilised by the authors to learn inner structures from raw data. The similarity of the image is de-termined using cosine similarity. Merrill and Huang (182) suggested that for the visual close-loop, an unsupervised automatic encoder network architecture is used. The Histogram of Oriented Gradients (HOG) technique provides geometric data and illumination invariance, which forces the encoder to reconstruct the HOG descriptor rather than the original image.

The resulting models do not require labelled training data or environment-specific training; instead, they extract strong to extreme changes in appearance directly from the raw photos.

Utilising the feature obtained through unsupervised DL can increase the loop-closure detection method's accuracy. PCANet, a deep cascade network, was utilised by Yifan Xia, et al. (183) to extract features as image descriptions. Principal component analysis (PCA), binary hashing and block-wise histograms are the three components that make up the PCANet, a straightforward DL

network. The PCA and deep CNN were used by Dai et al. (184) to execute a closed-loop detection process and to scale down the extracted feature dimensions. It is important to note the low detection accuracy of combination approaches. By using the pre-trained ResNet34 model to extract features, this issue is resolved. To reduce the dimension of the features, they then used Kernel PCA (KPCA) on the extraction features.

Seq-CAL, introduced by Xiong et al. (185), is a lightweight sequence-based unsupervised loop-closure-detection approach that integrates sequence information with PCA to achieve good detection accuracy and faster detection times. They reduced descriptor dimensions while retaining sufficient expressive power using PCA. An algorithm for lightweight loop-closure detection and product quantisation (PQ) was created by Huang et al. (186). By using the pre-trained CNN model, SSE-Net, they were able to extract the image's deep visual and semantic features and obtain a vector of feature descriptions. The loop is demonstrated by locating and matching the most comparable pair of candidate frames after PQ and encoding.

Zhu and Huang (187) developed fast and robust visual loop-closure detection using CNN. The authors improved the pre-training model using the Lite-shuffleNet network by extracting the semantic data and depth of the image to derive the feature descriptor, measuring the cosine similarity, choosing the best candidate frames and judging whether to loop.

To represent a picture, Jiayi Ma et al. (188) proposed the fast and robust loop-closure detection through the convolutional auto-encoder and motion consensus. To extract the features, they used a compact convolutional auto-encoder (CAE) network. They trained the network to offer the data of the visual loop closure-detection procedure using the deep perceptual similarity loss function. The principle of place sequence division is the foundation for the phase of loop-closure detection. The CAE network's mapped coding space is employed in the query job to determine which historical image is most similar to the current query image. They introduced an image-to-sequence section method based on place sequence division and distance-weighted voting for loop-closing selection.

In some works, the loop-closure process is addressed using a hybrid DL architecture (HDLA). To enhance spatial awareness and loop-closure detection using a hybrid CNN, Cai et al. (189) devised an effective way to produce high-level semantic image features. It is built using ResNet-18 and optimised with the split–transform–merge concept as well as the squeeze-and-excitation structure, allowing for the compensation of the network's ability to represent pictures without sacrificing performance. To save the time needed to measure the distance between deep semantic features, the authors provided a straightforward method of reducing dimensions during their fitting. Liu et al. (190) proposed a method for developing high-level semantic features that are resistant to changes in both viewpoint and lighting. The architecture of the network is a hybrid ConvNet network tuned to handle robust and real-time feature extraction. Although it shares AlexNet's fundamental structure, it functions best when the appearance is drastically altered. Shi and Li (191) employed a YOLOv4 model with an improved loss function to find the target in the camera-obtained images. The locality sensitive hash function is used to reduce the high-dimensional data dimension, and the cosine distance is used to detect loops.

Local3Ddeep descriptors (L3Ds) (192) is a method for loop detection that measures the overlap. It saves the loop candidate point cloud by their estimated relative positions and then deter-

mines the error metric between points that mutually correspond to the nearest neighbour descriptors. This technique enables precise loop recognition in the event of slight overlaps in 6-DoF estimation.

LoopNet (193) aims to discover important landmarks for the scene to focus on without being distracted by scene fluctuations. It is a plug-and-play algorithm. Additionally, it is a multi-scale attention-based Siamese convolutional model that learns feature embeddings that emphasise the distinguishable objects in the scene rather than comprehensive features.

MAQBOO (194) is a sophisticated algorithm. It increases the effectiveness of pre-trained models to boost visual recall and use them in real-time with multi-agent SLAM systems. In comparison to a high descriptor, the suggested approach achieves equivalent accuracy in a low descriptor dimension. Tab. 7 summarises loop-closing methods by DL in the order appearing in the text.

6. PROBLEMS AND CHALLENGES

We can infer from this study that the visual SLAM system changes over time. Every aspect of its architecture and computer vision tasks confronts issues. The researchers applied environmental perception research to address these problems, enhancing V-SLAM and enhancing resilience in real-world contexts caused by variations in lighting, dynamic objects and shifts in viewpoint. The use of low-level sensors has been found to be another significant SLAM issue. Numerous sources of ambiguity and problems must be solved in order to get a trustworthy SLAM. The three fundamental problems are temporal complexity, uncertainty and correspondence, sometimes known as data association. Classical difficulties and perception problems can be distinguished as issues in the development of visual SLAM. The classical problems result from algorithms whose tasks rely on computer vision-related issues. Among the most common problem:

- Estimation of intrinsic parameters is set before using visual SLAM systems because camera calibration is done before visual SLAM systems and is adjusted during the V-SLAM process.
- Pure rotation is a problem in the field of computer vision due to the inability to observe disparities in the monocular visual SLAM during purely rotational motion. To address this problem, several projection models were used (195)(196).
- The map initialisation presents the first estimation of the localisation and is the main task for the rest of the process of visual SLAM. Among the things that make a preliminary map accurate is to make the baseline wide.
- The scale ambiguity is a particular problem with monocular SLAM. It lies in their geometric inability to get the information of absolute scale about the trajectories and environment.
- Fusion of multi-sensors: The use of a single sensor in the SLAM process generates several limitations. The fusion of multiple sensors can provide rich data resulting in a more accurate and robust system. However, sensor fusion can cause problems on several levels.

Classical visual SLAM needs to address several issues, computation for large-scale environments, distortion of movement and achieving compatibility between accuracy and real-time process relationship.

- Perception problems give rise to algorithms that improve performance in all tasks and seek to implement a robust and precise system that confuses perception with optimization.

DL offers practical precision and robust object detection, and prediction, which can understand the scene this improves the processes of visual SLAM.

– Computation speed: DL offers many advantages in recognition. However, the low computational speed remains one of the most important problems. This makes dynamic SLAM not usable in embedded V-SLAM systems.

– Computation complexity: This problem is generated by incorporating the object-detection modules.

– Future research into SLAM perception will focus on solutions capable of handling real-world conditions and lighting changes and developing and improving tasks for performing visual SLAM in real-time scenes.

Tab. 7. The loop-closing methods of DL SLAM

Name		Year	Architecture/method	Main contribution
Light_unsupervised_D (182)		2018	-Unsupervised deep NN	-Efficient, and robust place recognition. -The visual loop closure that is both reliable and compact.
SuperPoint (181)		2019	-SuperPoint	-Simultaneously identifies interest spots and related descriptions. -By computing the cosine similarity of the respective vectors, it determines how similar the pictures are to one another.
HDLA	(189)	2018	-ResNet+ split-transform-merge strategy + squeeze-and-excitation structure.	-Produces high-level semantic picture features for better loop-closure detection and location recognition. -A simple while fitting dimension reduction algorithm, particularly useful for lowering the time required to estimate distance between deep semantic features.
	(190)	2019	-Hybrid CNN	-Provides high-level semantic picture characteristics specifically for loop closure detection. -By using locality-sensitive hashing (LSH) and employing the nearest neighbour of a single image to search for the key frame using the cosine similarity score, you may guarantee the real-time performance of loop closure detection.
YOLOv4 (191)		2020	-YOLOv4+ optimised loss function	-High-dimensional data can have its dimensions reduced by using the Locality Sensitive Hash function. -The cosine distance is used to determine the loop.
Local3DDeep descriptors (192)		2022	-L3Ds	-After registering the loop candidate point cloud by its estimated relative posture, computes the metric error between points that correspond to mutually-nearest-neighbour descriptors. -In the event of tiny overlaps, properly recognise loops and estimate 6-DoF postures.
LoopNet (193)		2022	-Plug-and-play model+LoopNet,	-Identifies similarities across scenes by identifying essential key landmarks to focus on while being unaffected by scene differences.
MAQBOO (194)		2022	-Multiple AcQuisitions of perceptiBle regiOns for priOr Learning	-Uses spatial information to improve the recall rate in image retrieval on pre- trained models

6-DoF, six degrees of freedom; CNN, convolutional neural network; DL, deep learning; HDLA, hybrid deep learning architecture; L3Ds, local 3D deep descriptors; NN, neural network; SLAM, simultaneous localisation and mapping; BoVW,bags of visual words; HOG,The Histogram of Oriented Gradients; PCA, Principal component analysis; CAE, convolutional auto-encoder; MAQBOOL, Multiple AcQuisitions of perceptiBle regiOns for priOrLearning.

7. CONCLUSION

The most-significant fundamental techniques and problems related to visual SLAM are highlighted in this paper's presentation of the emergence and development phases of the technology. The evolution of SLAM in this study was broken down into three phases: SLAM probabilities, vision SLAM and SLAM perception. Each phase tries to find solutions to the issues raised in the phase before it, while also fostering competency in visual SLAM. The study attempted to demonstrate the benefits, contributions and restrictions of each of the algorithms that were offered.

A lucrative field that also advances visual SLAM is created by the combination of DL methods with machine visions. DL has made numerous advances in recent years, notably for tasks like image analysis, processing and decision-making, which performs with great accuracy and speed. It is possible to use DL to enhance various SLAM tasks, including visual odometry, optimisation, relocalisation and loop closure. DL techniques are utilised in visual SALM to speed up computation and are crucial for fully comprehending the complex scene that is being viewed.

REFERENCES

1. Hans P Moravec. Obstacle Avoidance and Navigation by a Seeing Robot Rover in the Real World. SPittsburgh, Penna Carnegie-Mellon Univ Robot Institute. 1980.
2. D. Nister ON and JB. Visual odometry. Proc 2004 IEEE Comput Soc Conf Comput Vis Pattern Recognition,004 CVPR 2004. Washington DC. USA. 2004;1:I–I.
3. Longuet-Higgins H. A computer algorithm for reconstructing a scene from two projections. Nature. 1981;293:133–5.
4. CG Harris JMP. 3d positional integration from image sequences. Image Vis Comput Sci Direct. 1988;6(2):87–90.
5. Twinanda AP, Shehata S, Mutter D, Marescaux J, De Mathelin M and NP. Endonet: a deep architecture for recognition tasks on laparoscopic videos. IEEE Trans Med Imaging. 2016;36(1):86–97.
6. Bodenstedt S, Ohnemus A, Katic D, Wekerle AL, Wagner M, Kenngott H, Muller-Stich B, Dillmann R and SS. Real-time image-based instrument clas- sification for laparoscopic surgery. 2018. preprint arXiv:1808.00178.
7. Yang IC, Chen S. Precision cultivation system for greenhouse production. In Intelligent Environmental Sensing. Springer Berlin/Heidelberg. Ger Google Sch. 2015;191–211.

8. Borges DL, Guedes ST, Nascimento AR, Melo-Pinto P. Detecting and grading severity of bacterial spot caused by Xanthomonas spp. in tomato (Solanum lycopersicon) fields using visible spectrum images. Comput Electron Agric. 2016;149–159.

9. Liu X, Zhao D, Jia W, Ji W, Ruan C, Sun Y. Cucumber fruits detection in greenhouses based on instance segmentation. IEEE Access. 2019;139635–139642.

10. Asdemir S, Urkmez A, Inal S. Determination of body measurements on the Holstein cows using digital image analysis and estimation of live weight with regression analysis. Comput Electron Agric. 2011;76, 189–197.

11. Norton T, Chen C, Larsen MLV, Berckmans D. Precision livestock farming: Building 'digital representations' to bring the animals closer to the farmer. Anim 1. 2019;3:3009–3017.

12. Chou WC, Tsai WR, Chang HH, Lu SY, Lin KF, Lin P. Prioritization of pesticides in crops with a semi-quantitative risk ranking method for Taiwan postmarket monitoring program. J Food Drug Anal. 2019;27: 347–354.

13. Kalman RE. A new approach to linear filtering and prediction problems. Trans ASME. J Basic Eng. 1960;82(1):35–45.

14. Julier SJ, Uhlmann JK. A counter example to the theory of simultaneous localization and map building. Proc 2001 ICRA IEEE Int Conf Robot Autom (Cat No01CH37164). Seoul. Korea (South). 2001; 4: 4238-4243. doi:101109/ROBOT2001933280

15. Gamini Dissanayake MWM, Newman P, Clark S, Durrant-Whyte HF, Csorba M. A solution to the simultaneous localization and map building (SLAM) problem. IEEE Trans Robot Autom. 2001;17(3):229–41.

16. Smith R, Self M, Cheeseman P. A stochastic map for uncertain spatial relationships. Mach Intell Pattern Recognit [Internet]. 1988;5:435–61.
Available from: http://portal.acm.org/citation.cfm?id=57472

17. Moutarlier P, Chatila R. An Experimental System for Incremental Environment Modelling by an Autonomous Mobile Robot. LAAS-CNRS 7. Ave du Colonel Roche 31077 Toulouse.

18. Jazwinski AH. Stochastic Processes and Filtering Theory. 1970;64.

19. Zhu J, Zheng N, Yuan Z, Zhang QXZ and YH. A SLAM algorithm based on the central difference kaiman filter. IEEE Intell Veh Symp Xi'an. China. 2009;123–8.

20. Jiang X, Li T, Yu Y. A novel SLAM algorithm with Adaptive Kalman filter. ICARM 2016 Int Conf Adv Robot Mechatronics. 2016; 107–11.

21. Tian Y, Suwoyo H, Wang W, Mbemba D, Li L. An AEKF-SLAM Algorithm with Recursive Noise Statistic Based on MLE and EM. J Intell Robot Syst. 2020;97:339–55.

22. Julier SJ, Uhlmann JK. New extension of the Kalman filter to nonlinear systems. Proc Vol 3068, Signal Process Sens Fusion, Target Recognit VI. 1997;3068.

23. Havangi R. Robust SLAM: SLAM base on H ∞ square root unscented Kalman filter. Nonlinear Dyn. 2016;83(1):767–79.

24. Bahraini M, Bozorg M, Rad A. A new adaptive UKF algorithm to improve the accuracy of SLAM. Int J Robot. 2019;5(1):35–46.

25. Bahraini MS. On the Efficiency of SLAM Using Adaptive Unscented Kalman Filter. Iran J Sci Technol Trans Mech Eng [Internet]. 2020;44:727–35.
Available from: https://doi.org/10.1007/s40997-019-00294-z

26. Tang M, Chen Z, Yin F. SLAM with Improved Schmidt Orthogonal Unscented Kalman Filter. Int J Control Autom Syst. 2022;20(1598–6446):1327–35.

27. Liu D, Duan J and HS. A Strong Tracking Square Root Central Difference FastSLAM for Unmanned Intelligent Vehicle With Adaptive Partial Systematic Resampling. EEE Trans Intell Transp Syst. 2016;17(11):3110–20.

28. Maybeck PS. Stochastic Models, Estimation, and Control. Acad Press. 1979;1:282.

29. Garritsen T. Using the Extended Information Filter for Localization of Humanoid Robots on a Soccer Field. 2018;1–25.

30. Thrun S, Liu Y, Koller D, Ng AY, Ghahramani Z, Durrant-Whyte H. Simultaneous localization and mapping with sparse extended information filters. Int J Rob Res. 2004;23(7–8):693–716.

31. Walter MR, Eustice RM, Leonard JJ. Exactly sparse extended information filters for feature-based SLAM. Int J Rob Res. 2007;26(4):335–59.

32. He B, Liu Y, Dong D, Shen Y, Yan T, Nian R. Simultaneous localization and mapping with iterative sparse extended information filter for autonomous vehicles. Sensors (Switzerland). 2015;15(2): 19852–79.

33. Zhang H, Liu Y, Tan J, Xiong N. RGB-D SLAM Combining Visual Odometry and Extended Information Filter. Sensors [Internet]. 2015;15:18742–66. Available from: www.mdpi.com/journal/sensors

34. Ila V, Porta JM, Andrade-Cetto J. Information-based compact pose SLAM. IEEE Trans Robot. 2010;26(1):78–93.

35. Del Moral P. Nonlinear filtering: Interacting particle resolution. Comptes Rendus l'Académie des Sci - Ser I - Math. 1996;2(4):555–80.

36. Gordon NJ, Salmond DJ, Smith AFM. Novel approach to nonlinear/non-Gaussian Bayesian state estimation. IEE. 1993;140(2): 107–13.

37. Liu JS, Rong C. Sequential Monte Carlo methods for dynamic systems. J Am Stat Assoc. 1998;93(443):1032–1044.

38. Øivind Skare EB and LH. Improved Sampling-Importance Resampling and Reduced Bias Importance Sampling. Scand J Stat. 2003;30(4):719-737.

39. Bruno MGS. Regularized Particle Filters. Seq Monte Carlo Methods Nonlinear Discret Filtering Synth Lect Signal Process Springer. 2013.

40. Blackwell D. Conditional Expectation and Unbiased Sequential Estimation. Ann Math Stat. 1947;18(1):105–10.

41. Doucet A, Murphy K, Berkeley UC. Rao-Blackwellised Particle Filtering for Dynamic Bayesian Networks. 1999.

42. Murphy K SR. Rao-Blackwellised Particle Filtring for Dynamic Bayesian Networks. Springer New York. 2001;43(2):499–515.

43. Montemerlo M, Thrun S, Koller D, Wegbreit B. FastSLAM: A Factored Solution to the Simultaneous Localization and Mapping Problem. Eighteenth Natl Conf Artif Intell Menlo Park. 2002;593–598.

44. Montemerlo M, Thrun S, Siciliano B. FastSLAM:A Scalable Method for the Simultaneous Localization and Mapping Problem in Robotics. Springer. 2007;27.

45. Michael M, Thrun S, Koller D, Wegbreit B. FastSLAM 2.0: An Improved Particle Filtering Algorithm for Simultaneous Localization and Mapping that Provably Converges. IJCAI'03 Proc 18th Int Jt Conf Artif Intell. 2003;1151–6.

46. Kim C, Sakthivel R, Chung WK. Unscented FastSLAM : A Robust Algorithm for the Simultaneous Localization and Mapping Problem. 2008.

47. Eliazar A, Parr R. DP-SLAM: Fast, robust simultaneous localization and mapping without predetermined landmarks. IJCAI Int Jt Conf Artif Intell. 2003;1135–42.

48. Eliazar AI, Parr R. DP-SLAM 2.0. Dep Comput Sci Duke Univ North Carolina 27708.

49. Zikos N, Petridis V. 6-DoF Low Dimensionality SLAM (L-SLAM). J Intell Robot Syst. 2015;79:55–72.

50. Nie F, Zhang W, Yao Z, Shi Y, Li F, Huang Q. LCPF: A Particle Filter Lidar SLAM System with Loop Detection and Correction. IEEE Access. 2020;8:20401–12.

51. Hua J, Cheng M. Improved UFastSLAM algorithm based on particle filter. IEEE 9th Jt Int Inf Technol Artif Intell Conf. 2020;(2693–2865):1050–5.

52. Lin M, Member S, Canjun Yang, Li D. An Improved Transformed Unscented FastSLAM with Genetic Resampling. IEEE Trans Ind Electron. 2019;66(5):3583–94.

53. Tang M, Chen Z, Yin F. An Improved Adaptive Unscented FastSLAM with Genetic Resampling. Int J Control Autom Syst. 2021;19(4):1677–90.

54. Lu F, Milios E. Globally Consistent Range Scan Alignment for Environment Mapping. Auton Robots. 1997;4(4):333–49.

55. Thrun S. The GraphSLAM Algorithm with Applications to Large-Scale Mapping of Urban Structures. Int J Rob Res. 1998;25:403–29.

56. Grisetti G, Stachniss C, Grzonka S, Burgard W. A tree parameterization for efficiently computing maximum likelihood maps using gradient descent. Robot Sci Syst. 2008;3:65–72.

57. Frese U. Treemap: An O(log n) algorithm for indoor simultaneous localization and mapping. Auton Robots. 2006;103–22.

58. Grisetti G, Kümmerle R, Stachniss C, Frese U, Hertzberg C. Hierarchical optimization on manifolds for online 2D and 3D mapping. Proc - IEEE Int Conf Robot Autom. 2010;273–8.

59. Kaess M, Johannsson H, Roberts R, Ila V, Leonard JJ, Dellaert F. ISAM2: Incremental smoothing and mapping using the Bayes tree. Int J Rob Res. 2012;31(2):216–35.

60. Rainer K, Grisetti G, Hauke S, Kurt. K, Abstract—Many WB. g2o: A General Framework for Graph Optimization Rainer. IEEE Int Conf Robot Autom Shanghai Int Conf Cent. 2011;3607–13.

61. Dellaert F. Factor Graphs and GTSAM. A hands-on Introd Tech Rep (Georgia Tech, Atlanta 2012) [Internet]:1–27. Available from: http://tinyurl.com/gtsam.

62. Agarwal P, Tipaldi GD, Spinello L, Stachniss C, Burgard W. Robust map optimization using dynamic covariance scaling. Proc - IEEE Int Conf Robot Autom. 2013.

63. Strasdat H, Davison AJ, Montiel JMM, Konolige K. Double window optimisation for constant time visual SLAM. Int Conf Comput Vis. 2011.

64. M. Ruhnke R. Kümmerle G, Grisetti WB. Highly accurate 3D surface models by sparse surface adjustment. IEEE Int Conf Robot Autom. 2012;(10.1109/ICRA.2012.6225077).

65. Stachniss C, Leonard JJ, Thrun S. Simultaneous Localization and Mapping. In: Multimedia Contents 1153 springer Handbook Robotics Part E/46. 2016;1153–75.

66. Zhao L, Huang S, Dissanayake G. Linear SLAM: Linearising the SLAM problems using submap joining. Automatica. 2018;1–22.

67. Holder M, Hellwig S, Winner H. Real-time pose graph SLAM based on radar. IEEE Intell Veh Symp. 2019.

68. Youyang F, Qing W, Gaochao Y. Incremental 3-D pose graph optimization for SLAM algorithm without marginalization. Int J Adv Robot Syst. 2020;1–14.

69. Fan T, Wang H, Rubenstein M, Murphey T. Cpl-slam: Efficient and certifiably correct planar graph-based slam using the complex number representation. IEEE Trans Robot. 2020;36(6):1719–37.

70. Sun Z, Wu B, Xu CZ, Sarma SE, Yang J, Kong H. Frontier Detection and Reachability Analysis for Efficient 2D Graph-SLAM Based Active Exploration. IEEE/RSJ Int Conf Intell Robot Syst. 2020;2051–8.

71. Pierzchała M, Giguère P, Astrup R. Mapping forests using an unmanned ground vehicle with 3D LiDAR and graph-SLAM. Comput Electron Agric. 2018;145:217–25.

72. Press W, Keukolsky S WV and BF. Levenberg Marquardt Method. Numer Recipes C Art Sci Comput. 1992;542–54.

73. Shum HY, Ke Q and ZZ. Efficient Bundle Adjustment with Virtual Key Frames: A Hierarchical Approach to Multi-frame Structure from Motion. IEEE Comput Soc Conf Comput Vis Pattern Recognition. 1999.

74. Hartley R, Zisserman A. Multiple View Geometry in Computer Vision. Cambridge Univ Press. 2000;18.

75. Melbouci K, Collette SN, Gay-Bellile V, Ait-Aider O, Carrier M, Dhome M. Bundle adjustment revisited for SLAM with RGBD sensors. Proc 14th IAPR Int Conf Mach Vis Appl MVA. 2015;166–9.

76. Frost D, Prisacariu V, Murray D. Recovering Stable Scale in Monocular SLAM Using Object-Supplemented Bundle Adjustment. IEEE Trans Robot. 2018;34(3):1–11.

77. Schops T, Sattler T, Pollefeys M. Bad slam: Bundle adjusted direct RGB-D slam. IEEE/CVF Conf Comput Vis Pattern Recognit. 2019;134–44.

78. Zhao Y, Smith JS, Vela PA. Good Graph to Optimize: Cost-Effective, Budget-Aware Bundle Adjustment in Visual SLAM. Comput Vis Pattern Recognit [Internet]. 2020;1–20. Available from: http://arxiv.org/abs/2008.10123

79. Wang K, Ma S, Ren F, Lu J. SBAS: Salient Bundle Adjustment for Visual SLAM. J LATEX Cl FILES.arxiv201211863v1[csRO]. 2015;14(8):1–11.

80. Campos C, Elvira R, Rodriguez JJG, Montiel JMM, Tardos JD. ORB-SLAM3: An Accurate Open-Source Library for Visual. Visual-Inertial and Multimap SLAM. IEEE Trans Robot. 2021;37(6):1874–90.

81. Gonzalez M, Marchand E, Kacete A, Royan J. S3LAM: Structured Scene SLAM. Robotics [Internet]. 2022. Available from: http://arxiv.org/abs/2109.07339

82. Tanaka T, Sasagawa Y, Okatani T. Learning to Bundle-adjust: A Graph Network Approach to Faster Optimization of Bundle Adjustment for Vehicular SLAM. Proc IEEE Int Conf Comput Vis. 2021;6230–9.

83. Rosten E, Drummond T. Machine Learning for High-Speed Corner Detection. Leonardis A, Bischof H, Pinz A Comput Vis – ECCV 2006ECCV 2006 Lect Notes Comput Sci Springer. Berlin. Heidelb. 2006;3951:430–43.

84. Bay H, Ess A, Tuytelaars T, Gool L Van. Speeded-Up Robust Features (SURF). Comput Vis Image Underst. 2008;110(3):346–59.

85. Calonder M, Lepetit V, Strecha C, Fua P. BRIEF: Binary robust independent elementary features. ECCV 2010 Lect Notes Comput Sci Springer. Berlin. Heidelberg. 2010;6314:778–92.

86. E. Rublee, V. Rabaud KK and GB. ORB: an efficient alternative to SIFT or SURF. Int Conf Comput Vision. Barcelona. Spain. 2011;2564–71.

87. Harris C, Stephens M. A Combined Corner and Edge Detector. Proc 4th Alvey Vis Conf. 1988;147--151.

88. Civera J, Lee SH. RGB-D Odometry and SLAM. Rosin, P, Lai, YK, Shao, L, Liu, Y RGB-D Image Anal Process Adv Comput Vis Pattern Recognition Springer. Cham. 2019;117–144.

89. Davison AJ, Reid ID NDM, Stasse O. Monoslam: real-time single camera SLAM. Pattern Anal Mach Intell IEEE. 2007;29(6):1052–67.

90. Davison AJ. Real-time simultaneous localisation and mapping with a single camera. Proc Ninth IEEE Int Conf Comput Vision. Nice. Fr. 2003;2:1403–10.

91. Klein G, Murray D. Parallel tracking and mapping for small AR workspaces. 2007 6th IEEE ACM Int Symp Mix Augment Reality. ISMAR. 2007;225–34.

92. Klein G, Murray D. Parallel tracking and mapping on a camera phone. th IEEE Int Symp Mix Augment Reality. Orlando FL. USA. 2009. 2009;83–6.

93. Endres F, Hess J, Engelhard N, Sturm J DC and WB. An evaluation of the RGB-D SLAM system. IEEE Int Conf Robot Autom Saint Paul. MN. USA. 2012;3(c):1691–6.

94. Mur-Artal R, Montiel JMM, Tardos JD. ORB-SLAM: A Versatile and Accurate Monocular SLAM System. IEEE Trans Robot. 2015;31(5):1147–63.

95. Tardos DG-L and JD. Bags of Binary Words for Fast Place Recognition in Image Sequences. IEEE Trans Robot. 28(5):1188–97.

96. Strasdat H, Davison AJ, Montiel. JMM. Scale Drift-Aware Large Scale Monocular SLAM. Robot Sci Syst. 2010.

97. Mei C, Sibley G, Newman P. Closing loops without places. IEEE/RSJ 2010 Int Conf Intell Robot Syst IROS 2010 - Conf Proc. 2010;3738–44.

98. Mur-Artal R, Tardós JD. ORB-SLAM: Tracking and Mapping Recognizable Features. Conf Work Multi VIew Geom Robot - RSS 2014 [Internet]. 2014. Available from: http://vindelman.technion.ac.il/events/mvigro/MurArtal14rss_ws.pdf

99. Mur-Artal R, Tardos JD. ORB-SLAM2: An Open-Source SLAM System for Monocular, Stereo and RGB-D Cameras. IEEE Trans Robot. 2017;33(5):1255–62.

100. Sumikura S, ShibuyaM KS. OpenVSLAM: A versatile visual SLAM framework. MM '19 Proc 27th ACM Int Conf Multimedia. 2019;2292–5.

101. Muñoz-Salinas R, Medina-Carnicer R. UcoSLAM: Simultaneous localization and mapping by fusion of keypoints and squared planar markers. Comput Vis Pattern Recognit. 2019;

102. Sun Q, Yuan J, Zhang X, Duan F. Plane-Edge-SLAM: Seamless Fusion of Planes and Edges for SLAM in Indoor Environments. IEEE Trans Autom Sci Eng. 2021;18(4):2061–75.

103. Newcombe RA, Lovegrove SJ, Davison AJ. DTAM: Dense Tracking and Mapping in Real-Time. Int Conf Comput Vision, Barcelona, Spain. 2011;2320–7.

104. J Engel JS and DC. Semi-Dense Visual Odometry for a Monocular Camera. IEEE Int Conf Comput Vision. Sydney. NSW. Aust. 2013;1449–56.

105. Engel J, Sturm J, Cremers D. LSD-SLAM: Large-Scale Direct Monocular SLAM. Proc IEEE Int Conf Comput Vis. 2013;1449–56.

106. Engel J, Stuckler J DC. Large-scale direct SLAM with Stereo Cameras. IEEE/RSJ Int Conf Intell Robot Syst (IROS). Hamburg Ger. 2015;1935–42.

107. Engel J, Cremers, Daniel, Caruso D. Large-scale direct SLAM for omnidirectional cameras. IEEE/RSJ Int Conf Intell Robot Syst (IROS). Hamburg Ger. 2015;141–8.

108. Forster C, Pizzoli M, Scaramuzza D. SVO : Fast Semi-Direct Monocular Visual Odometry. IEEE Int Conf Robot Autom (ICRA)Hong Kong. China. 2014;15–22.

109. Engel J, Koltun V, Cremers D. Direct Sparse Odometry. IEEE Trans Pattern Anal Mach Intell. 2018;40(3):611–25.

110. Gao X, Wang R, Demmel N, Cremers D. LDSO: Direct Sparse Odometry with Loop Closure. IEEE Int Conf Intell Robot Syst Spain. 2018;2198–204.

111. Sheng C, Pan S, Gao W, Tan Y, Zhao T. Dynamic-DSO: Direct sparse odometry using objects semantic information for dynamic environments. Appl Sci. 2020;10(4):1–20.

112. Newcombe RA, Izadi S, Hilliges O, Molyneaux D, Kim D, Davison AJ et al. KinectFusion: Real-time dense surface mapping and tracking. 210th IEEE Int Symp Mix Augment Reality. Basel. Switzerland. 2011;127–36.

113. Concha A, Civera J. RGBDTAM: A cost-effective and accurate RGB-D tracking and mapping system. IEEE Int Conf Intell Robot Syst Concha J Civera. RGBDTAM A cost-effective accurate RGB-D Track Mapp Syst 2017 IEEE/RSJ Int Conf Intell Robot Syst (IROS). Vancouver. 2017;6756–63.

114. Fontán A JC and RT. Information-Driven Direct RGB-D Odometry. IEEE/CVF Conf Comput Vis Pattern Recognit (CVPR). Seattle. WA. USA. 2020;4928–36.

115. Ma L, Kerl C, Stückler J, Cremers D. CPA-SLAM: Consistent Plane-Model Alignment for Direct RGB-D SLAM. IEEE Int Conf Robot Autom (ICRA). Stock Sweden [Internet]. 2016;1:1285–91. Available from: https://pdfs.semanticscholar.org/d41a/4ab403d6c7611047f83f575c f4c16bfd5282.pdf

116. Dai A, Nießner M, Zolloer M, Izadi S and C. BundleFusion: Real-time Globally Consistent 3D Reconstruction using On-the-fly Surface Re-integration. IEEE Int Conf Progr Compr. 2022;1(1):19.

117. Hsiao M, Westman E, Zhang G, Kaess M. Keyframe-based dense planar SLAM. IEEE Int Conf Robot Autom (ICRA). Singapore. 2017;5110–7.

118. Dong X, Cheng L, Peng H, Li T. FSD-SLAM: a fast semi-direct SLAM algorithm. Complex Intell Syst [Internet]. 2022;8:1823–34. Available from: https://doi.org/10.1007/s40747-021-00323-y

119. Bloesch M, Omari S, Hutter M, Siegwart R. Robust visual inertial odometry using a direct EKF-based approach. IEEE/RSJ Int Conf Intell Robot Syst (IROS). Hamburg Ger. 2015;298–304.

120. Sun K, Mohta K, Pfrommer B, Watterson M, Liu S, Mulgaonkar Y, et al. Robust Stereo Visual Inertial Odometry for Fast Autonomous Flight. IEEE Robot Autom Lett. 2018;3(2):965–72.

121. Mourikis AI, Roumeliotis SI. A multi-state constraint Kalman filter for vision-aided inertial navigation. Proc 2007 IEEE Int Conf Robot Autom Rome. Italy. 2007;3565–72.

122. Leutenegger S, Lynen S, Bosse M, Siegwart R, Furgale P. Keyframe-Based Visual-Inertial Odometry Using Nonlinear Optimization. Int J Rob Res. 2014;34(3):1–26.

123. Schneider T, Dymczyk M, Fehr M, Egger K, Lynen S, Gilitschenski I et al. Maplab: An Open Framework for Research in Visual-Inertial Mapping and Localization. IEEE Robot Autom Lett. 2018;3(3):1418–25.

124. Liu H, Chen M, Zhang G, Bao H, Bao Y. ICE-BA: Incremental, Consistent and Efficient Bundle Adjustment for Visual-Inertial SLAM. IEEE/CVF Conf Comput Vis Pattern Recognition. Salt Lake City. UT USA. 2018;1974–82.

125. Forster C, Carlone L, Dellaert F, Scaramuzza D. On-Manifold Preintegration for Real-Time Visual-Inertial Odometry. Iin IEEE Trans Robot. 2017;33(1):1–20.

126. Von Stumberg L, Usenko V, Cremers D. Direct Sparse Visual-Inertial Odometry Using Dynamic Marginalization. IEEE Int Conf Robot Autom (ICRA). Brisbane QLD. Aust. 2018;2510–7.

127. Qin T, Li P, Shen S. VINS-Mono: A Robust and Versatile Monocular Visual-Inertial State Estimator. IEEE Trans Robot. 2018;34(4):1004–20.

128. Yang Z, Shen S. Monocular visual-inertial state estimation with online initialization and camera-IMU extrinsic calibration. IEEE Trans Autom Sci Eng. 2017;14(1):39–51.

129. Mur-Artal R, Tardos JD. Visual-Inertial Monocular SLAM with Map Reuse. IEEE Robot Autom Lett. 2017;2(2):796–803.

130. He Y, Zhao J, Guo Y, He W, Yuan K. PL-VIO: Tightly-coupled monocular visual–inertial odometry using point and line features. Sensors (Switzerland). 2018;18(4):1–25.

131. Zheng F, Tsai G, Zhang Z, Liu S, Chu CC, Hu H. Trifo-VIO: Robust and Efficient Stereo Visual Inertial Odometry Using Points and Lines. IEEE/RSJ Int Conf Intell Robot Syst (IROS). Madrid Spain. 2018;3686–93.

132. Li X, Li Y, Ornek EP, Lin J, Tombari F. Co-Planar Parametrization for Stereo-SLAM and Visual-Inertial Odometry. IEEE Robot Autom Lett. 2020;5(4):6972–9.

133. Rosinol A, Sattler T, Pollefeys M, Carlone L. Incremental visual-inertial 3d mesh generation with structural regularities. Int Conf Robot Autom (ICRA). Montr QC. Canada. 2019;8220–6.

134. Seiskari O, Rantalankila P, Kannala J, Ylilammi J, Rahtu E, Solin A. HybVIO: Pushing the Limits of Real-time Visual-inertial Odometry. IEEE/CVF Winter Conf Appl Comput Vis (WACV). Waikoloa HI, USA. 2022;287-296.

135. Kaushik V, Jindgar K, Lall B. ADAADepth: Adapting data augmentation and attention for self-supervised monocular depth estimation. IEEE Robot Autom Lett. 2021;6(4):7791–8.

136. Tateno K, Tombari F, Laina I NN. CNN-SLAM: Real-time dense monocular SLAM with learned depth prediction. Comput Vis Pattern Recognit. 2017;6243–52.

137. Bloesch M, Czarnowski J, Clark R, Leutenegger S AJD. CodeSLAM-Learning a Compact. Optimisable Representation for Dense Visual SLAM. 2018;2560–8. Available from: http://openaccess.thecvf.com/content_cvpr_2018/papers/Bloesch_ CodeSLAM_--_Learning_CVPR_2018_paper.pdf

138. Mohanty V, Agrawal S, Datta S, Ghosh A, Vishnu Dutt Sharma DC. DeepVO: A Deep Learning approach for Monocular Visual Odometry. 2016. Available from: http://arxiv.org/abs/1611.06069

139. Li R, Wang S, Long Z, Gu D. UnDeepVO: Monocular Visual Odometry Through Unsupervised Deep Learning. IEEE Int Conf Robot Autom (ICRA). Brisbane QLD. Aust. 2018;7286–91.

140. Yang N, Von Stumberg L, Wang R, Cremers D. D3VO: Deep Depth, Deep Pose and Deep Uncertainty for Monocular Visual Odometry. Proc IEEE Comput Soc Conf Comput Vis Pattern Recognit. 2020;1278–89.

141. Yin Z, Shi J. GeoNet: Unsupervised Learning of Dense Depth, Optical Flow and Camera Pose http: 2018;1983–92. Available from: geonet: Unsupervised Learning of Dense Depth, Optical Flow and Camera Pose http://arxiv.org/abs/1803.02276v2

142. Zhao C, Sun L, Purkait P, Duckett T, Stolkin R. Learning Monocular Visual Odometry with Dense 3D Mapping from Dense 3D Flow. IEEE/RSJ Int Conf Intell Robot Syst (IROS). Madrid Spain. 2018;6864–71.

143. Zhou T, Brown M, Snavely N DGL. Unsupervised Learning of Depth and Ego-Motion from Video. CEEE Conf Comput Vis Pattern Recognit (CVPR), Honolulu, HI, USA [Internet]. 2017;6612–9. Available from: https: //github.com/tinghuiz/SfMLearner.%0A2

144. Zagoruyko S, Komodakis N. Learning to Compare Image Patches via Convolutional Neural Networks Sergey. IEEE Conf Comput Vis Pattern Recognit (CVPR). Boston MA. USA. 2015;4353–61.

145. G VKB, Carneiro G, Reid I. Learning Local Image Descriptors with Deep Siamese and Triplet Convolutional Networks by Minimizing Global Loss Functions. IEEE Conf Comput Vis Pattern Recognit (CVPR)IEEE Conf Comput Vis Pattern Recognit (CVPR). Las Vegas NV. USA [Internet]. 2016;5385–94. Available from: http://openaccess.thecvf.com/content_cvpr_2016/supplemental/G_Learning_Local_Image_2016_CVPR_supplemental.pdf

146. Mayer N, Ilg E, Hausser P, Fischer P. A Large Dataset to Train Convolutional Networks for Disparity, Optical Flow, and Scene Flow Estimation. IEEE Conf Comput Vis Pattern Recognit (CVPR). Las Vegas NV. USA. 2016;4040–8.

147. Tankovich V, Häne C, Zhang Y, Kowdle A, Fanello S, Bouaziz S. HitNet: Hierarchical Iterative Tile Refinement Network for Real-time Stereo Matching. IEEE/CVF Conf Comput Vis Pattern Recognit (CVPR). Nashville TN. USA. 2021;14357–67.

148. Huang PH, Matzen K, Kopf J, Ahuja N, Huang J Bin. DeepMVS: Learning Multi-view Stereopsis. IEEE/CVF Conf Comput Vis Pattern Recognition. Salt Lake City UT. USA. 2018;2821–30.

149. Song X, Zhao X, Hu H, Fang L. EdgeStereo: A Context Integrated Residual Pyramid Network for Stereo Matching. Comput Vis – ACCV 2018 ACCV 2018 Lect Notes Comput Sci. 2018;arXiv:1803.05196.

150. Shao C, Zhang C, Fang Z, Yang G. A Deep Learning-Based Semantic Filter for RANSAC-Based Fundamental Matrix Calculation and the ORB-SLAM System. IEEE Access. 2020;8:3212–23.

151. Zhang W, Liu G, Tian G. A Coarse to Fine Indoor Visual Localization Method Using Environmental Semantic Information. IEEE Access. 2019;7:21963–70.

152. Lin YF, Yang LJ, Yu CY, Peng CC, Huang DC. Object recognition and classification of 2D-SLAM using machine learning and deep learning techniques. Int Symp Comput Consum Control (IS3C). Taichung City. Taiwan. 2020;473–6.

153. Wang S, Clark R, Wen H, Trigoni N. End-to-End, Sequence-to-Sequence Probabilistic Visual Odometry through Deep Neural Networks. Int J Robot Res 37 [Internet]. 2018;37:513–42. Available from: doi.org/10.1177/0278364917734298

154. Li J, Li Z, Feng Y, Liu Y, Shi G. Development of a Human-Robot Hybrid Intelligent System Based on Brain Teleoperation and Deep Learning SLAM. IEEE Trans Autom Sci Eng. 2019;16(4):1664–74.

155. Lan E. A Novel Deep Learning Architecture By Integrating Visual Simultaneous Localization And Mapping (Vslam) Into Cnn For Real-Time Surgical Video Analysis. 19th Int Symp Biomed Imaging (ISBI). Kolkata. India. 2022;1–5.

156. Hu S, Li D, Tang G, Xu X. A 3D semantic visual SLAM in dynamic scenes. 6th IEEE Int Conf Adv Robot Mechatronics (ICARM). Chongqing. China. 2021;522–8.

157. Almalioglu Y, Saputra MRU, De Gusmao PPB, Markham A, Trigoni N. GANVO: Unsupervised deep monocular visual odometry and depth estimation with generative adversarial networks. Proc - IEEE Int Conf Robot Autom Conf Robot Autom (ICRA), Montr QC. Canada. 2019;5474–80.

158. Ban X, Wang H, Chen T, Wang Y, Xiao Y. Monocular Visual Odometry based on depth and optical flow Using deep learning. IEEE Trans Instrum Meas. 2021;70:1–19.

159. Liang HJ, Sanket NJ, Fermuller C, Aloimonos Y. SalientDSO: Bringing Attention to Direct Sparse Odometry. IEEE Trans Autom Sci Eng. 2019;16(4):1619–26.

160. Tang J, Ericson L, Folkesson J, Jensfelt P. GCNv2: Efficient Correspondence Prediction for Real-Time SLAM. IEEE Robot Autom Lett. 2019;4(4):3505–10.

161. Detone D, Malisiewicz T, Rabinovich A. SuperPoint: Self-supervised interest point detection and description. EEE/CVF Conf Comput Vis Pattern Recognit Work (CVPRW). Salt Lake City UT. USA. 2018;337–49.

162. Kwang Moo Yi, Eduard Trulls, Vincent Lepetit PF. LIFT: Learned Invariant Feature Transform Kwang. Springer Int Publ AG 2016.

163. Ganti P, Waslander S. Network uncertainty informed semantic feature selection for visual SLAM. 16th Conf Comput Robot Vis (CRV) Kingston QC. Canada. 2019;121–8.

164. Gu X, Wang Y, Ma T. DBLD-SLAM: A Deep-Learning Visual SLAM System Based on Deep Binary Local Descriptor. Int Conf Control Autom Inf Sci (ICCAIS). Xi'an China. 2021;325–30.

165. Krishnan KS, Sahin F. ORBDeepOdometry - A feature-based deep learning approach to monocular visual odometry. 14th Annu Conf Syst Syst Eng (SoSE). Anchorage AK. USA. 2019;296–301.

166. Huang Z, Wang X, Huang L, Huang C, Wei Y, Liu W. CCNet: Criss-cross attention for semantic segmentation. IEEE Trans Pattern Anal Mach Intell. 2019;603–12.

167. Qin Z, Wang J, Lu Y. MonoGRNet: A General Framework for Monocular 3D Object Detection. IEEE Trans Pattern Anal Mach Intell. 2021;44(9):5170–84.

168. Ronald Clark, Sen Wang, Hongkai Wen, Andrew Markham NT. VINet: Visual-Inertial Odometry as a Sequence-to-Sequence Learning Problem. Proc Thirty-First AAAI Conf Artif Intell. 2017;31(1):3995–4001.

169. G. Costante, M. Mancini PV and TAC. "Exploring Representation Learning With CNNs for Frame-to-Frame Ego-Motion Estimation,. IEEE Robot Autom Lett. 2016;1:18-25.

170. Gu X. DBLD-SLAM : A Deep-Learning Visual SLAM System Based on Deep Binary Local Descriptor. 2021;325–30.

171. Vijayanarasimhan S, Ricco S, Schmid C. SfM-Net: Learning of Structure and Motion from Video. 2017. arXiv preprint arXiv:1704.07804.

172. Konda K, Memisevic R. Learning visual odometry with a convolutional network. Proc ofthe 10th Int Conf Comput Vis Theory Appl. 2015;1:486–90.

173. Wang S, Clark R, Wen H, Trigoni N. DeepVO: Towards end-to-end visual odometry with deep Recurrent Convolutional Neural Networks. IEEE Int Conf Robot Autom (ICRA). Singapore. 2017;2043–50.

174. Clark R, Wang S, Markham A, Trigoni N, Wen H. VidLoc : A Deep Spatio-Temporal Model for 6-DoF Video-Clip Relocalization. IEEE Conf Comput Vis Pattern Recognit (CVPR). Honolulu HI. USA. 2017;2652–60.

175. Mahattansin N, Sukvichai K PB and TI. Improving Relocalization in Visual SLAM by using Object Detection. 9th Int Conf Electr Eng Comput Telecommun Inf Technol (ECTI-CON). Pr Khiri Khan. Thail. 2022;1–4.

176. Li R, Liu Q, Gui J DG and HH. Indoor Relocalization in Challenging Environments With Dual-Stream Convolutional Neural Networks. IEEE Trans Autom Sci Eng. 2018;15(2):651–62.

177. Dong S, Fan Q, Wang H, Shi J, Yi L, Funkhouser T, et al. Robust Neural Routing Through Space Partitions for Camera Relocalization in Dynamic Indoor Environments. IEEE/CVF Conf Comput Vis Pattern Recognit (CVPR), Nashville, TN, USA. 2021;8540–50.

178. Nakashima R, Seki A. SIR-Net : Scene-Independent End-to-End Trainable Visual Relocalizer Ryo Nakashima. Int Conf 3D Vis (3DV). Quebec City QC. Canada. 2019;472–81.

179. Zhou L. Visual Relocalization using Long-Short Term Memory Fully Convolutional Network. IEEE Int Symp Mix Augment Real Adjun (ISMAR-Adjunct), Munich, Ger. 2018;258–63.

180. Duong ND, Kacete A, Sodalie C, Oierre-Yves R JR. xyzNet: towards Machine learning camera relocalization by using a scene coordinate prediction network. IEEE Int Symp Mix Augment Real Adjun. 2018;2–7.

181. Wu X, Tian X, Zhou J, Xu P, Chen J. Loop Closure Detection for Visual SLAM Based on SuperPoint Network. 2019 Chinese Autom Congr (CAC). Hangzhou. China. 2019;3789–93.

182. Merrill N, Huang G. Lightweight Unsupervised Deep Loop Closure. Conf Robot Sci Syst . 2018;1–10.

183. Xia Y, Li J, Qi L, Fan H. Loop Closure Detection for Visual SLAM Using PCANet Features. Int Jt Conf Neural Networks (IJCNN), Vancouver BC. Canada,. 2016;2274–81.

184. Dai K, Cheng L, Yang R, Yan G. Loop Closure Detection Using KPCA and CNN for Visual SLAM. 40th Chinese Control Conf (CCC). Shanghai. China. 2021;8088–93.

185. Xiong F, Ding Y, Yu M, Zhao W NZ and PR. A Lightweight sequence-based Unsupervised Loop Closure Detection. Int Jt Conf Neural Networks (IJCNN). Shenzhen. China. 2021;1–8.

186. Huang L, Zhu M, Zhang M. Visual Loop Closure Detection Based on Lightweight Convolutional Neural Network and Product Quantization. IEEE 12th Int Conf Softw Eng Serv Sci (ICSESS). Beijing. China. 2021;122–6.

187. Zhu M, Huang L. Fast and Robust Visual Loop Closure Detection with Convolutional Neural Network. IEEE 3rd Int Conf Front Technol Inf Comput (ICFTIC). Greenville SC. USA. 2021;3681–91.

188. Ma J, Wang S, Zhang K, He Z, Huang J XM. Fast and Robust Loop-Closure Detection via Convolutional Auto-Encoder and Motion Consensus. IIEEE Trans Ind Informatics. 2022;18(6):3681–91.

189. Cai S, Zhou D, Guo R, Zhou H, Peng K. Implementation of Hybrid Deep Learning Architecture on Loop-Closure Detection. 2018; 521–6.

190. Liu Y, Xiang R, Zhang Q, Ren Z, Cheng J. Loop Closure Detection based on Improved Hybrid Deep Learning Architecture. IEEE Int Conf Ubiquitous Comput Commun Data Sci Comput Intell Smart Comput Netw Serv (SmartCNS). Shenyang. China. 2019;312–7.

191. Shi X, Li L. Loop Closure Detection for Visual SLAM Systems Based on Convolutional Netural Network. IEEE 24th Int Conf Comput Sci Eng (CSE). Shenyang. China. 2021;123–9.

192. Zhou Y, Wang Y, Poiesi F, Qin Q, Wan Y. Loop Closure Detection Using Local 3D Deep Descriptors. IEEE Robot Autom Lett. 2022;7(3):6335–42.

193. Osman H, Darwish N, Member S, Bayoumi A. LoopNet: Where to Focus? Detecting Loop Closures in Dynamic Scenes. IEEE Robot Autom Lett. 2022;7(2):2031–8.

194. Bhutta MUM, Sun Y, Lau D, Liu M, Member S. Why-So-Deep : Towards Boosting Previously Trained Models for Visual Place Recognition. 1824 IEEE Robot Autom Lett. 2022;7(2):1824–31.

195. Gauglitz S, Sweeney C, Ventura J MT and TH. Live Tracking and Mapping from Both General and Rotation-Only Camera Motion. IEEE Int Symp Mix Augment Real (ISMAR). Atlanta GA. USA. 2012;13–22.

196. Daniel HC, Kim K, Kannala J, Pulli K, Heikkilä J. DT-SLAM: Deferred triangulation for robust SLAM. 2nd Int Conf 3D Vision. Tokyo. Japan. 2014;609–16.

Zoulikha Bouhamatou: https://orcid.org/0000-0002-3985-0147

Foudil Abdessemed: https://orcid.org/0000-0003-0935-3147

PREDICTIVE ANALYSIS ON THE INFLUENCE OF Al2O3 AND CuO NANOPARTICLES ON THE THERMAL CONDUCTIVITY OF R1234yf-BASED REFRIGERANTS

Baiju S. BIBIN* ⬤**, Panitapu BHRAMARA**** ⬤**, Arkadiusz MYSTKOWSKI***** ⬤**, Edison GUNDABATTINI*** ⬤

*Department of Thermal and Energy Engineering, School of Mechanical Engineering,
Vellore Institute of Technology (VIT), Vellore 632 014, India
**Department of Mechanical Engineering, JNTUH College of Engineering, Hyderabad 500085, India
***Department of Automatic Control and Robotics, Faculty of Electrical Engineering, Bialystok University of Technology,
Wiejska 45D, 15-351, Bialystok, Poland

bibinb.s2019@vitstudent.ac.in, bhramara74@jntuh.ac.in, a.mystkowski@pb.edu.pl, edison.g@vit.ac.in

Abstract: Nano-enhanced refrigerants are substances in which the nanoparticles are suspended in the refrigerantatthe desired concentration. They have the potential to improve the performance of refrigeration and air-conditioning systems that use vapour compression. This study focuses on the thermal conductivity of alumina (Al2O3) and cupric oxide (CuO) nanoparticles immersed in 2,3,3,3-tetrafluoropropene (R1234yf). The thermal conductivity of nano-refrigerants was investigated using appropriate models from earlier studies where the volume concentration of particles and temperatures were varied from 1% to 5% and from 273 K to 323K, respectively. The acquired results are supported by prior experimental investigations on R134a-based nano-refrigerants undertaken by the researchers. The main investigation results indicate that the thermal conductivity of Al2O3/R1234yf and CuO/R1234yf is enhanced with the particle concentrations, interfacial layer thickness, and temperature. Also, the thermal conductivity of Al2O3/R1234yf and CuO/R1234yf decreases with particle size. The thermal conductivity of Al2O3/R1234yf and CuO/R1234yf nano-refrigerants become enhanced with a volume concentration of nano-sized particles by 41.2% and 148.1% respectively at 5% volume concentration and 323K temperature. The thermal conductivity of Al2O3/R1234yf reduces with temperature, by upto 3% of nanoparticle addition and after that, it enhances. Meanwhile, it declines with temperature, by upto 1% of CuO nanoparticle inclusion for CuO/R1234yf. CuO/R1234yf has a thermal conductivity of 16.69% greater than Al2O3/R1234yf at a 5% volume concentration. This paper also concludes that, among the models for thermal conductivity study, Stiprasert's model is the most accurate and advanced.

Keywords: heat transfer fluid, nanoparticles, nano-refrigerants, thermal conductivity models, thermo-physical properties, nanofluids, temperature, volume concentration

1. INTRODUCTION

Nanofluids have piqued the interest of researchers all over the world as a viable option for improving the efficiency of heat transfer. Nanofluids were first exhibited by Choi and Eastman [1] at Argonne National Laboratory as immersed nano-sized materials in base fluids with a standard particle size of 1–100 nm. Nanofluid is a heat transport medium, which is inflexible and established by suspending nano-sized particles unvarying isolated into heat transfer fluid (HTF). The merit of these particles progresses bythe fluid heat transportation created by their advanced thermal conductivity. Nanofluids are of boundless worth due to their higher enactment on the advancement of the Thermal Act. Nanofluids are used in a variety of heat transfer applications, including electronic cooling, aerospace industries, refrigeration, and air-conditioning [2]. Two approaches are used to improve the heat transfer coefficient of nanofluids. The first approach is to boost the Nusselt number, which is impacted by the Reynolds, Prandtl, and geometry. The second strategy is to shorten the characteristic length, which is inversely related to the heat transfer coefficient. The heat transfer coefficient is deter-mined using the Nusselt number, nanofluid thermal conductivity, and characteristic length in this approach. These approaches are used to create microchannel heat sinks [3]. Recent studies analyse the thermal performance of Joule heating, Brownian motion and thermo-phoretic diffusion on Carreau, Casson-Willianson and Maxwell nanofluids [4–14]. These investigations are leading to more practical applications of nanofluids. The thermophysical properties of nano-fluids influence the performance of the systems. One of their most critical thermo-physical properties is thermal conductivity, which is related to the convective heat transfer and boiling coefficients. As a result, thermal conductivity has gained a lot of attention in current research, and the non-linear dependency of temperature on thermal conductivity has been verified [15–23]. Alhajaj et al. [24] identified the thermal conductivity of 0.5% of Al₂O₃ and TiO₂ nanoparticles with water and ethylene glycol. Plant and Saghir [25] studied the heat transfer performance of the three-channel heat exchanger using 1% Al₂O₃/water and 2% Al₂O₃/water nanofluids. In their study, they concluded that heat transfer is enhanced by the increase in the volume concentration of the nanoparticle in the base fluid. From their analysis, it is clear that the thermal conductivity of the nanofluid greatly impacts the performance of systems.

Nano-refrigerants are a novel kind of nanofluids, in which the nanoparticles are suspended on the base refrigerants. Many researchers have identified the 'k' value of different refrigerants [26–28]. This novel kind of refrigerant is used for enhancing the performance of refrigeration systems. The 'k' value of the working fluid influences the heat transfer of the system [29]. Three key benefits could be achieved by using nanoparticles in refrigeration systems [30]: (a) nanoparticles as an additive could improve the mixing of the lubricant and the refrigerant. (b) the refrigerant's 'k' value and other thermo-physical properties could be enhanced and (c) dispersion of nanoparticles in the refrigerant could increase the coefficient of performance and cooling effect and enhance the energy efficiency of the system. The experiments showed that the nano-refrigerant has a greater thermal conductivity compared with that of the base refrigerant [15]. In addition, the vapour compression refrigeration system works with the nano-enhanced refrigerant and outperforms the traditional pure refrigerant refrigeration system [31–33]. On the contrary, the clustering or settlement of nanoparticles in refrigerants could reduce the steadiness of the nano-refrigerant and restrict its use in refrigeration systems. Many studies have been carried out on the thermo-physical properties and the ability of heat transport of nano-refrigerants. According to the literature, suspending nanoparticles in refrigerants could improve the thermal conductivity (k) and heat transfer efficiency [16, 17].

Recently researchers have tried to identify the 'k' values of various nano-enhanced refrigerants. Jiang et al. [17] calculated the value of 'k' for four distinct forms of carbon nanotubes at particle concentrations of 0.2%–1% added in R113 refrigerant. With regulated volume concentrations of 0.1%–1.2%, Jiang et al. [15] studied the 'k' value of R113-based nano-refrigerant, in which they used various nanoparticles such as Cu, Al, Ni, CuO, and Al_2O_3. Alawi et al.[18] analysed the value of 'k' for R134a with CuO nanoparticles for different volume concentrations varying from 1% to 5% and temperatures in the range of 300–320K. Mahbubulet al. [19] also explored the influence of 'k' on the R134a-based nano-refrigerant. They used Al_2O_3 as the suspended particle with concentrations of 1–5 vol%. Another work by Alawi et al. [34] analysed the value of 'k' for nano-refrigerants containing alumina nanoparticles and R141b as the base refrigerant with the variation in concentration from 1 vol% to 4 vol%. The impacts of the volume concentration of nanoparticles present in the base fluid, types of nanoparticles, refrigerants, sizes of particles, and shapes of particles could all affect the value of 'k'. Thermal conductivity could be improved as interfacial layers are created in nanofluids [35–37]. Likewise, Patil et al. [38] identified that the 'k' value of the nano-refrigerant declined with the particle size. According to Zawawi et al. [39], the 'k' value of nanofluids is influenced by their volume concentration and temperature. The value of 'k' was improved with the concentration of nanomaterial and it declined with temperature. Yang et al. [40] found that increasing the volume fraction of TiO_2 increased the value of 'k' of R134a/TiO_2, 30% of the 'k' value was enhanced by adding 5% of the volume concentration of TiO_2 to the standard fluid. They also realised a decrease in thermal conductivity with an increase in temperature. Ammar Hassan et al. [41] explored the thermal conductivity of the R134a-based nano-refrigerant. They used Al_2O_3, SiO_2, ZrO_2, and CNT (Carbon nano-tubes) as nano inhibitors with different mass concentrations. From these, they identified that the thermal conductivity was improved with the nanomaterial concentration in the refrigerant. The main advantagesof the nanoparticles in the refrigeration systems are given in Fig. 1.
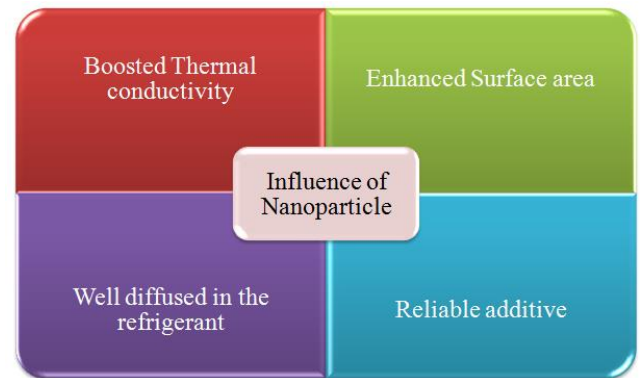
Fig.1. Significant benefits of nanoparticles in refrigeration system

Hydro-fluoro-olefin (HFO) refrigerants are the fourth generation, which have low global warming potential (GWP) and zero ozone depletion potential (ODP). Hence, it is a green refrigerant. These mixtures can replace HFCs like R125, R134a, R32 and others legally or with a small system modification [42]. There are two important types of HFO refrigerants: R1234ze(E) and R1234yf. The major limitation of R1234yf is its lesser performance than R134a. So, this article focuses on the system's efficiency improvement with R1234yf using nanoparticles.

However, investigations on the thermo-physical properties of HFO-based (R1234yf) nano-refrigerants are limited. The major goal of this analysis is to evaluate the thermal conductivity of Al_2O_3/R1234yf and CuO/R1234yf nano-refrigerants using mathematical models. The study examines the impact of nanoparticle concentrations ranging from 1% to 5% by volume. The investigation of nano-refrigerant thermal conductivity, on the contrary, is broadened by identifying and including the effects of nano-refrigerant temperature and particle size, as well as the interfacial layer on the thermal conductivity of the novel refrigerant.

2. METHODOLOGY

The 'k' values of Al_2O_3/R1234yf and CuO/R1234yf nano-refrigerants are predicted by advanced Stipresert's model and compared with classical models such as those of Maxwell, Crosser & Hamilton and Yu and Choi. In Tabs. 1 and 2 are tabulated the properties of Al_2O_3 and CuO nanomaterials and the R1234yf refrigerant. The nano-refrigerant efficacy is investigated by studying the effects of Al_2O_3 and CuO nanoparticle concentration of 1–5 vol% in R1234yf refrigerant. The average particle size of nanoparticles was 40nm. However, when studying the influence of the temperature and size of nanoparticles on the 'k' value of nano-added refrigerant, the nano-sized particles radius is considered to be 10–50nm and the temperature varied from 273 K to 323K. For the stability of the nano-refrigerant, no surfactant was used. As a result, the effect of surfactants was overlooked during the research. The mathematical models of nanofluids and nano-refrigerants from different authorities and the thermal conductivity of CuO/R1234yf and Al_2O_3/R1234yf nano-enhanced refrigerants were analysed.

Tab. 1. The properties of Al2O3 and CuO nanoparticles

Nano particle	Molecular weight (kg/kmol)	Density (kg/m³)	Thermal conductivity (W/mK)	Specific heat (J/kgK)	Source
Al₂O₃	101.96	3,880	40	729	[43]
CuO	-	6,320	32.9	550.5	[44]

Tab. 2. Properties of R1234yf refrigerant [45]

`Temperature (K)	Pressure (kPa)	Density (kg/m³)	Viscosity (µPa/s)	Thermal conductivity (W/mK)	Specific heat (kJ/kgK)
273	315	1,175	220	0.0746	1.259
283	436	1,144	194	0.0713	1.293
293	590	1,111	171	0.0672	1.332
303	782	1,075	152	0.0631	1.379
313	1017	1,037	134	0.0586	1.498
323	1301	993.3	118	0.054	1.566

3. MATHEMATICAL MODELS

3.1. Maxwell's Model

Thermal conductivity models were built based on Maxwell's classic research into conduction through amalgamated media. This model was created by Maxwell under the premise that the solid phase is spherical. According to Maxwell's model, the thermal conductivity of the base fluid, sphere-shaped particles and particle volume percentage impact the '*k*' value of the nano-refrigerant. In addition, this model was found to be valid only in the case when $\varphi \ll 2.5\%$. However, the model considers the mixture of fluid and nanoparticle as a homogeneous fluid. Maxwell [46] looked at the efficient thermal conductivity of a multiphase mixture of solid and liquid states, and the efficient '*k*' value of nano-refrigerant, k_nr is given as

$$\frac{k_{\text{nr}}}{k_{\text{r}}} = \frac{k_{\text{np}} + 2k_{\text{r}} - 2\varphi(k_{\text{r}} - k_{\text{np}})}{k_{\text{np}} + 2k_{\text{r}} + \varphi(k_{\text{r}} - k_{\text{np}})} \tag{1}$$

According to Maxwell's model, the '*k*' value of the nano-refrigerant is influenced by the particle volume concentration and thermal conductivity of the nano-sized particles and fluids. Meanwhile, the other classical models include the impact of particle shape, particle distribution and particle interactions. These models predict almost identical improvements at low concentrations. However, the classical models donot account for the particle size and the interfacial layer.

3.2. Hamilton & crosser model

The shape factor (*f*) was introduced by Hamilton and Crosser [47], who extended Maxwell's work to non-sphere-shaped particles and termed it a new parameter that is related to the structure of the particles. The shape factor has been measured experimentally for various materials. They tried to establish a model that took particle structure, composition, and conductivity into account in both the solid and fluid phases. The Hamilton & Crosser model estimates the value of '*k*' for a nano-refrigerant as follows:

$$\frac{k_{\text{nr}}}{k_{\text{r}}} = \frac{k_{\text{np}} + (f-1)k_{\text{r}} - (f-1)(k_{\text{r}} - k_{\text{np}})\varphi}{k_{\text{np}} + (f-1)k_{\text{r}} + (k_{\text{r}} - k_{\text{np}})\varphi} \tag{2}$$

The shape factor (*f*) is calculated by the formula $f = 3/\chi$ and χ is known as sphericity, which is the ratio of a particle's surface area to its volume. The sphericity of spherical and cylindrical shapes is 1 and 0.5 respectively.

3.3. Yu and Choi Model

By assuming that, the pure refrigerant molecules near the solid surface of nanoparticles form layered structures that approximate solids, Yu and Choi [35] altered Maxwell's model. Consequently, the interfacial nano-layer works as a thermal link between the pure refrigerant and the solid nano-sized particles, enhancing the value of '*k*'. To account for the role of the nanolayer in measuring efficient '*k*' value, Yu and Choi considered a spherical nano-sized particle of radius enclosed by an interfacial nano-layer of thickness. They also believed that the thermal conductivity of the interfacial nano-layer is greater than thatof the refrigerant. The value of '*k*' of the nano-enhanced refrigerant is determined as:

$$\frac{k_{\text{nr}}}{k_{\text{r}}} = \frac{k_{\text{np}} + 2k_{\text{r}} - 2\varphi(k_{\text{r}} - k_{\text{np}})(1+\beta)^3}{k_{\text{np}} + 2k_{\text{r}} + \varphi(k_{\text{r}} - k_{\text{np}})(1+\beta)^3} \tag{3}$$

The '*k*' of solid nano-sized particles and the base refrigerant, as well as the volume fraction of suspended particle, particle structure, interfacial nanolayer thickness and '*k*', all affect the '*k*' value of nano-refrigerants. The β is the ratio of the equivalent particle (with interfacial nano-layer) radius to the nanoparticle radius. This is influenced by the nano-particle size and the interfacial thickness.

3.4. Stiprasert's Model

The thermal conductivity of nano-refrigerant was measured using the Sitprasert et al. [36] correlation. This model could be used to quantify the influences of nanoparticle size, nanoparticle volume, and a temperature-dependent interfacial layer. Fig.2 shows a diagrammatic representation of the interfacial layer.
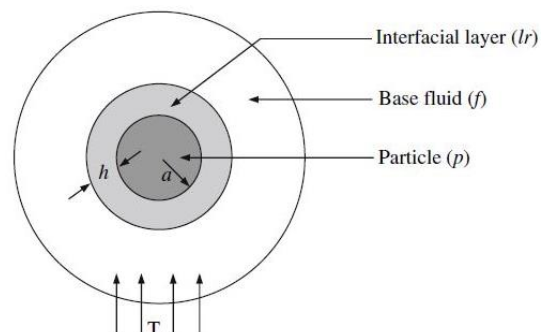
Fig.2. Solid nanoparticle interfacewithfluid medium [29]

The effective thermal conductivity is determined as:

$$k_{nr} = \frac{(k_{np}-k_{lr})\varphi k_{lr}(2\beta_1^3-\beta^3+1)+(k_{np}+2k_{lr})\beta_1^3[\varphi\beta^3(k_{lr}-k_r)+k_r]}{\beta_1^3(k_{np}+2k_{lr})-(k_{np}-k_{lr})\varphi[\beta_1^3+\beta^3-1]} \quad (4)$$

Where,

$$\beta = 1 + \frac{h}{a} \quad (5)$$

$$\beta_1 = 1 + \frac{h}{2a} \quad (6)$$

$$h = 0.01(T-273)a^{0.35} \quad (7)$$

$$k_{lr} = C\frac{h}{a}k_r \quad (8)$$

The value of C = 30 for Al_2O_3 and C = 110 for CuO nanoparticles [36].

4. RESULTS AND DISCUSSION

4.1. Influence of nano-enhanced refrigerant volume fraction and temperature on thermal conductivity

The volume concentration of nanoparticles in the base refrigerant and temperature are major parameters influencing thermal conductivity. When intensifying the amount of nanoparticles in the refrigerant the properties will change. In addition, the properties vary with the temperature. Due to this fact, all the mathematical models have considered volume concentration and temperature as prime parameters influencing the propertiesof the refrigerant.

Figs. 3 and 4 depict the 'k' value of Al_2O_3/R1234yf and CuO/R1234yf nano-refrigerant respectively. Here the volume concentrations change from 1% to 5% and the temperature ranges from 273K to 323K. The 'k' value of nano-refrigerant is predicted using Stiprasert's model. The nano-enhanced refrigerants show the highest 'k' value at 5% of nanoparticle concentration and a temperature of 323K. The rise in the value of 'k' is proportional to the increase in the volume fraction of Al_2O_3 and CuO, as seen in the figures. Fig.3 indicates that the thermal conductivity of the Al_2O_3/R1234yf refrigerant is enhanced by 41.2% at 323K temperature and 5% volume concentration. From Fig.4, the improvement in the 'k' value of CuO/R1234yf refrigerant is obtained as 148.1% at 5% volume concentration and 323K temperature. The enhancement of 'k' values with the addition of nanoparticles to the base refrigerant is due to the enlargement of specific surface areas for heat transfer. The nanoparticles have more surface area than the bulk materials, hence when these are added to the base refrigerant, heat transfer increases due to more specific surface area. Therefore the 'k' value of the nano-added refrigerant increases together with the concentration (vol.%) of nanoparticles.

From Figs.3 and 4 it is seen that the variation of the 'k' value of both nano-enhanced refrigerants with temperature is different. The 'k' value of Al_2O_3/R1234yf is continuously reduced with the rise in temperature of upto 3% of Al_2O_3 nanoparticle addition. At higher concentrations, it enhanced with temperature. At 323K, the 'k' value is reduced by 27.61% with 1% addition and it is enhances by 2.2% with 5% inclusion of nanoparticles as compared with 273K. However, for CuO/R1234yf nano-enhanced refrigerant the 'k' value is reduced with temperature by upto 1%

inclusion of CuO nanoparticle. After that, the thermal conductivity of the nano-refrigerant improved with temperature. The 'k' value is reduced by upto 7.2% at 323K and 1% volume concentration. Meanwhile, it is enhanced by 79.59% with 5% CuO nano-sized particle addition and 323K temperature as compared with 273K. At low volume concentrations of the nanoparticle in the base-refrigerant, the molecules of the refrigerant move apart from each other when heated, this leads to increasing their mean path. Thus this reduces the probability of molecular collision. Hence at lower concentrations, the 'k' value of the nano-refrigerants reduces with a rise in temperature [48]. However, at high-volume concentrations, the density of the nanoparticles is high,so that when heat is added to the fluid, the molecules move rapidly. Consequently, the nanoparticles in the refrigerant tend to move faster due to the Brownian motion and high energy content, which leads to improving the 'k' value with the temperature rise [39, 49].
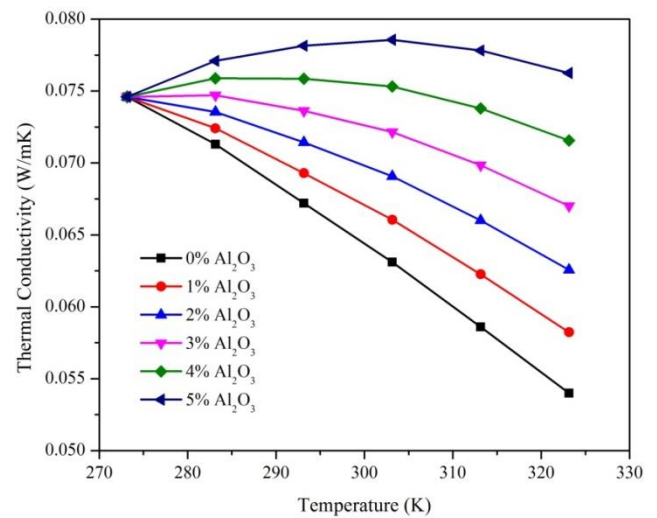
Fig.3. Dependency of Al2O3 on thermal conductivity at different temperatures obtained by Stiprasert's model
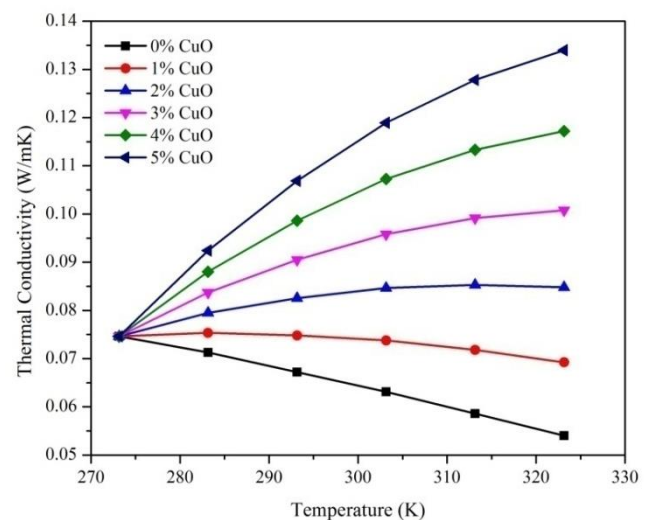
Fig.4. Impact of CuO on thermal conductivity at different temperatures obtained by Stiprasert's model
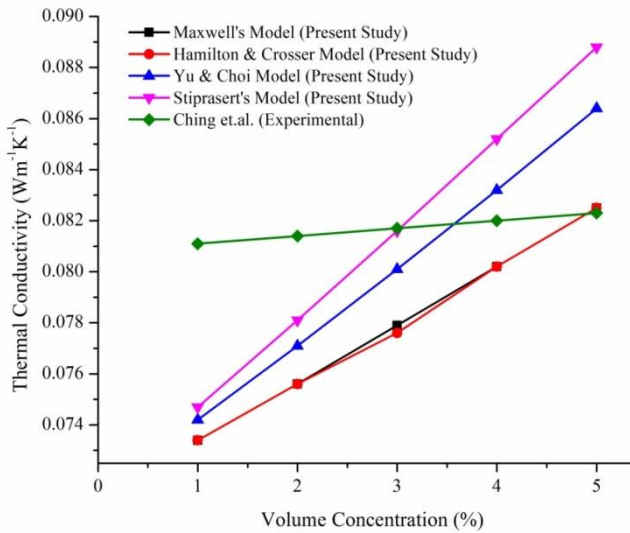
Fig.5. Dependency of volume concentration of thermal conductivity

Fig.6. Validation of results with experimental and simulation studies of Al2O3/R134a

The R134a-based nano refrigerant is chosen for validation of the mathematical models. Fig.5 compares the 'k' value of R134a-based nano refrigerant with different nanoparticle volume concentrations of Al₂O₃ in the refrigerant obtained by various thermal conductivity models at constant temperature (283K) and the particle having constant size (40nm). The 'k' value of the nano-enhanced refrigerant is improved with increase in volume fractions. The value of 'k' varies with models due to the assumptions made by each model being different. From the figure, it is established that the highest value of 'k' is obtained by Stiprasert's model. This is because Stiprasert's model considers the influence of the interfacial nano-layer between nanoparticles as well as refrigerant. As the spherical shape of nanoparticles is considered, the 'k' value of Maxwell and Hamilton & Crosser models is the same. Fig.5 also indicates that the present study result shows a similar variation to the experimental data by Jwo et al. [50], that is, the 'k' value enhances with the volume concentration of nanoparticles. The experimental values are different from the predicted values. This is because of the assumptions taken for the predictions. However, the variation of the 'k' value with volume concentration is in the same trend. They studied experimentally to identify the 'k' value of nano-refrigerant containing R134a as refrigerant and Al₂O₃ as the nanoparticle. In this case, the 'k' value of the nano-refrigerant is also slightly enhanced with the volume fraction of nanoparticles in the mixture.

Fig.6 shows the influence of temperature on the thermal conductivity of nano-added refrigerants by various mathematical models at a particular volume concentration (equals 5%) and constant particle size (40nm). The Maxwell and Yu and Choi models, demonstrate that as the temperature rises, the value of 'k' for nano-enhanced refrigerant is lessened. The value of 'k' for the nano-refrigerant improved with temperature according to Stiprasert's model. The same kind of variation is obtained in the studies conducted by Jwo et al. [50] and Mahbubul et al. [19] with a 5% particle concentration. The thermal conductivity of the nano-refrigerant is analysed using mathematical correlation by Mahbubul et al. [19]. The researchers worked with an R134a-based nano-refrigerant and they obtained that the 'k' value of the Al₂O₃/R134a refrigerant is enhanced with temperature. Therefore Stiprasert's model is closer to the real case in comparison with other investigated models.
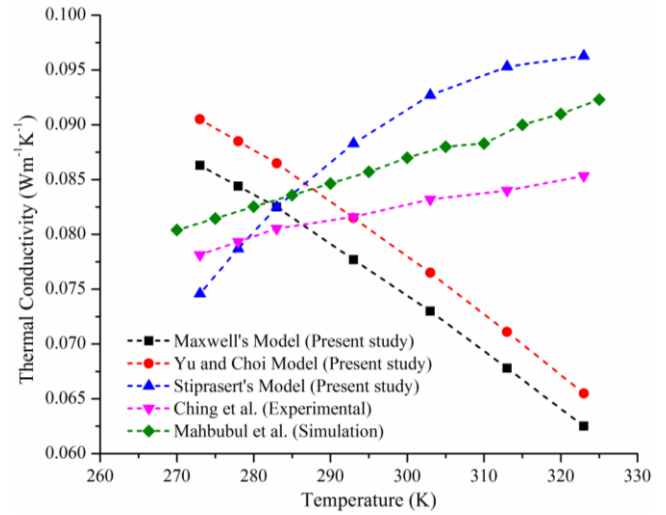
Since the value of 'k' for nano-sized particles is usually more compared to the base refrigerant, the 'k' value of the nano-refrigerant must also be higher [17]. The Brownian motion of the nanoparticle is intensified with the temperature rise and the heat transfer is boosted with the contribution of micro convection, which leads to the thermal conductivity enhancement [19]. The result shows as the temperature rises, the 'k' value of nano-refrigerant also rises.

4.2. Impact of the nano-enhanced refrigerant particle size on thermal conductivity

Particle size is a critical parameter that acts on the property of nano-refrigerant. This parameter is considered in the Yu and Choi model as well as in the Stiprasert model.
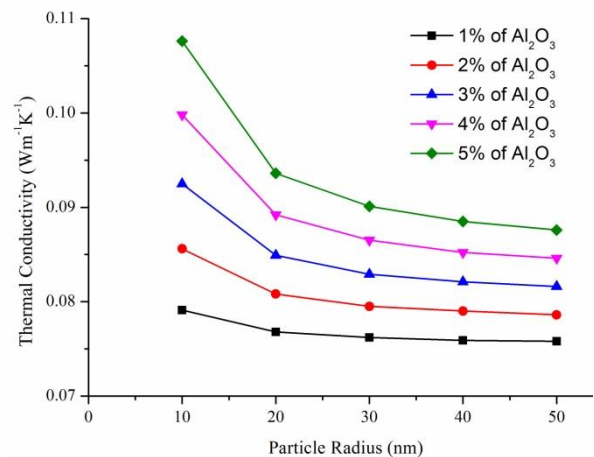
Fig.7. Influence of Al2O3 particle radius on thermal conductivity of R1234yf

Fig.7 depicts the influences of the size of Al₂O₃ nanoparticles on thermal conductivity at a constant temperature (283K) obtained by the Yu and Choi model. From the figure, it is seen that the 'k' value is reduced with nanoparticle size increment. A similar trend is observed with Stipersert's model as shown in Fig.8. The

'*k*' value of the CuO-based nano-refrigerant is decreased with a particle size that is revealed in Fig.8. Both Figs.7 and 8 also indicate that the '*k*' values of both nano-enhanced refrigerants increased with the increase in volume fraction of Al_2O_3 and CuO nanoparticles.
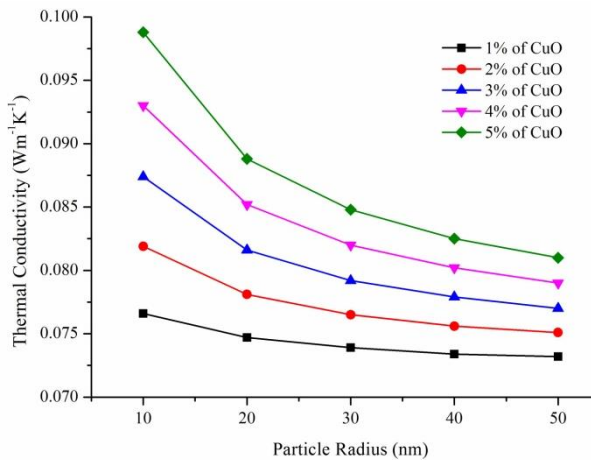
Fig.8. Influence of particle radius on thermal conductivity of CuO/R1234yf refrigerant

The '*k*' value of nano-enhanced refrigerants is decreased with particle size. This is because when the particle size increases the bulkiness of the material increases. The material's specific surface area is reduced as a result of its increased bulkiness. Heat transfer is diminished as the specific surface area is decreased. Hence the thermal conductivity reduces with the enlargement of particles.
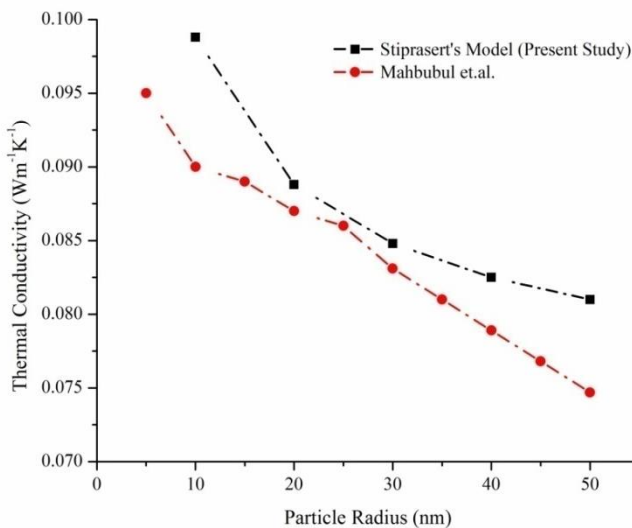
Fig.9. Comparison of the influence of particle radius on thermal conductivity with the previous study

Fig.9 shows a comparison of the trend in variation of the '*k*' value of the nano-refrigerant with particle radius with a previous study by Mahbubul et al. [19]. They studied the thermal conduc-

tivity of R134a-based nano-refrigerants whereas the present study is on R1234yf-based nano-refrigerants. Due to the insufficient experimental work carried out with R1234yf-based nano-refrigerant, the comparison is done with R134a-based nano-refrigerant. For comparison, the result obtained using Stiprasert's model in R1234yf-based nano-refrigeranthas been considered. From Fig.9, it is clear that the '*k*' value of both studies decreases with an increase in nanoparticle radius. Hence the mathematical model developed by Stiprasert is suitable for the analysis of the effect of particle radius on the '*k*' value in nano-refrigerant studies.

4.3. Impact of the interfacial nanolayer between nanoparticle and refrigerant on thermal conductivity

By adding nanoparticles to the refrigerant, a layer between the substances is formed. This layer is called the interfacial layer. The thickness of this layer also influences the properties of the nano-refrigerant. Yu and Choi and Stiprasert have considered that interfacial nanolayer is an important parameter that affects thermal conductivity.

Figs.10 and 11 show the influence of nanolayer or interfacial layer thickness on thermal conductivity at a constant particle radius (equals 20nm). The '*k*' value of nano-refrigerant is improved by enhancement in nanolayer thickness in both cases of CuO/R1234yf and Al_2O_3/R1234yf nano-refrigerants. Due to the imperfect contact between the solid–solid interfaces, a resistance is created in the interface. This interface resistance is known as Kapitza resistance, which acts as a barrier to heat transfer and lowers the effective thermal conductivity. However, this resistance phenomenon is not predominant at the solid–liquid interface of particles in liquid suspension. Due to these, the nanolayer enhances the thermal conductivity of the nano-refrigerant [35]. The interfacial layers surrounding the nanoparticles are enhancement mechanisms that raise the '*k*' value of nano-refrigerant by raising the specific surface area of the nanoparticles [19]. The total surface area per unit mass of a nanoparticle is referred to as the specific surface area.
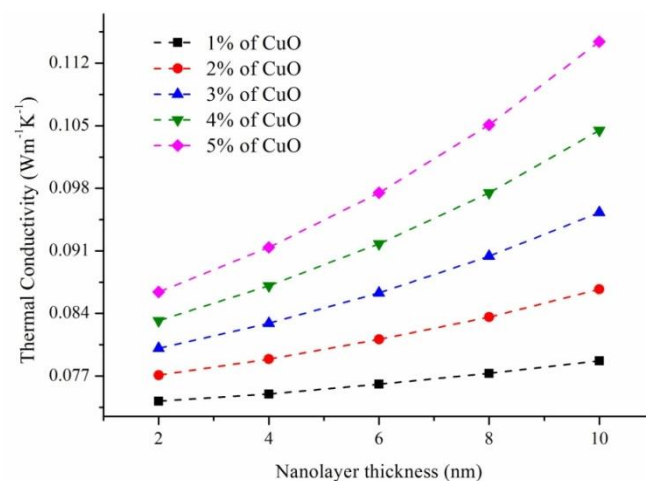
Fig.10. Variation of thermal conductivity with nano interfacial layer for CuO/R1234yf

sciendo

Baiju S. Bibin, Panitapu Bhramara, Arkadiusz Mystkowski, Edison Gundabattini
Predictive Analysis on the Influence of Al₂O₃ and CuO Nanoparticles on the Thermal conductivity of R1234yf-Based Refrigerants

DOI 10.2478/ama-2024-0050
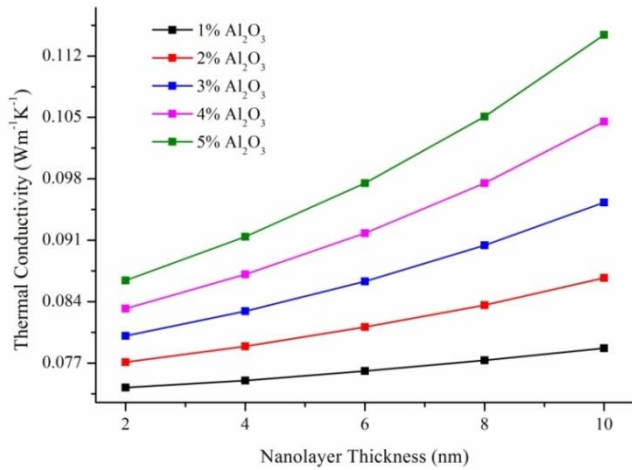
Fig.11. Rise in thermal conductivity with nano interfacial layer for Al2O3/R1234yf

4.4. Effect of type of nanoparticle on the thermal conductivity

In common, each nanoparticlehas different values of thermal conductivity. However, based on the models developed by Maxwell, Hamilton & Crosser and Yu and Choi, the value of 'k' of Al₂O₃ and CuO nanoparticles added to the refrigerant gives similar values and does not influence the type of nanoparticles. However, this is not practically correct as the type of nanoparticle and refrigerant affects the 'k' value of the nanofluid. The nanoparticle type dependency is seen in the Stipresert model.

Fig. 12 compares the thermal conductivity of CuO/R1234yf and Al₂O₃/R1234yf nano-refrigerant at 283K and the particles having a size of 40nm, which is obtained by Stiprasert's mathematical model. From the figure, it is clear that the 'k' value of the nano-refrigerant depends on the kind of nanoparticles which are added to the base refrigerants. The highest values of 'k' are obtained when the R1234yf refrigerant is mixed with CuO. At 5% volume concentration, the thermal conductivity observed by adding CuO in R1234yf is 16.69% greater than that of Al₂O₃ nanoparticles added to the same refrigerant.
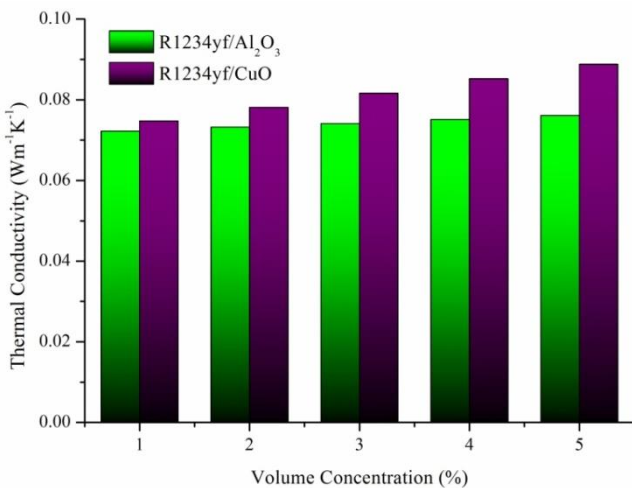
Fig.12. Comparison of thermal conductivity of R1234yf-based nano-refrigerant

5. CONCLUSION

The thermal conductivity of the R1234yf refrigerant-based nano-refrigerant has been investigated using mathematical models such as Maxwell, Crosser & Hamilton, Yu and Choi and Stiprasert's models, and the results have been validated using previous experimental studies conducted by different researchers. The thermal conductivity of Al₂O₃/R1234yf and CuO/R1234yf nano-refrigerants improves with increasing nanoparticle volume concentrations, similar to nanofluids, according to the findings. The thermal conductivity of Al₂O₃/R1234yf and CuO/R1234yf nano-refrigerants was enhanced with a volume concentration of nano-sized particles by 41.2% and 148.1% respectively at 5% volume concentration and 323K temperature. The thermal conductivity of Al₂O₃/R1234yf is reduced with temperature, by upto 3% of nanoparticle addition and after that, it is enhanced. Meanwhile, it declined with temperature, by upto 1% of CuO nanoparticle inclusion for CuO/R1234yf. However, the value of thermal conductivity is improved with temperature at higher concentrations as observed in theStiprasert model which is a positive feature for all real-world applications. By increasing the particle size, the thermal conductivity of CuO/R1234yf and Al₂O₃/R1234yf nano-refrigerants is decreased. The interfacial nanolayer also impacted the thermal conductivity value of the nano-refrigerant such that as the thickness ofthe interfacial nanolayer increased the value of thermal conductivity also increased. The thermal conductivity value of CuO/R1234yf is more than that of Al₂O₃/R1234yf about 16.69% at a 5% volume concentration. Compared with the results from the mathematical models in previous experimental studies, Stiprasert's model gives more realistic results. Therefore, the Stiprasert model is the best and most advanced model for study of thermal conductivity. The only drawback of this model is that it applies only to Al₂O₃ and CuO nanoparticles.

Nomenclature

Acronyms		Subscripts:	
a	Particle Radius (nm)	lr	Interfacial
f	Shape factor	layer	
GWP	Global Warming Potential	nr	Nano-refrigerant
h	Thickness of Interfacial nanolayer (nm)	np	Nanoparticle
HFO	Hydro fluoro olefin	r	Refrigerant
HTF	Heat transfer fluid	Greek Symbols:	
k (W/mK)	Thermal Conductivity	ϕ	Particle Volume fraction
ODP	Ozone Depletion Potential	χ	Sphericity
T	Temperature (K)		
Vol	Volume		

REFERENCES

1. Choi S, Eastman J. Enhancing thermal conductivity of fluids with nanoparticles. In: 1995 International mechanical engineering congress and exhibition. San Francisco. CA (United States). 1995.
2. Shukla RK, Dhir VK. Effect of Brownian motion on thermal conductivity of nanofluids. Journal of Heat Transfer [Internet]. 2008 Mar 18;130(4). Available from: https://doi.org/10.1115/1.2818768

3. Jang SP, Choi SJ. Effects of various parameters on nanofluid thermal conductivity. Journal of Heat Transfer [Internet]. 2006 Aug 2;129(5):617–23. Available from: https://doi.org/10.1115/1.2712475

4. Irfan M. Energy transport phenomenon via Joule heating and aspects of Arrhenius activation energy in Maxwell nanofluid. Waves in Random and Complex Media [Internet]. 2023 Apr 12;1–16. Available from: https://doi.org/10.1080/17455030.2023.2196348

5. Irfan M. Influence of thermophoretic diffusion of nanoparticles with Joule heating in flow of Maxwell nanofluid. Numerical Methods for Partial Differential Equations [Internet]. 2022 Sep 23;39(2):1030–41. Available from: https://doi.org/10.1002/num.22920

6. Rafiq K, Irfan M, Khan MA, Anwar MS, Khan W. Arrhenius activation energy theory in radiative flow of Maxwell nanofluid. Physica Scripta [Internet]. 2021 Jan 28;96(4):045002. Available from: https://doi.org/10.1088/1402-4896/abd903

7. Irfan M, Khan M, Khan WA. Heat sink/source and chemical reaction in stagnation point flow of Maxwell nanofluid. Applied Physics A [Internet]. 2020 Oct 27;126(11). Available from: https://doi.org/10.1007/s00339-020-04051-x

8. Irfan. Study of Brownian motion and thermophoretic diffusion on non-linear mixed convection flow of Carreau nanofluid subject to variable properties. Surfaces and Interfaces. 2021 Apr;23(100926).

9. Irfan M, Anwar MS, Kebail I, Khan WA. Thermal study on the performance of Joule heating and Sour-Dufour influence on nonlinear mixed convection radiative flow of Carreau nanofluid. Tribology International [Internet]. 2023 Oct 1;188:108789. Available from: https://doi.org/10.1016/j.triboint.2023.108789

10. Ali U, Irfan M. Thermal aspects of multiple slip and Joule heating in a Casson fluid with viscous dissipation and thermo-solutal convective conditions. International Journal of Modern Physics B [Internet]. 2022 Sep 22;37(05). Available from: https://doi.org/10.1142/s0217979223500431

11. Irfan M, Rafiq K, Khan M, Waqas M, Anwar MS. Theoretical analysis of new mass flux theory and Arrhenius activation energy in Carreau nanofluid with magnetic influence. International Communications in Heat and Mass Transfer [Internet]. 2021 Jan 1;120:105051. Available from: https://doi.org/10.1016/j.icheatmasstransfer.2020.105051

12. Ali U, Irfan M, Akbar NS, Rehman KU, Shatanawi W. Dynamics of Soret–Dufour effects and thermal aspects of Joule heating in multiple slips Casson–Williamson nanofluid. International Journal of Modern Physics B [Internet]. 2023 Jun 9. Available from: https://doi.org/10.1142/s0217979224502060

13. Irfan M, Aftab R, Khan M. Thermal performance of Joule heating in Oldroyd-B nanomaterials considering thermal-solutal convective conditions. Chinese Journal of Physics [Internet]. 2021 Jun 1;71:444–57. Available from: https://doi.org/10.1016/j.cjph.2021.03.010

14. Irfan M, Khan W, Pasha AA, Alam MI, Islam N, Zubair M. Significance of non-Fourier heat flux on ferromagnetic Powell-Eyring fluid subject to cubic autocatalysis kind of chemical reaction. International Communications in Heat and Mass Transfer [Internet]. 2022 Nov 1;138:106374. Available from: https://doi.org/10.1016/j.icheatmasstransfer.2022.106374

15. Jiang W, Ding G, Peng H, Gao Y, Wang K. Experimental and model research on nanorefrigerant thermal conductivity. Science and Technology for the Built Environment [Internet]. 2009 May 1;15(3): 651–69. Available from: https://doi.org/10.1080/10789669.2009.10390855

16. Mahbubul IM, Saadah AR, Saidur R, Khairul MA, Kamyar A. Thermal performance analysis of Al₂O₃/R-134a nanorefrigerant. International Journal of Heat and Mass Transfer [Internet]. 2015 Jun 1;85:1034–40. Available from: https://doi.org/10.1016/j.ijheatmasstransfer.2015.02.038

17. Jiang W, Ding G, Peng H. Measurement and model on thermal conductivities of carbon nanotube nanorefrigerants. International Journal of Thermal Sciences [Internet]. 2009 Jun 1;48(6):1108–15. Available from: https://doi.org/10.1016/j.ijthermalsci.2008.11.012

18. Alawi OA, Sidik NAC. Influence of particle concentration and temperature on the thermophysical properties of CuO/R134a nanorefrigerant. International Communications in Heat and Mass Transfer [Internet]. 2014 Nov 1;58:79–84. Available from: https://doi.org/10.1016/j.icheatmasstransfer.2014.08.038

19. Mahbubul IM, Fadhilah SA, Saidur R, Leong KY, Afifi AM. Thermophysical properties and heat transfer performance of Al₂O₃/R-134a nanorefrigerants. International Journal of Heat and Mass Transfer [Internet]. 2013 Jan 1;57(1):100–8. Available from: https://doi.org/10.1016/j.ijheatmasstransfer.2012.10.007

20. Mahbubul IM, Saidur R, Afifi AM. Thermal Conductivity, Viscosity and Density of R141b Refrigerant based Nanofluid. Procedia Engineering [Internet]. 2013 Jan 1;56:310–5. Available from: https://doi.org/10.1016/j.proeng.2013.03.124

21. Mahbubul IM, Saidur R, Afifi AM. Influence of particle concentration and temperature on thermal conductivity and viscosity of Al₂O₃/R141b nanorefrigerant. International Communications in Heat and Mass Transfer [Internet]. 2013 Apr 1;43:100–4. Available from: https://doi.org/10.1016/j.icheatmasstransfer.2013.02.004

22. Alawi OA, Sidik NAC. Mathematical correlations on factors affecting the thermal conductivity and dynamic viscosity of nanorefrigerants. International Communications in Heat and Mass Transfer [Internet]. 2014 Nov 1;58:125–31. Available from: https://doi.org/10.1016/j.icheatmasstransfer.2014.08.033

23. Alawi OA, Sidik NAC. The effect of temperature and particles concentration on the determination of thermo and physical properties of SWCNT-nanorefrigerant. International Communications in Heat and Mass Transfer [Internet]. 2015 Oct 1;67:8–13. Available from: https://doi.org/10.1016/j.icheatmasstransfer.2015.06.014

24. Al-Hajaj Z, Bayomy AM, Saghir MZ. A comparative study on best configuration for heat enhancement using nanofluid. International Journal of Thermofluids [Internet]. 2020 Nov 1;7–8:100041. Available from: https://doi.org/10.1016/j.ijft.2020.100041

25. Plant, Saghir. Numerical and experimental investigation of high concentration aqueous alumina nanofluids in a two and three channel heat exchanger. International Journal of Thermofluids. 2021 Feb;9(100055).

26. Avsec J, Marčič M. The calculation of equilibrium and nonequilibrium thermophysical properties [Internet]. 35th AIAA Thermophysics Conference. 2001. Available from: https://doi.org/10.2514/6.2001-2766

27. Avsec J, Marčič M. The calculation of the thermophysical properties for pure refrigerants and their mixtures [Internet]. 33rd Thermophysics Conference. 1999. Available from: https://doi.org/10.2514/6.1999-3676

28. Yılmaz F, Özdemir AF, Şahin AŞ, Selbaş R. Prediction of thermodynamic and thermophysical properties of carbon dioxide. Journal of Thermophysics and Heat Transfer [Internet]. 2014 Jul 1;28(3): 491–8. Available from: https://doi.org/10.2514/1.t4042

29. Wang KJ, Ding GL, Jiang WT. Development of nanorefrigerant and its rudiment property. In 8th International Symposium on Fluid Control, Measurement and Visualization. Chengdu. China: China Aerodynamics Research Society 2005 Aug.

30. Bi S, Guo K, Liu Z, Wu J. Performance of a domestic refrigerator using TiO₂-R600a nano-refrigerant as working fluid. Energy Conversion and Management [Internet]. 2011 Jan 1;52(1):733–7. Available from: https://doi.org/10.1016/j.enconman.2010.07.052

31. Wang, Hao, Xie, Li. A refrigerating system using HFC134A and mineral lubricant appended with N-TiO₂ (R) as working fluids. In: Heating, ventilating and air conditioning. ISHVAC 2003. Tsinghua University Press. 2003.

32. Wang, Shiromoto, Mizogami. Experiment study on the effect of nanoscale particle on the condensation process. In: Proceeding of the 22nd International Congress of Refrigeration. Beijing. China. 2007.

33. Bi S, Song L, Zhang L. Application of nanoparticles in domestic refrigerators. Applied Thermal Engineering [Internet]. 2008 Oct 1;28(14–15):1834–43. Available from: https://doi.org/10.1016/j.applthermaleng.2007.11.018

Baiju S. Bibin, Panitapu Bhramara, Arkadiusz Mystkowski, Edison Gundabattini

Predictive Analysis on the Influence of Al₂O₃ and CuO Nanoparticles on the Thermal conductivity of R1234yf-Based Refrigerants

DOI 10.2478/ama-2024-0050

34. Alawi OA, Salih JM, Mallah AR. Thermo-physical properties effectiveness on the coefficient of performance of Al₂O₃/R141b nano-refrigerant. International Communications in Heat and Mass Transfer [Internet]. 2019 Apr 1;103:54–61. Available from: https://doi.org/10.1016/j.icheatmasstransfer.2019.02.011

35. Yu W, Choi SJ. The role of interfacial layers in the enhanced thermal conductivity of nanofluids: a renovated Maxwell model. Journal of Nanoparticle Research [Internet]. 2003 Apr 1;5(1/2):167–71. Available from: https://doi.org/10.1023/a:1024438603801

36. Sitprasert C, Dechaumphai P, Juntasaro V. A thermal conductivity model for nanofluids including effect of the temperature-dependent interfacial layer. Journal of Nanoparticle Research [Internet]. 2008 Nov 4;11(6):1465–76. Available from: https://doi.org/10.1007/s11051-008-9535-4

37. Leong KC, Yang C, Murshed SMS. A model for the thermal conductivity of nanofluids – the effect of interfacial layer. Journal of Nanoparticle Research [Internet]. 2006 Apr 1;8(2):245–54. Available from: https://doi.org/10.1007/s11051-005-9018-9

38. Patil MS, Kim SC, Seo JH, Lee M. Review of the Thermo-Physical Properties and Performance Characteristics of a refrigeration system using Refrigerant-Based Nanofluids. Energies [Internet]. 2015 Dec 31;9(1):22. Available from: https://doi.org/10.3390/en9010022

39. Zawawi NNM, Azmi WH, Redhwan A a. M, Sharif MZ, Sharma KV. Thermo-physical properties of Al₂O₃-SiO₂/PAG composite nanolubricant for refrigeration system. International Journal of Refrigeration [Internet]. 2017 Aug 1;80:1–10. Available from: https://doi.org/10.1016/j.ijrefrig.2017.04.024

40. Yang L, Hu Y. Toward TIO2 Nanofluids—Part 2: Applications and Challenges. Nanoscale Research Letters [Internet]. 2017 Jul 6;12(1). Available from: https://doi.org/10.1186/s11671-017-2185-7

41. Mohammed AHSMMMS Karam Hashim. Energy observation technique for vapour absorption using nano fluid refrigeration [Internet]. 2020. Available from: http://sersc.org/journals/index.php/IJAST/article/view/22608

42. Wang X, Amrane K, Johnson P. Low Global Warming Potential (GWP) Alternative Refrigerants Evaluation Program (Low-GWP AREP) [Internet]. Purdue e-Pubs. Available from: http://docs.lib.purdue.edu/iracc/1222

43. Chandrasekar M, Suresh S, Bose AC. Experimental investigations and theoretical determination of thermal conductivity and viscosity of Al₂O₃/water nanofluid. Experimental Thermal and Fluid Science [Internet]. 2010 Feb 1;34(2):210–6. Available from: https://doi.org/10.1016/j.expthermflusci.2009.10.022

44. Akhavan-Behabadi MA, Sadoughi M, Darzi M, Fakoor-Pakdaman M. Experimental study on heat transfer characteristics of R600a/POE/CuO nano-refrigerant flow condensation. Experimental Thermal and Fluid Science [Internet]. 2015 Sep 1;66:46–52. Available from: https://doi.org/10.1016/j.expthermflusci.2015.02.027

45. Wang CC. An overview for the heat transfer performance of HFO-1234yf. Renewable & Sustainable Energy Reviews [Internet]. 2013 Mar 1;19:444–53. Available from: https://doi.org/10.1016/j.rser.2012.11.049

46. Maxwell JC. A treatise on electricity and magnetism. Journal of the Franklin Institute [Internet]. 1954 Dec 1;258(6):534. Available from: https://doi.org/10.1016/0016-0032(54)90053-8

47. Hamilton R, Crosser OK. Thermal conductivity of heterogeneous Two-Component systems. Industrial & Engineering Chemistry Fundamentals [Internet]. 1962 Aug 1;1(3):187–91. Available from: https://doi.org/10.1021/i160003a005

48. Sharif MZ, Azmi WH, Redhwan A a. M, Mamat R. Investigation of thermal conductivity and viscosity of Al₂O₃/PAG nanolubricant for application in automotive air conditioning system. International Journal of Refrigeration [Internet]. 2016 Oct 1;70:93–102. Available from: https://doi.org/10.1016/j.ijrefrig.2016.06.025

49. Stacy SC, Zhang X, Pantoya ML, Weeks BL. The effects of density on thermal conductivity and absorption coefficient for consolidated aluminum nanoparticles. International Journal of Heat and Mass Transfer [Internet]. 2014 Jun 1;73:595–9. Available from: https://doi.org/10.1016/j.ijheatmasstransfer.2014.02.050

50. Jwo, Jeng, Chang, Teng. Experimental study on thermal conductivity of lubricant containing nanoparticles. Reviews on Advanced Materials Science. 2008;18:660–6.

Baiju S. Bibin: https://orcid.org/0000-0002-9142-5068

Panitapu Bhramara: https://orcid.org/0000-0003-0756-5488

Arkadiusz Mystkowski: https://orcid.org/0000-0002-5742-7609

Edison Gundabattini: https://orcid.org/0000-0003-4217-2321

AN APPLICATION OF THE SCALED–PIXEL–COUNTING PROTOCOL
TO QUANTIFY THE RADIOLOGICAL FEATURES OF ANATOMICAL STRUCTURES
OF THE NORMAL TARSAL JOINT IN HORSES

Marta BOROWSKA*⬤, **Bernard TUREK****⬤, **Paweł LIPOWICZ***⬤, **Tomasz JASIŃSKI****⬤,
Katarzyna SKIERBISZEWSKA**⬤, **Małgorzata DOMINO****⬤

*Institute of Biomedical Engineering, Faculty of Mechanical Engineering, Bialystok University of Technology,
ul. Wiejska 45C, 15-351 Bialystok, Poland
**Department of Large Animal Diseases and Clinic, Institute of Veterinary Medicine, Warsaw University of Life Sciences,
ul. Nowoursynowska 166, 02-787 Warsaw, Poland

m.borowska@pb.edu.pl, bernard_turek@sggw.edu.pl, p.lipowicz@pb.edu.pl, jasinski_tomasz@sggw.edu.pl,
katarzyna_skierbiszewska@sggw.edu.pl, malgorzata_domino@sggw.edu.pl

Abstract: Background: As high-performance human and equine athletes train and compete at the highest level of effort, the prevention of high-performance-cased diseases, such as osteoarthritis (OA), requires knowledge of the anatomy and physiology of the subjected bones. Objective: Implementation of the scaled–pixel–counting protocol to quantify the radiological features of anatomical structures of the normal equine tarsal joint as the first step in the prevention of the tarsal joints OA in high-performance sport horses. Methods: A radiographic examination was performed in six cadaverous equine pelvic limbs. The dorso–plantar projection of the tarsal joint was performed using density standard (DS) attached to the radiographic cassette, standard X-ray equipment and standard diagnostic imaging protocol. On each of the radiographs, pixel brightness (PB) was extracted for each of the 10 steps (S1–S10) of DS. On each of the radiographs, seven regions of interest (ROIs) were annotated representing four bones (II tarsal bone [II TB], III tarsal bone [III TB], IV tarsal bone [IV TB] and central tarsal bone [CTB]) and three joints (proximal intertarsal joint [PIJ], distal intertarsal joint [DIJ] and tarsometatarsal joint [TMJ]), respectively. For each ROI, the percentage (%) of number of pixels (NP) from each range was calculated. Results: The % of NP was lower in bones than in joint spaces for S1–S6 and was higher in bones than in joint spaces for S8–S10. The % of NP was higher in PIJ than TMJ for S1 and higher in PIJ than DIJ for S4. No differences were found between consecutive bones for all examined steps of DS. Conclusions: An application of the scaled–pixel–counting protocol provides the quantitative radiological features of normal bone and joint structures of the tarsal joint in horses, making possible differentiation of the lucency of joint space and opacity of bone structure.

Key words: diagnostic imaging, radiographs, radiodensity, osteoarthritis, prevention, horse

1. BACKGROUND

Nowadays, the sport uses of horses in show jumping, eventing, dressage, driving, reining, vaulting or endurance are characterised by the horse being perceived as an athlete in its own right [1]. In high-performance sports, both human and equine athletes train and compete at the highest level of effort. However, one should keep in mind that the level is close to the limit of injury, as exceeding their physiological limits allows for the achievement of the best performance results [2]. The powerhouse of athletic movement is the muscle [3], whereas the skeletal system withstands the stressors placed on bones and joints during high-performance exercises [2]. Therefore, the cultivation of a deep knowledge of the anatomy and physiology of the equine skeletal system is the first step in the prevention of high-performance-cased diseases.

One of the important causes of lameness, and thus the exclusion of a sport horse from high-performance, is osteoarthritis (OA) [4]. OA is a joint disease that results from joint cartilage and underlying bone destruction. Cartilage destruction and subchondral bone sclerosis are central and irreversible steps in OA, although the entire joint is affected [2]. As OA occurs more commonly in the overload joint, repeated stress on bones and ligaments has been suggested to be important in the development of OA. Show jumping horses are often affected by the metacarpophalangeal joint OA due to forces acting during landing after a jump, whereas dressage and reining horses often suffer from the tarsal joints OA due to strong involvement of the hindquarters during exercise [5]. The OA is diagnosed based on the clinical symptoms such as lameness and stiffness, which in many cases reduce the working capacity of the horse [6–8]. The clinical symptoms are relevant to the low-grade chronic inflammation typical for OA. Therefore, the diagnostic strategy considers the basic clinical examination and the detailed orthopaedic examination supported by the flexion tests, response to local anaesthesia and diagnostic imaging of the affected joint [6, 7]. While radiographic changes in the tarsal joint have variable correlation with lameness [6, 8–10], radiographs are frequently used to identify structural changes in the clinically suspected joint [6]. Within the structural changes identified by radiographs in affected joints, joint space narrowing, lysis of subchondral bone, sclerosis of subchondral bone, mineralisation front defects and osteophyte/enthesophyte development should be considered [6, 7, 10, 11].

One of the equine limb joints most commonly diagnosed with OA is the tarsal joint. In the experience from clinical cases, there

is evidence indicating that the OA in the tarsal joint is a slowly progressing disease where clinical symptoms may precede radiographic signs of degenerative joint disease [8]. Spontaneous radiographic signs of OA in the tarsal joint have been reported in 30%–60% of the horse population, with a higher prevalence in mature horses [6, 8–10]. The OA in the tarsal joint has been suggested to be most frequent in older horses. Type and load of work, limb conformation and developmental abnormalities have been proposed as factors predisposing to the OA in the tarsal joint [8]. The OA in the distal intertarsal joint (DIJ), tarsometatarsal joint (TMJ) and less proximal intertarsal joint (PIJ) occurs commonly in sport horses; thus, structural changes' identification may help in OA prognosis determination and the monitoring of disease [7, 10, 11]. The onset of the OA in the tarsal joint is often followed by a period during which degenerative changes worsen and clinical symptoms of varying degrees are present continuously or intermittently. However, the progression of the disease as evaluated by radiological examination remains to be investigated as not all individuals reach the final stage of ankylosis [8]. The monitoring of the radiographic signs is particularly important when the outcome of both spontaneous ankylosis [8] and surgical arthrodesis [12, 13] of tarsal joints is assessed.

One may observe that OA structural changes are quantified using the radiographic rating systems, which are commonly used in humans [14, 15], occasionally used in dogs [16] and increasingly used in equines [12, 13]. These systems attempt to assign an aggregate score based on the appearance and severity of recognised disease features [12–17]. Besides the adaptation of the Delphi technique that considers the expert consultation process [18], none of the developed rating systems allow radiographic signs grading in detail for each tarsal bone (TB) and joint independently, and therefore none of the existing radiographic rating systems is unlikely to be useful in the clinic [18]. Labens et al. [18] concluded that they do not advocate the clinical use of the radiographic rating scale developed in the recent literature unless each user assesses his repeatability first. Moreover, they concluded that the knowledge of radiographic features selected by the experts according to their diagnostic value for the tarsal joint OA may be of considerable importance to the development of other rating scales. They suggested that the new rating systems may include a scale based on the recognition of the selected radiographic features and the assignment of a numeric value. To fill the gap in the existing state of quantification of the radiological signs of the tarsal joint OA, the objective of this study is to implement the scaled–pixel–counting protocol to quantify the radiological features of the anatomical structures of the equine normal tarsal joint as the first step in the prevention of the tarsal joints OA in the high-performance sport horses.

2. METHODS

The study was conducted on six cadaverous equine pelvic limbs with no clinical symptoms and radiological signs of the tarsal joint OA. Limbs were collected during post-mortem at the commercial slaughterhouse in Poland, which does not fall under the legislation for the protection of animals used for scientific purposes, national decree-law (Dz. U. 2015 poz. 266) and the European Union law (2010-63-EU directive). Thus, no Ethical Committee's permission was needed for sample retrieval after slaughter. The absence of clinical symptoms of OA was determined as the initial

inclusion criterion; thus, the orthopaedic examination, limited to slaughter specimen, was performed considering swelling, joint deformities and impaired function of the affected joint manifested by decreased motion range. The orthopaedic examination was performed by palpation in lateral and anterior-posterior recumbency. During the orthopaedic examination, the lack (0) or presence (1) of swelling, joint deformities and decreased motion range was noted. The initial exclusion criterion was the presence of at least one clinical symptom of tarsal joint OA. Based on the initial exclusion criterion, no limb was excluded.

Then, the radiographic examination of the tarsal joint was performed using an X-ray tube (Orange 9020HF, Ecoray Co., Seoul, South Korea), a radiographic cassette (Saturn 8000, Vievorks Co., Seoul, South Korea) and a portable computer (HP Inc. UK Ltd., Reading, UK). The X-ray tube settings were 1.25 mAs and 60 kV, and the distance between the X-ray tube and radiographic cassette was 80 cm (Fig. 1A). The density standard (DS) was attached to the radiographic cassette while each radiograph was taken. The DS was positioned perpendicular to the surface of the cassette so that the long axis of the DS was parallel to the long axis and the thick end was caudally of the cassette (Fig. 1B, C). Concomitant with the cassette and DS being subject to such preparation, the dorso–plantar projection of the tarsal joint, with the centre of the X-ray beam positioned on the central tarsal bone (CTB), was performed. As Labens et al. [18] did not find significant differences between the equine tarsal joint ratings for each radiographic projection, considering the lateromedial view, dorso–plantar view and planto–dorsal view, the only one dorso–plantar projection of the tarsal joint in the current preliminary study was used. The radiographs were acquired as .jpg files.
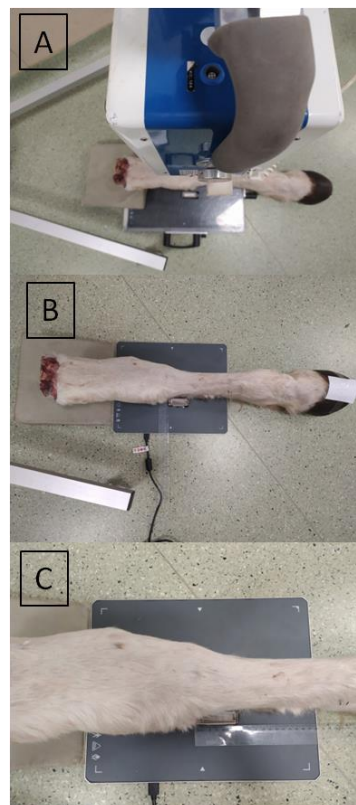
Fig. 1. The position of the X-ray tube on a tripod in relation to the cadaverous pelvic limb and radiographic cassette (A). The position of the density standard (DS) in relation to the long (B) and short (C) edges of the radiographic cassette
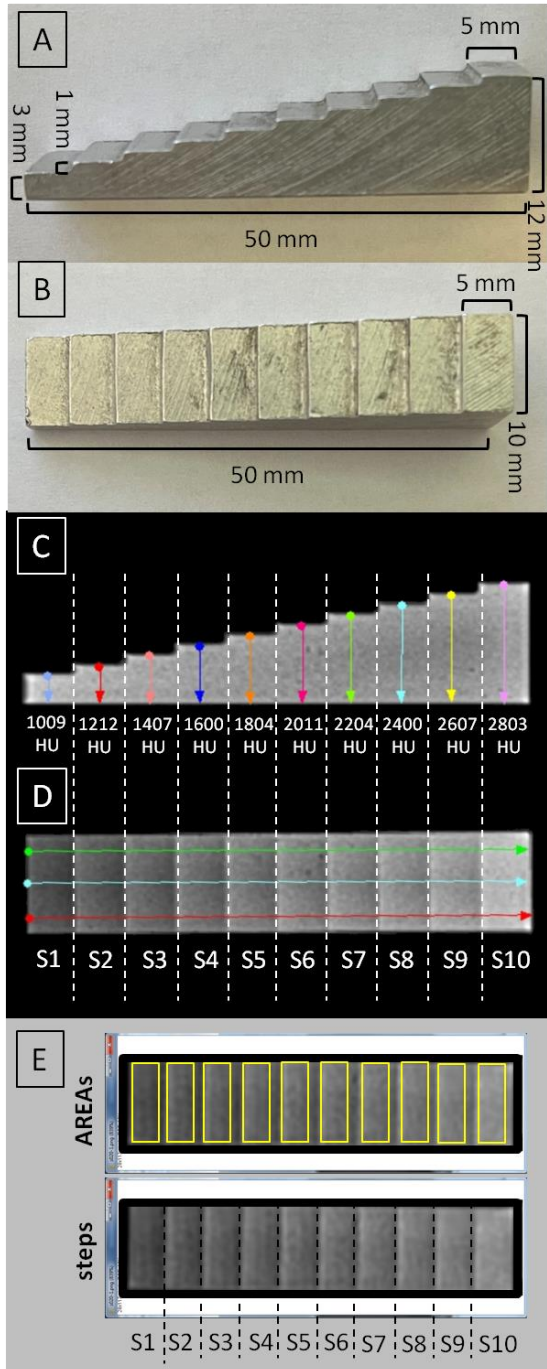
Fig. 2. The density standard (DS) in visible light (A, B) and under X-ray beam (C, D, E) – side view (A, C) and top view (B, D, E). Views concerning 10 steps (S1–S10) of an irregular cuboid. Views with marked dimensions (A, B), attenuation of the X-ray beam (C, D) and rectangular regions of interest representing S1–S10 (AREAs) (E)

Each radiograph was visually assessed for the presence of radiological signs of the tarsal joint OA. The absence of radiological signs of OA was determined as the secondary inclusion criterion; thus, the radiograms evaluation considered joint space narrowing, lysis of subchondral bone, sclerosis of subchondral bone, mineralisation front defects and osteophyte/enthesophyte [6, 7, 10, 11]. The radiograph evaluation was visually assessed. The lack (0) or presence (1) of radiological signs of joint space narrowing, lysis of subchondral bone, sclerosis of subchondral bone,

mineralisation front defects and osteophyte/enthesophyte was noted. The secondary exclusion criterion was the presence of at least one radiological sign of tarsal joint OA. Based on the secondary exclusion criterion, no limb was excluded.

The DS is an irregular cuboid with 10 steps (S1–S10) made of aluminium (Al; 95.20–98.88 Mass%; 92.71–98.92 Atom%). The DS dimensions are 55 mm in length on the basis, 12 mm high in the highest place, 3 mm high in the lowest place and 10 mm wide (Fig. 2A, B). The attenuation of the X-ray beam passing through the DS was reported in the form of the Hounsfield unit (HU) measured for each of S1–S10 as follows: S1 = 1009 HU, S2 = 1212 HU, S3 = 1407 HU, S4 = 1600 HU, S5 = 1804 HU, S6 = 2011HU, S7 = 2204 HU, S8 = 2400 HU, S9 = 2607 HU and S10 = 2803 HU (Fig. 2C, D).

On each radiograph (Fig. 3A), seven polymorphic regions of interest (ROIs) were manually annotated using ImageJ software (version 1.46r, Wayne Rasband, Bethesda, MD, USA). ROIs represented seven anatomical structures of the normal tarsal joint including four bones (II tarsal bone [II TB], III tarsal bone [III TB], IV tarsal bone [IV TB] and central tarsal bone [CTB]) (Fig. 3B–E) and three joints (proximal intertarsal joint [PIJ], distal intertarsal joint [DIJ] and tarsometatarsal joint [TMJ]) (Fig. 3F–H).
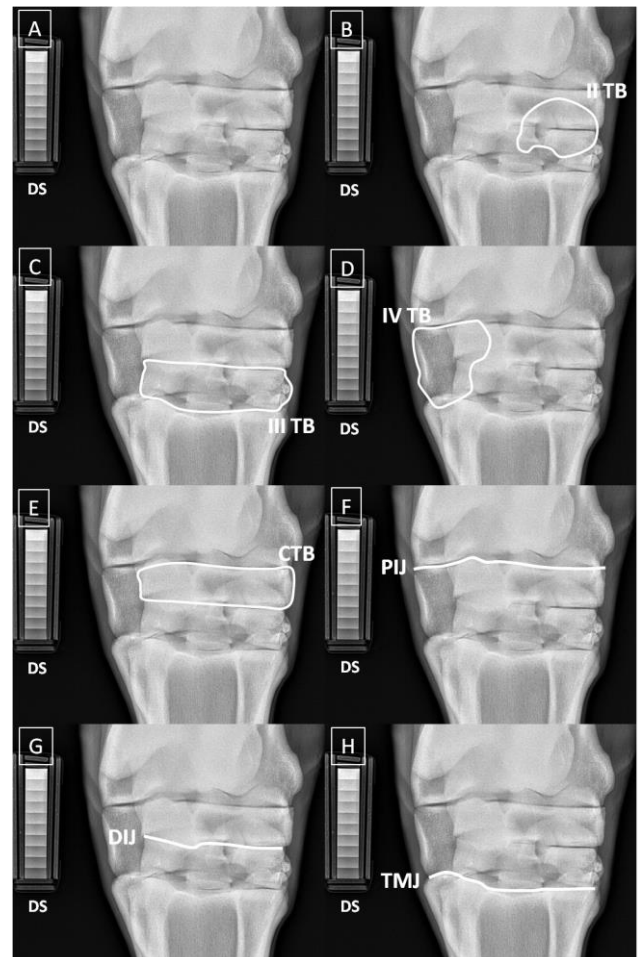
Fig. 3. The radiograph of the normal tarsal joint in horses without (A) and with (B–H) annotated regions of interest (ROIs). ROIs represented II tarsal bone (II TB) (B), III tarsal bone (III TB) (C), IV tarsal bone (IV TB) (D), central tarsal bone (CTB) (E), proximal intertarsal joint (PIJ) (F), distal intertarsal joint (DIJ) (G) and tarsometatarsal joint (TMJ) (H). The density standard (DS) is visible in each image

sciendo

Marta Borowska, Bernard Turek, Paweł Lipowicz, Tomasz Jasiński, Katarzyna Skierbiszewska, Małgorzata Domino   DOI 10.2478/ama-2024-0051
An Application of the Scaled–Pixel–Counting Protocol to Quantify the Radiological Features of Anatomical Structures of the Normal Tarsal Joint in Horses

Additionally, 10 rectangular regions of interest representing S1–S10 (AREAs) were manually annotated using ImageJ software (Fig. 2E). The AREAs represented 10 steps of DS with various degrees of X-ray beam attenuation. Each AREAs returned the values of Pixel Brightness (PB) <0; 255> and determined the ranges of PB change in each AREAs. AREAs that attenuated a small amount of the X-ray beam (e.g. soft tissues) were dark and represented an area of lucency. AREAs that attenuate a large amount of the X-ray beam (e.g. bones) were bright and represented an area of opacity. The algorithm of the scaled pixel counting protocol was implemented in Python:

```
M = image*mask
start = 1
for i in range(R):
    if i=R-1: end = PBMax
    else: end = AREA[i]+(Area[i+1] AR-
EA[i])/2)
    NP = len(M[(M>=start)&(M<=end)])
    %NP =
    start = end+1
```

where the notations are as follows:
− M – segmented image
− image – input image
− mask – mask image
− start – beginning of the interval
− end – ending of the interval
− R – number of S-labelled data series
− Area – vector of intervals S-labelled data series
− NP – pixel counting result
− % NP – normalisation of NP
  The features of the algorithm are:
− the algorithm is designed for grayscale images;
− the result depends on the image intensity of the ROIs;
− in each image, AREAs were annotated to compare the beam attenuation of individual ROIs; and
− the result is independent of artefacts generated during registration, and thus each measurement is independent.

The number of pixels (NP) values was grouped for all bone's and joint's anatomical structures, assigned to 10 S-labelled data series (S1, S2, S3, S4, S5, S6, S7, S8, S9 and S10) and tested independently for univariate distributions using the Kolmogorov–Smirnov normality test. The S-labelled NP data series were compared between all bone's and joint's anatomical structures, using the Mann–Whitney test. The NP values were grouped for the consecutive bone's and joint's anatomical structures, assigned to 10 S-labelled data series and tested independently for univariate distributions using the Kolmogorov–Smirnov normality test. The S-labelled NP data series were compared between bone's (II TB vs III TB vs IV TB vs CTB) and joint's (PIJ vs DIJvs TMJ) anatomical structures separately, using the Kruskal–Wallis test followed by the Dunn's multiple comparisons test. For each data set, at least one data series was non-Gaussian distributed. The alpha value was established as α = 0.05. All statistical analyses were performed using Graph Pad Prism 6 software (GraphPad Software Inc., Avenida De La Playa La Jolla, CA, USA).

3. RESULTS

The % of NP was lower in bones than in joint spaces for S1 (Fig. 4A), S2, (Fig. 4B), S3 (Fig. 4C), S4 (Fig. 4D), S5 (Fig. 4E) and S6 (Fig. 4F), which made it possible to indicate the lucency of joint space. Moreover, the % of NP was higher in bones than in joint spaces for S8 (Fig. 4H), S9 (Fig. 4I) and S10 (Fig. 4J), which made it possible to indicate the opacity of bone structure. No difference was found between bones and joints for S7 (Fig. 4G).

Similarly, no differences were found between consecutive bones for all examined steps of DS (Fig. 5). However, the % of NP was higher in PIJ than TMJ for S1 (Fig. 6A) and higher in PIJ than DIJ for S4 (Fig. 6D), which indicate on the potential possibility of discrimination of narrow and wider tarsal joints not influenced by the type of normal TB.
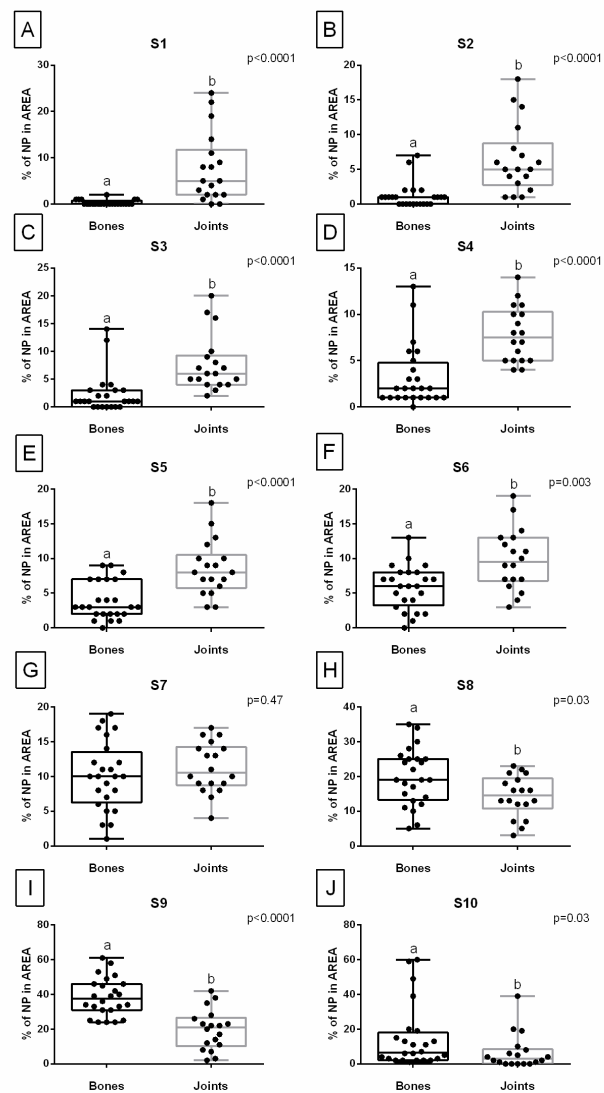
Fig. 4. The comparison of percentage (%) of the number of pixels (NP) in each AREA between all bones and joints anatomical structures. Data displayed separately for consecutive 10 steps (A, S1; B, S2; C, S3; D, S4; E, S5; F, S6; G, S7; H, S8; I, S9; J, S10) of density standard (DS). Data on box plots are represented by lower quartile, median and upper quartile, whereas whiskers represent minimum and maximum values. The single realisations are represented by dots. Lowercase letters indicate differences between groups for p < 0.05
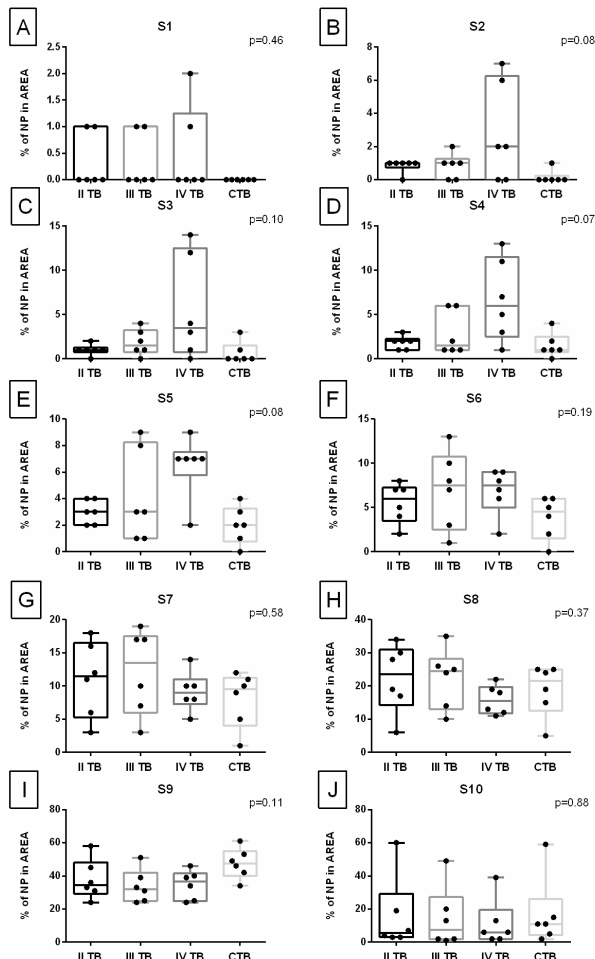
Fig. 5. The comparison of percentage (%) of the number of pixels (NP) in each AREA between consecutive bones' anatomical structures: II tarsal bone (II TB); III tarsal bone (III TB); IV tarsal bone (IV TB) and central tarsal bone (CTB). Data displayed separately for consecutive 10 steps (A, S1; B, S2; C, S3; D, S4; E, S5; F, S6; G, S7; H, S8; I, S9; J, S10) of density standard (DS). Data on box plots are represented by lower quartile, median and upper quartile, whereas whiskers represent minimum and maximum values. The single realisations are represented by dots. Lowercase letters indicate differences between groups for p < 0.05

4. DISCUSSION

Highlighting the most relevant results, one may observe that the application of this scaled–pixel–counting protocol allows quantification of the radiological features of joint spaces and bones and makes it possible to indicate the lucency of joint space and the opacity of bone structure. Thus, we confirm that the quantification of the radiological features of the anatomical structures of the equine normal tarsal joint is feasible. Knowing that the joint space narrowing, lysis of subchondral bone, sclerosis of subchondral bone, mineralisation front defects and osteophyte/enthesophyte development in OA-affected TB are the most important radiological signs [6, 7, 10, 11], the quantification of lucency and opacity may be considered as the first step for the tarsal joints rating. One may observe that on the radiographs, joint space narrowing is recognised as thin and uneven lucency between the adjacent cortical bones [6–9].
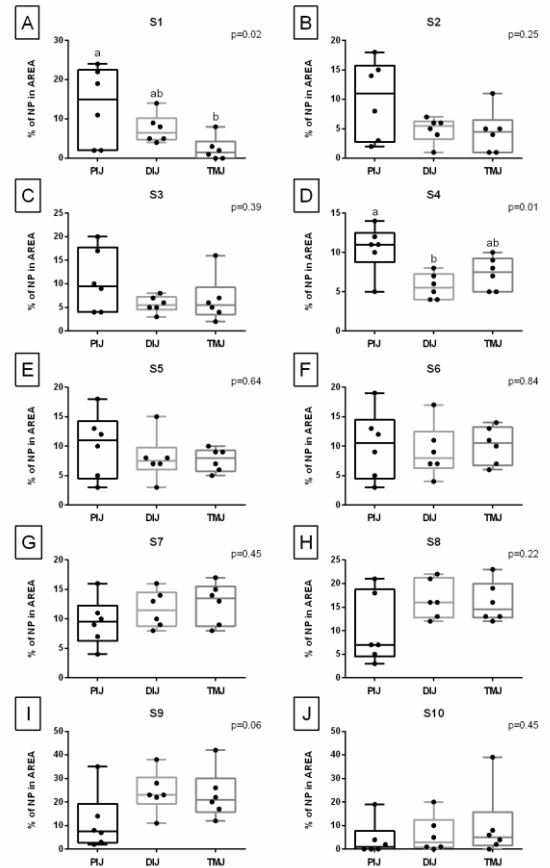
Fig. 6. The comparison of percentage (%) of the number of pixels (NP) in each AREA between consecutive joints' anatomical structures: proximal intertarsal joint (PIJ); distal intertarsal joint (DIJ) and tarsometatarsal joint (TMJ). Data displayed separately for consecutive 10 steps (A, S1; B, S2; C, S3; D, S4; E, S5; F, S6; G, S7; H, S8; I, S9; J, S10) of density standard (DS). Data on box plots are represented by lower quartile, median and upper quartile, whereas whiskers represent minimum and maximum values. The single realisations are represented by dots. Lowercase letters indicate differences between groups for p < 0.05

Therefore, in the case of OA when the lucency between bones is less than it should be, one can expect an increase in %NP of the bright steps (S7–S8) and a decrease in %NP of the dark steps (S1–S3) at the anatomical location of the joint spaces (PIJ, DIJ and/or TMJ ROIs).

This hypothesis needs to be tested in further studies using the OA-affected tarsal joints and comparing the results of the OA group with the results for normal joints presented in the current research. On the radiographs, lysis of the subchondral bone is recognised as the area of increased lucency within the cortical and subcortical bone [6–9]. Therefore, in the case of OA when the lucency within bones' anatomical locations is more than it should be, one can expect an increase in %NP of the dark steps (S1–S3) and a decrease in %NP of the bright steps (S8–S10) at the anatomical location of the bones (II TB, III TB, IV TB and/or CTB ROIs). On the radiographs, sclerosis of the subchondral bone is recognised as the area of increased opacity within the cortical and subcortical bones [6–9]. Therefore, in the case of OA when the opacity within bones' anatomical locations is more than it should be, one can expect an increase in %NP of the bright steps (S8–S10) and a decrease in %NP of the dark steps (S1–S3) at the

anatomical location of the bones (II TB, III TB, IV TB and/or CTB ROIs). On the radiographs, mineralisation front defects and osteophyte/enthesophyte development are recognised as the areas of increased opacity between the adjacent cortical bones [6–9]. Therefore, in the case of OA when the opacity between bones is more than it should be, one can expect an increase in %NP of the bright steps (S9–S10) and a decrease in %NP of the dark steps (S1–S3) at the anatomical location of the joint spaces (PIJ, DIJ and/or TMJ ROIs). The differentiation between the rating of joint space narrowing and mineralisation/osteophyte/enthesophyte should be assessed in detail since both radiological signs concern an increase in %NP of the bright steps and a decrease in %NP of the dark steps at the joint spaces. We suspect that estimation of bone density with HU will be helpful, with the advantage of %NP of S9–S10 in the case of new bone formations and %NP of S7–S8 in the case of narrowing. However, all these hypotheses need to be tested in further studies using the OA-affected tarsal joints and comparing the results of the OA group with the results for normal joints presented in the current research.

In the recent research, a radiographic rating scale, developed through the Delphi process, was employed for assessing OA radiographic signs in the tarsal joint. The rating was determined based on the total extent of the affected distal tarsal joint surface, leading to a qualitative verbal descriptive assessment of the

The four-point scale was used for the assessment of OA in the entire distal tarsus concerning none OA, mild OA, moderate OA and severe OA [18]. In the current research, a scaled–pixel– counting protocol was applied to all anatomical structures of the tarsal joint, resulting in a quantitative numerical rating of the normal joint. As no OA-affected joints were examined, the 4-point scale for the assessment of OA in the entire tarsal joint was not applied. Labens et al. [18] stated that if used clinically, this generalised assessment may not allow the progression of individual radiographic abnormalities to be monitored. Contrarily, the current scaled–pixel–counting protocol will be potentially able to distinguish the individual radiographic abnormalities since each bone and joint are quantified separately. However, all these hypotheses need to be tested in further studies using the OA-affected tarsal joints and comparing the results of the OA group with the results for normal joints presented in the current research.

In recent research, the basis of the radiographic rating scale was a visual analogue rating scale in which assessors indicated the severity of each radiographic feature as they perceived it [18]. In the current research, the basis of the radiographic rating protocol was the use of the DS and the automatic algorithm. In this protocol, assessors did not indicate the severity of each radiographic feature but only annotated the ROIs within the algorithm quantifying the relative values. The two-step relativisation, concerning the use of individual X-ray images of DS and the normalisation of NP values, makes the final results for each image as individual as possible. However, validation of the method on a larger set of valid and OA-affected data is necessary to continue the work that has been commenced in the current study.

5. CONCLUSION

An application of the scaled–pixel–counting protocol provides the quantitative radiological features of normal bone and joint structures of the tarsal joint in horses, making possible differentiation of the lucency of joint space and opacity of bone structure.

Moreover, the scaled–pixel–counting protocol allows for the discrimination of narrow and wider tarsal joints not influenced by the type of normal TB.

REFERENCES

1. Reed SA. Horses as athletes: the road to success. Animal Frontiers. 2022;12(3): 3-4.
   https://doi.org/10.1093/af/vfac024
2. Baccarin RYA, Seidel SRT, Michelacci YM, Tokawa PKA, Oliveira TM. Osteoarthritis: a common disease that should be avoided in the athletic horse's life. Animal Frontiers. 2022;12(3):25-36.
   https://doi.org/10.1093/af/vfac026
3. Latham CM, Guy CP, Wesolowski LT, White-Springer SH. Fueling equine performance: importance of mitochondrial phenotype in equine athletes. Animal Frontiers. 2022;12(3):6–14.
   https://doi.org/10.1093/af/vfac023
4. Goodrich LR, Nixon AJ. Medical treatment of osteoarthritis in the horse–a review. The Veterinary Journal. 2006;171(1): 51-69.
   https://doi.org/10.1016/j.tvjl.2004.07.008
5. De Sousa NR, Luna SP, Pizzigatti D, Martins MT, Possebon FS, Aguiar AC. Relation between type and local of orthopedic injuries with physical activity in horses. Ciência Rural. 2017;47: 1–7.
   https://doi.org/10.1590/0103-8478cr20151218
6. Coppelman EB, David FH, Tóth F, Ernst NS, Trumble TN. The association between collagen and bone biomarkers and radiographic osteoarthritis in the distal tarsal joints of horses. Equine Veterinary Journal. 2020;52(3): 391-398.
   https://doi.org/10.1111/evj.13187
7. Baxter GM, Southwood LL, Dechant JE. Diagnosis of distal tarsal osteoarthritis in horses. Compendium On Continuing Education For The Practising Veterinarian-North American Edition. 2003;25(2): 138–147.
8. Eksell P, Axelsson M, Broström H, Ronéus B, Häggström J, Carlsten J. Prevalence and risk factors of bone spavin in Icelandic horses in Sweden: a radiographic field study. Acta Veterinaria Scandinavica. 1998;39: 339-348.
   https://doi.org/10.1186/BF03547782
9. Björnsdóttir S, Axelsson M, Eksell P, Sigurdsson H, Carlsten J. Radiographic and clinical survey of degenerative joint disease in the distal tarsal joints in Icelandic horses. Equine Veterinary Journal (2000);32(3): 268-272.
   https://doi.org/10.2746/042516400776563590
10. Ley CJ, Björnsdóttir S, Ekman S, Boyde A, Hansson K. Detection of early osteoarthritis in the centrodistal joints of Icelandic horses: Evaluation of radiography and low-field magnetic resonance imaging. Equine Veterinary Journal. 2016;48(1): 57-64.
    https://doi.org/10.1111/evj.12370
11. Fairburn A, Dyson S, Murray R. Clinical significance of osseous spurs on the dorsoproximal aspect of the third metatarsal bone. Equine Veterinary Journal. 2010;42: 591–599.
    https://doi.org/10.1111/j.2042-3306.2010.00097.x
12. Dechant JE, Baxter GM, Southwood LL, et al. Use of a three-drilltract technique for arthrodesis of the distal tarsal joints in horses with distal tarsal osteoarthritis: 54 cases (1990–1999). Journal of the American Veterinary Medical Association. 2003;223(12): 1800–1805.
    https://doi.org/10.2460/javma.2003.223.1800
13. Zubrod CJ, Schneider RK, Hague BA, Ragle CA, Gavin PR, Kawcak CE. Comparison of three methods for arthrodesis of the distal intertarsal and tarsometatarsal joints in horses. Veterinary Surgery. 2005;34: 372–382.
    https://doi.org/10.1111/j.1532-950X.2005.00057.x
14. Gunther KP, Sun Y. Reliability of radiographic assessment in hip and knee osteoarthritis. Osteoarthritis Cartilage. 1999;7: 239–246.
    https://doi.org/10.1053/joca.1998.0152

15. Kessler S, Dieppe P, Fuchs J, Sturmer T, Gunther KP. Assessing the prevalence of hand osteoarthritis in epidemiological studies. The reliability of a radiological hand scale. Annals of the Rheumatic Diseases. 2000;59: 289–292. doi: 10.1136/ard.59.4.289
Available from: https://ard.bmj.com/content/59/4/289

16. Innes JF, Costello M, Barr FJ, Rudorf H, Barr ARS. Radiographic progression of osteoarthritis of the canine stifle joint: a prospective study. Veterinary Radiology & Ultrasound. 2004;45(2): 143-148.
doi: 10.1111/j.1740-8261.2004.04024.x

17. Bjornsdottir S, Ekman S, Eksell P, Lord P. High detail radiography and histology of the centrodistal tarsal joint of Icelandic horses age 6 months to 6 years. Equine Veterinary Journal. 2004;36: 5–11.
https://doi.org/10.2746/0425164044864679

18. Labens R, Innocent GT, Voûte LC. Reliability of a quantitative rating scale for assessment of horses with distal tarsal osteoarthritis. Veterinary Radiology & Ultrasound. 2007;48(3): 204-211.
https://doi.org/10.1111/j.1740-8261.2007.00230.x

Marta Borowska: https://orcid.org/0000-0003-0148-9912

Bernard Turek: https://orcid.org/0000-0002-1065-5284

Paweł Lipowicz: https://orcid.org/0000-0003-3598-0945

Tomasz Jasiński: https://orcid.org/0000-0003-2906-9944

Katarzyna Skierbiszewska: https://orcid.org/0000-0002-1765-5340

Małgorzata Domino: https://orcid.org/0000-0001-9436-1074

COMPARING THE HEATING RATE OF THE PROXIMAL PHALANX OF THE FINGERS IN RHEUMATOID ARTHRITIS AND HEALTHY SUBJECTS

Mikhail IHNATOUSKI*⊚, Jolanta PAUK⊚, Kristina DAUNORAVICIENE***⊚, Jurgita ZIZIENE***⊚**

*Scientific and Research Department, Yanka Kupala State University of Grodno, Elizy Azeska 22, 230023 Grodno, Belarus
**Mechanical Engineering Department, Bialystok University of Technology, Wiejska 45C, 15-351 Bialystok, Poland
***Faculty, Department, Vilnius Gediminas Technical University, Sauletekio al. 11, LT-10223 Vilnius, Lithuania

mii_by@mail.ru, j.pauk@pb.edu.pl, kristina.daunoraviciene@vilniustech.lt, jusrgita.ziziene@vilniustech.lt

Abstract: Thermography is a non-invasive imaging technique that has been used for the assessment of rheumatoid arthritis (RA). The purpose of this research was to compare the heating rate of the proximal phalanx of the fingers and the whole palms in RA and that of healthy subjects. The study was conducted on 48 patients with high disease activity, hospitalised for RA, and 45 healthy subjects. The thermograms were taken with the FLIR camera E60bx. Subjects were instructed to immerse both hands up to the wrist in water thermostatically controlled at 0°C for 30 s. Then, the hands were pulled out of the water; the warm-up period was 180 s. Image pre-processing included: segmentation, extraction and anatomy identification. The mean value of the heating rate for whole palms and the proximal phalanx of the fingers in the RA group was lower than that in the control group ($p < 0.05$). This coincides with the uneven flow of the heat-transfer blood caused by the disease. However, the difference between the heating rates of the proximal phalanx of the fingers was greater than that of the entire hand. In addition, the proximal phalanx heating rates of the second, third and fourth fingers were higher than those of the outermost two fingers. The study may be used to develop clinical tools in the detection of abnormal heat signatures in the phalanx proximal of the fingers.

Key words: active thermography, rheumatoid arthritis, proximal phalanx, image processing

1. INTRODUCTION

Thermography is a non-invasive imaging technique that has been used for the assessment of rheumatoid arthritis (RA) in recent years. It is a promising tool for the early detection and monitoring of RA, as it can detect changes in temperature distribution and blood flow associated with inflammation in affected joints [1-4]. The state-of-the-art approach for thermography imaging in rheumatoid patients typically involves the use of high-resolution infrared cameras to capture thermal images of the affected joints. These images are then processed using advanced image-processing techniques to extract temperature and blood flow information [4-7]. Recently, research has focused on the development of machine learning algorithms that can automatically analyze thermography images and provide quantitative measurements of disease severity. These algorithms use deep learning techniques to extract features from the images and can accurately classify patients based on disease severity [8-9]. Some common methods include image segmentation [10] e.g. active contour models, superpixel-based methods, and deep learning algorithms to accurately identify and isolate the regions of the image corresponding to inflamed joints or regions of interest, to facilitate the quantification of disease activity and response to treatment. Some studies explored [11] various methods for feature extraction in thermal images of RA patients, including identifying hotspots, quantifying temperature changes and analysing blood flow patterns. Several researchers have written about the classification in

RA to prediction disease activity and treatment response [12-18]. The authors explored various methods for classification of RA disease activity based on thermography images, including using machine learning algorithms such as deep learning, support vector machines and decision trees. Studies [19-23] explored various methods for image fusion in RA, including combining different modalities of imaging such as thermal imaging, ultrasound and magnetic resonance imaging (MRI) in aim to improving the accuracy of diagnosis and assessment of RA by providing a more comprehensive and detailed view of the affected joints and surrounding tissues.

In healthy individuals, the skin temperature of the proximal phalanx of the fingers is typically regulated by factors such as environmental conditions and overall blood flow. Individual variations can occur, but significant heating or abnormal temperature patterns are not typically observed in the absence of any underlying pathology. The degree of temperature elevation can vary depending on the severity of inflammation and disease activity. In RA, joint inflammation and damage can lead to increased blood flow and warmth in affected joints [24-26], but there have been no studies comparing the heating rate of the healthy and RA fingers. The purpose of this research was to compare the heating rate of the proximal phalanx of the fingers and the whole palms in RA and healthy subjects. Consequently, the results may deliver tools for the detection of abnormal heat signatures in the phalanx proximal of the fingers.

2. METHODS

2.1. Subjects

The study was performed at the University Clinical Hospital in Białystok [5,6]. The measurement was taken for a total group of 48 patients with high disease activity hospitalised due to RA. The criteria for patients' inclusion were: age >18 years, and duration of anti-rheumatic biological therapies over 1 year. The exclusion criteria were: age <18 years old, and non-biological therapies. The control group was 45 healthy subjects. The patients participated in the study with their consent, according to the declaration of Helsinki. The Polish Regional Committees have approved this study for Medical and Health Research Ethics.

2.2. Measurement protocol

The subjects were asked to not consume alcohol or caffeinated drinks for 24 hours, smoke for 2 hours, and exercise for 1 hour before the image-acquisition phase. All subjects were introduced to the infrared camera and the course of the study. Next, the subjects' images for the affected joints of the hand were captured. The patients were instructed to immerse both hands up to the wrist in water thermostatically controlled at 0°C for 5 s; the local skin temperature dropped by 4–5°C and is stable [5,6]. Then removed their hands from the water, and the rewarming period was for 180 s, Fig. 1. The water temperature was continuously monitored with a digital thermometer. The thermograms of the hand fingers for each subject were taken by an experienced experimenter to ensure both rigor and repeatability [25] with the FLIR thermal imaging camera E60bx (Systems Inc., USA) with a resolution of 320 × 240 pixels and thermal sensitivity of <0.05°C, following the guidelines of the American Thermology Association. Thermography images were taken in a controlled environment (air humidity 55%, emissivity 0.98 with stable temperature 23±1°C and minimal airflow). FLIR Tools software for the analysis of images was used.
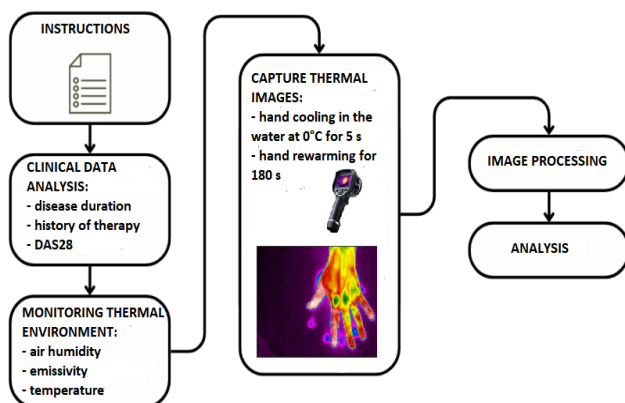
Fig. 1. Flowchart of the investigation procedure.

The heating rate of the proximal phalanx of the fingers (Fig. 2) was analysed in two steps: (1) hand cooling; and (2) hand rewarming.
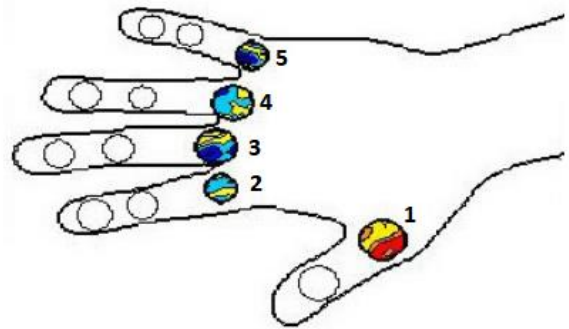
Fig. 2. Definitions of the regions of interest used in this study: 1 – Pollex; 2 – Index; 2 – Medius; 4 – Annularis; 5 – Digitus minimus.

2.3. Image processing

The temperature data was considered as two-dimensional matrices $L_i: w \times h \to T_i$, where w and h determine the frame size in Cartesian coordinates $x = 0..\ w$, $y = 0..\ h$, and the number of frames are $i = 1..\ n$. Secondly, the data were used as a table function. The frame number corresponds to the time $t = 0..\ t_n$. Therefore, $L_i: w \times h \to T(x, y, t)$. In this case, the temperature measured at the initial state of a hand represented only one frame, which means that $i \equiv n = 1$. Therefore, $L_{init}: w \times h \times 1 \to T_{init}$. Each frame contains information about the temperature of two objects: hand and a background. At the first stage, a hand was separated from a background. Anjos [27] proved that the balanced histogram thresholding method delivers good results for segmentation of the frame L_{init} and the frame L_n, and contains information on the completion of the heating process of pre-chilled hands (Figs. 3 and 4a). The change in the hand temperature from the first measurement to the last was determined as below:

$$\Delta T(x,y) = T_n\ (x,y) - T_1\ (x,y). \tag{1}$$

In case of uniform and incomplete heating, the heating rate was determined according the formula [27]:

$$v'(x,y) = \frac{\Delta T(x,y)}{t_n}. \tag{2}$$

We assumed that the temperature peaks are reached not simultaneously at different points of the surface. Consequently, the instant of the temperature peaks in the points are $t_{max}^j \leq t_n$, where $j = 1..\ w \times h$ is a number of element of a two-dimensional matrix L_i, corresponding to a pair of Cartesian coordinates (x, y) of the point on the surface. Therefore, $T(x, y, t_{max}^j) \geq T(x, y, t_n^j)$. If $T(x, y, z)$ is a three-dimensional function of temperature in some heat-conducting medium, but the temperature varies only along one axis 0Z, then the density vector of the heat flux transferred by the thermal conductivity is proportional to the temperature gradient vector $q \propto -grad(T) = -\bar{e}_z\ \partial T/\partial z$. The function of temperature is also three-dimensional, and the temperature varies only along one axis 0 t too. In a general case, the heating rate at each point on the surface by differentiating the temperature function with respect to time is obtained as follows (Fig. 4b):

$$v(x,y) = \frac{\partial T(x,y,t^j)}{\partial t}, \quad t^j = 0..\ t_{max}^j. \tag{3}$$

The vector function $v(x, y)$ reflects the processes of heating in more detail than the distribution of the scalar values of the instants (Fig. 5). It is a two-dimensional vector field; therefore, a divergence of the heating rate field must be found (Fig. 6):

$$div\big(v(x,y)\big) = \nabla \cdot \frac{\partial T(x,y,t^j)}{\partial t}, \qquad (4)$$

$$div\big(v(x,y)\big) = \frac{\partial}{\partial x} \cdot \frac{\partial T(x,y,t^j)}{\partial t} + \frac{\partial}{\partial y} \cdot \frac{\partial T(x,y,t^j)}{\partial t} + + \frac{\partial}{\partial t} \cdot \frac{\partial T(x,y,t^j)}{\partial t}. \qquad (5)$$

The partial derivatives of the heating rate with respect to the Cartesian coordinates are zero because the heating rate vector is orthogonal to the axes $0X$ and $0Y$:

$$div\big(v(x,y)\big) = \frac{\partial}{\partial t} \cdot \frac{\partial T(x,y,t^j)}{\partial t}. \qquad (6)$$

2.4. Statistical analysis

The temperature data were presented as means and standard deviations (SD). To verify the hypothesis of a normal distribution of the analysed variables, the Shapiro–Wilk test was used. To further analyse variables with normal distribution, a parametric test (Student's test) and other nonparametric tests were used. Statistical analyses were performed using Statistica software version 13.1 (StatSoft, Poland).

3. RESULTS

The mean age for healthy subjects was 52.9 ± 4.2 years and for RA patients was 52.3 ± 5.4 years. The image data were recorded on 5,400 frames (thermograms). The initial temperature (hand cooling) and final temperature (hand rewarming) are presented in Fig. 3 and in Tab. 1.

Tab. 1. The temperature (standard deviation) during hand cooling and hand rewarming in the control group and RA patients, [°C] ±SD

Phalanx proximalis	Control group		RA group	
	cooling	rewarming	cooling	rewarming
Pollex	23.5 (1.75)	29.5 (1.95)	23.0 (2.75)	29.5 (2.45)
Index	24.2 (2.40)	32.0 (1.75)	23.0 (2.35)	31.5 (1.35)
Medius	24.0 (2.10)	32.5 (2.15)	23.5 (1.60)	30.5 (2.15)
Annularis	23.5 (1.90)	32.3 (1.95)	24.0 (2.50)	31.0 (1.75)
Digitus minimus	23.0 (2.50)	31.0 (1.65)	22.0 (1.80)	31.0 (2.75)

The process is stationary, and its parameters are stable at any time interval. This is a consequence of the ability of a living organism to maintain viability.

Fig. 4a shows the coldest and the warmest temperature of a subject hand as well as the diagram of the average values of the frame sequence that reflect the process of the subject's hand heating (Fig. 4b).

Fig. 5 shows the time spatial distributions to reach temperature peaks at points $(t^j_{\max})$ and the heating rate at each point $(v(x, y))$ of the subject's hand.
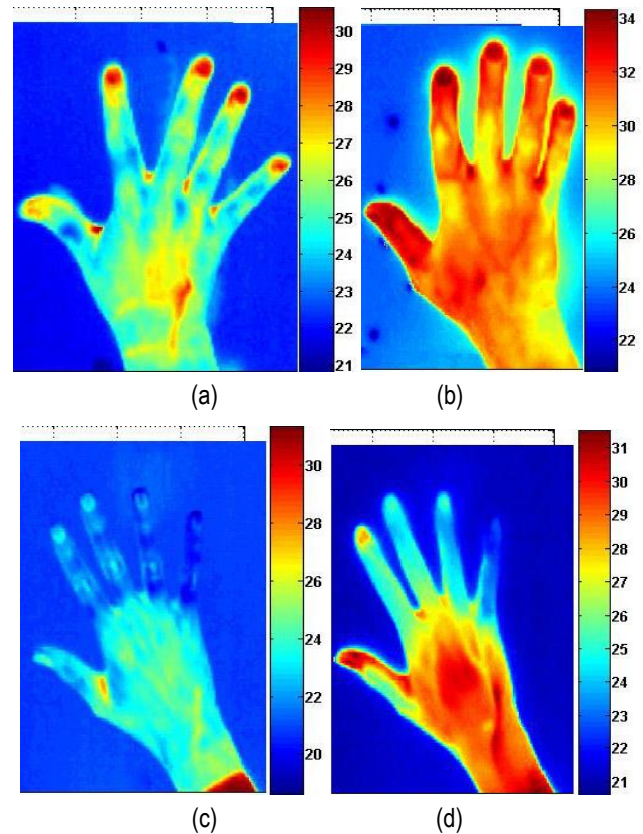
Fig. 3. Thermogram: (a) hand cooling in typical; (b) hand rewarming in typical; (c) hand cooling in RA patient; (d) hand rewarming in RA patient. RA, rheumatoid arthritis.
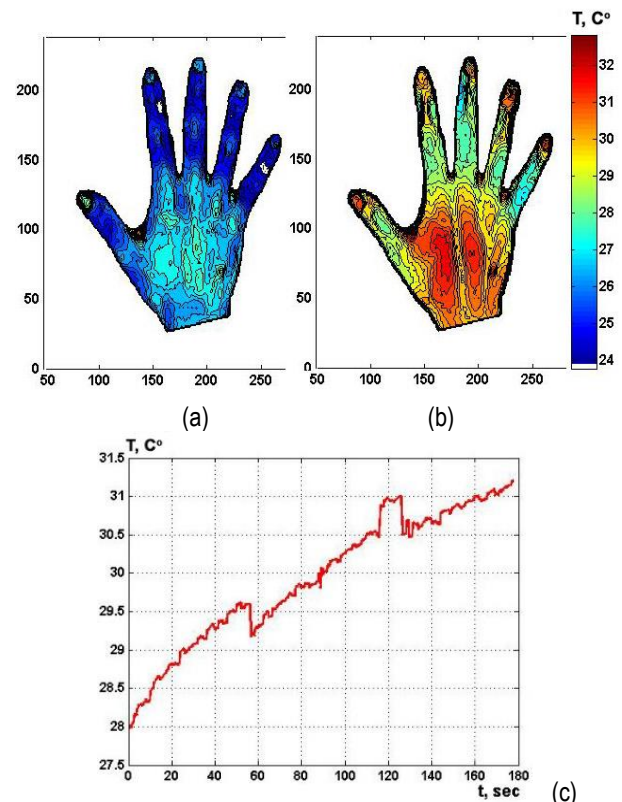
Fig. 4. The thermograms after segmentation: (a) the coldest and the warmest temperature of a hand; (b) a diagram of the heating rate.

(a)   (b)

(c)
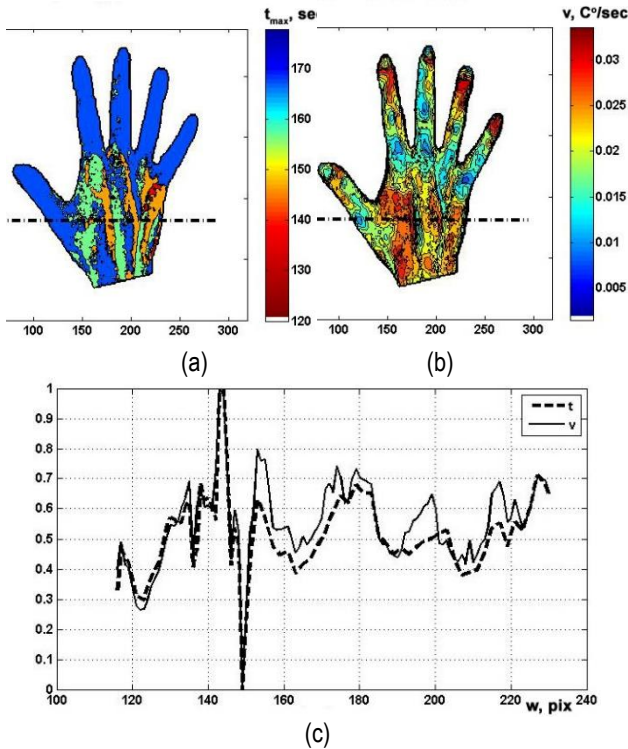
Fig. 5. Space distribution: (a) the time of the temperature peaks; (b) the heating rate; (c) the cross-sections of the time and the heating rate

(a)

(b)
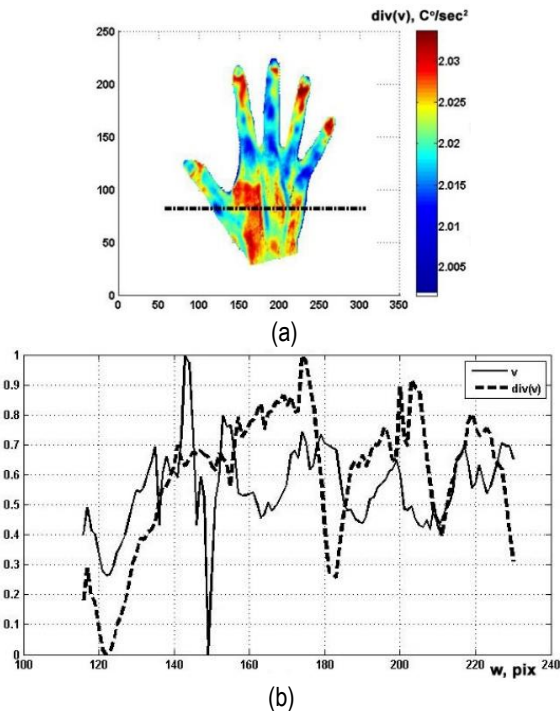
Fig. 6. The divergence of the heating rate: (a) space distribution; (b) the cross-sections of the heating rate and its divergence.

The divergence of the heating rate (Fig. 6a) shows the density of the sources of the central field of the heating rate at each point of the heated surface of the hand. Fig. 6b shows the normalised cross-sections of the heating rate and its divergence.

The maximum temperature on the thermograms was reached at different times in restoring the hand's temperature after cooling. Therefore, it can be assumed that the rate of temperature in-

creases (degree/min) at the points will not be the same. The heating rate depends on the spatial distribution of the inflow of the coolant (blood) and may be an important diagnostic parameter. At the first stage of the study, we did not take into account the anatomical correlation of the heating rate at the measured points but analysed the static parameters of the distributions. Typical histograms of the heating rate at the points of the image of the hand are shown in the images (Fig. 7). The rate of temperature rise (degree/min) is reflected in the x-axis. The values of the y-axis are normalised to the distribution integral, which made it possible to impose a Gaussian curve on the histogram.

(a)   (b)
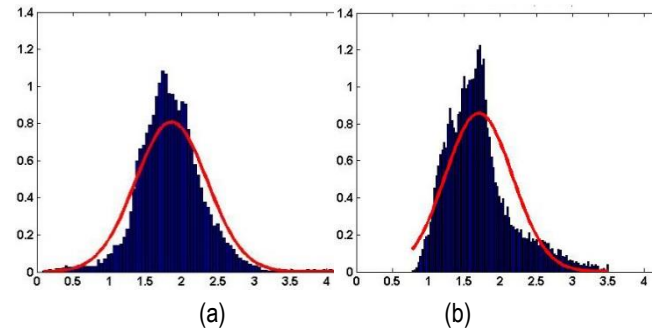
Fig. 7. Histograms of the heating rate (degree/min) at the points of the hand image: (a) control group; (b) RA. RA, rheumatoid arthritis.

Mean and SD were used to compare the distributions of the heating rate of people with RA and a control group of healthy people (Tab. 2). The average value in the RA group was slightly less than in the control group. The SD of the RA group was greater than that of the control group and was significant ($p < 0.05$). This is in good agreement with the uneven flow of the coolant (blood) caused by the disease.

Tab. 2. The parameters of the heating rate of hand by groups (RA, rheumatoid arthritis; SD, standard deviation)

Groups	Mean ($p > 0.05$)	SD ($p < 0.05$)
Control	1.865	0.292
RA	1.711	0.552
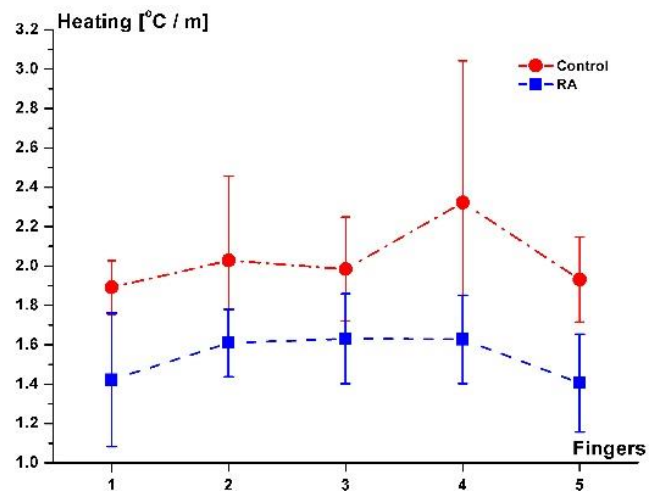
Fig. 8. The heating rate (degree/min) of proximal phalanx of the fingers ($p < 0.05$). RA, rheumatoid arthritis

sciendo

Mikhail Ihnatouski, Jolanta Pauk, Kristina Daunoraviciene, Jurgita Ziziene
Comparing the Heating Rate of the Proximal Phalanx of the Fingers in Rheumatoid Arthritis and Healthy Subjects

DOI 10.2478/ama-2024-0052

The next stage of the study was the localisation of the assessment of the heating rate in the anatomically important areas of the hand. The technique of finger extraction by skeletonisation has been described in Ref. [6]. We identified areas corresponding to the phalanges on the pre-discovered midlines of the fingers. Fig. 8 shows the heating rate of the proximal phalanx of the fingers ($p < 0.05$).

The heating rates of the control group were higher than those of the RA group, as well as the average heating rate of the hand as a whole. The difference in heating rates between the two groups on the proximal phalanx of the fingers was more than the difference between these groups across the whole hand. The heating rates of the proximal phalanx of the fingers 2, 3 and 4 were more than that of the two outermost fingers.

4. DISCUSSION AND CONCLUSIONS

Recent studies have investigated the use of thermography to detect early signs of joint inflammation in RA patients before the onset of visible joint damage. Some studies [2,3,6] found that thermography was able to detect joint inflammation in RA patients with normal physical exam findings, suggesting its potential as a sensitive screening tool. Other studies have also looked at the use of thermography to monitor disease activity and treatment response in RA patients [7,8]. In RA, the joints of the fingers are commonly affected, including the proximal interphalangeal (PIP) joints. That's why we decided to compare the heating rate phalanx proximal of the fingers in RA and healthy subjects. Previously, active thermography has been used in research studies to investigate the thermal properties of joints in RA patients [28]. In these studies, a dry cooling of the skin for 1 min was used to induce a thermal contrast on the joint, which was then measured with an infrared camera. Other heat sources, such as surface cooling with a mixture of CO_2 and air from a cryotherapy unit and also laser or microwave heating, may be more suitable in some cases [29]. In all cases, it is very important to follow established measurement protocols and safety guidelines. We used water as a heat source in active thermography studies [5,6] to minimise thermal exposure to the patient. In Ref. [5], the outcomes included the mean temperature of five fingers of a hand: In static, post-cooling, and post-rewarming, the total change in mean temperature of fingers due to cold provocation, the total change in mean temperature of fingers due to rewarming, the area under the cooling curve, the area under the heating curve, the difference between the area under the rewarming and the cooling curve, and temperature intensity distribution maps. For patients with high disease activity, a lower area under the heating curve and a lower difference between the area under the rewarming curve and the cooling curve were observed, as well as a smaller total change in mean temperature due to rewarming, compared to patients with moderate disease activity. In turn, the study [6] proved that the cold provocation test discriminates between RA patients and controls and detected inflammation in RA patients by the measurement of temperature profiles along the fingers using an infrared camera.

This study examines the dynamical change in the temperature field over time. Integral parameters were also taken into account, as well as parameters of specific anatomical areas. We found that there were significant differences in the heating rate on the phalanx proximalis of RA patients and healthy control subjects. Specifically, the RA-affected joints had lower peak temperatures and

slower cooling rates compared to the healthy joints. The use of water as a heat source in active thermography for RA patients has some limitations. RA patients may have joint deformities, inflammation and other conditions that can affect the thermal properties of the joint and the ability of water to induce a thermal contrast.

Overall, these studies suggest that active thermography can be used to detect thermal differences between RA-affected joints and healthy joints in the same patient and may have the potential as a non-invasive tool for monitoring disease activity and treatment efficacy in RA.

REFERENCES

1. McInnes IB, Schett G. The pathogenesis of rheumatoid arthritis. N Engl J Med. 2011;365:2205–19.
   https://doi.org/10.1056/NEJMra1004965
2. Branco JHL, Branco RLL, Siqueira TC, de Souza LC, Dalago KMS, Andrade A. Clinical applicability of infrared thermography in rheumatic diseases: A systematic review. J Therm Biol. 2022; 104:103172. https://doi.org/10.1016/j.jtherbio.2021.103172
3. Sanchez BM, Lesch M, Brammer D, Bove SE, Thiel M, Kilgore KS. Use of a portable thermal imaging unit as a rapid, quantitative method of evaluating inflammation and experimental arthritis. J Pharmacol Toxicol Methods. 2008;57(3):169-75.
   https://doi.org/10.1016/j.vascn.2008.01.003
4. Kow J, Tan YK. An update on thermal imaging in rheumatoid arthritis. Joint Bone Spine. 2023;90(3):105496.
   https://doi.org/10.1016/j.jbspin.2022.105496
5. Pauk J, Wasilewska A., Ihnatouski M. Infrared thermography sensor for disease activity detection in rheumatoid arthritis patients. Sensors. 2019;19(16):3444. https://doi.org/10.3390/s19163444
6. Pauk J, Ihnatouski M, Wasilewska A. Detection of inflammation from finger temperature profile in rheumatoid arthritis. Med Biol Eng Comput. 2019;57(12):2629-2639.
   https://doi.org/10.1007/s11517-019-02055-1.
7. Morales-Ivorra I, Narváez J, Gómez-Vaquero C, Moragues C, Nolla JM, Narváez JA, Marín-López MA. A Thermographic Disease Activity Index for remote assessment of rheumatoid arthritis. RMD Open. 2022;8(2):e002615. https://doi.org/10.1136/rmdopen-2022-002615
8. Morales-Ivorra I, Narváez J, Gómez-Vaquero C, Moragues C, Nolla JM, Narváez JA, Marín-López MA. Assessment of inflammation in patients with rheumatoid arthritis using thermography and machine learning: a fast and automated technique. RMD Open. 2022;8(2): e002458. https://doi.org/10.1136/rmdopen-2022-002458
9. Bardhan S, Bhowmik MK. 2-Stage classification of knee joint thermograms for rheumatoid arthritis prediction in subclinical inflammation. Australas Phys Eng Sci Med. 2019;42(1):259-277.
   https://doi.org/10.1007/s13246-019-00726-9
10. Ahalya RK, Snekhalatha U, Dhanraj VJ. Automated segmentation and classification of hand thermal images in rheumatoid arthritis using machine learning algorithms: A comparison with quantum machine learning technique. Therm Biol. 2023;111:103404.
    https://doi.org/10.1016/j.jtherbio.2022.103404
11. Snekhalatha U, Anburajan M, Sowmiya V, Venkatraman B, Menaka M: Automated hand thermal image segmentation and feature extraction in the evaluation of rheumatoid arthritis, Proc Inst Mech Eng H 2015;229(4):319-31. https://doi.org/10.1177/0954411915580809
12. Tripoliti EE, Fotiadis D, Argyropoulou M. Automated segmentation and quantification of inflammatory tissue of the hand in rheumatoid arthritis patients using magnetic resonance imaging data. Artif Intell Med 2007;40(2):65-85. https://doi.org/10.1016/j.artmed.2007.02.003
13. Venerito V, Angelini O, Cazzato G, Lopalco G, Maiorano E, Cimmino A, et al. A convolutional neural network with transfer learning for automatic discrimination between low and high-grade synovitis: a pilot study. Intern Emerg Med. 2021;16:1457–65.
    https://doi.org/10.1007/s11739-020-02583-x

14. Folle L, Meinderink T, Simon D, Liphardt AM, Krönke G, et al. Deep learning methods allow fully automated segmentation of metacarpal bones to quantify volumetric bone mineral density. Sci Rep. 2021; 11:9697–706. https://doi.org/10.1038/s41598-021-89111-9

15. Norgeot B, Glicksberg BS, Trupin L, Lituiev D, Gianfrancesco M, Oskotsky B, et al. Assessment of a deep learning model based on electronic health record data to forecast clinical outcomes in patients with rheumatoid arthritis. JAMA Netw Open. 2019;2:e190606. https://doi.org/10.1001/jamanetworkopen.2019.0606

16. Fukae J, Isobe M, Hattori T, Fujieda Y, Kono M, Abe N, et al. Convolutional neural network for classification of two-dimensional array images generated from clinical information may support diagnosis of rheumatoid arthritis. Sci Rep. 2020;10:5648. https://doi.org/10.1038/s41598-020-62634-3

17. Üreten K, Erbay H, Maraş HH. Detection of rheumatoid arthritis from hand radiographs using a convolutional neural network. Clin Rheumatol. 2020;39:969–74. https://doi.org/10.1007/s10067-019-04487-4

18. Christensen ABH, Just SA, Andersen JKH, Savarimuthu TR. Applying cascaded convolutional neural network design further enhances automatic scoring of arthritis disease activity on ultrasound images from rheumatoid arthritis patients. Ann Rheum Dis. 2020;79:1189–93. https://doi.org/10.1136/annrheumdis-2019-216636

19. Tan YK, Hong C, Li H, Allen JC Jr, Thumboo J. Thermography in rheumatoid arthritis: a comparison with ultrasonography and clinical joint assessment. Clin Radiol. 2020;75(12):963.e17-963.e22. https://doi.org/10.1016/j.crad.2020.08.017

20. Umapathy S, Thulasi R, Gupta N, Sivanadhan S. Thermography and colour Doppler ultrasound: a potential complementary diagnostic tool in evaluation of rheumatoid arthritis in the knee region. Biomed Tech (Berl) 2020;26;65(3):289-299. https://doi.org/10.1515/bmt-2019-0051

21. Mountz JM, Alavi A, Mountz JD. Emerging optical and nuclear medicine imaging methods in rheumatoid arthritis. Nat Rev Rheumatol. 2012;8(12):719-28. https://doi.org/10.1038/nrrheum.2012.148

22. Tan YK, Hong C, Li H, Allen JC Jr, Thumboo J. A novel use of combined thermal and ultrasound imaging in detecting joint inflammation in rheumatoid arthritis. Eur J Radiol. 2021;134:109421. https://doi.org/10.1016/j.ejrad.2020.109421

23. Dreher R, Müller K, Grebe SF, Altaras J, Federlin K. [Scintigraphic, thermographic and radiographic findings in rheumatoid arthritis (RA) and their value for diagnosis and therapy]. Verh Dtsch Ges Inn Med. 1978;(84):1492-6.

24. Tegelberg A, Kopp S. Skin surface temperature over the temporomandibular and metacarpophalangeal joints in individuals with rheumatoid arthritis. Acta Odontol Scand. 1987;45(5):329-36. https://doi.org/10.3109/00016358709096355

25. Gatt A, Mercieca C, Borg A, Grech A, Camilleri L, Gatt C, Chockalingam N, Formosa C. A comparison of thermographic characteristics of the hands and wrists of rheumatoid arthritis patients and healthy controls. Sci Rep. 2019;25;9(1):17204. https://doi.org/10.1038/s41598-019-53598-0

26. Fischer M, Mielke H, Glaefke S, Deicher H. Generalized vasculopathy and finger blood flow abnormalities in rheumatoid arthritis. J Rheumatol. 1984;11(1):33-7.

27. Anjos A, Leite R, Cancela ML, Shahbazkia H. MAQ – A bioinformatics tool for automatic macroarray analysis. International Journal of Computer Applications 2010;4(3). https://doi.org/10.5120/843-1066

28. Rusch D, Follmann M, Boss B, Neeck G. Dynamic thermography of the knee joints in rheumatoid arthritis (RA) in the course of the first therapy of the patient with methylprednisolone. Z Rheumatol. 2000;59(2):II/131-5. https://doi.org/10.1007/s003930070009

29. Nowakowski A. Problems of active dynamic thermography measurement standarization in medicine. Pomiary Automatyka Robotyka 2021;3: 51-56. https://doi.org/10.14313/PAR_241/51

Mikhail Ihnatouski: https://orcid.org/0000-0003-4350-3401

Jolanta Pauk: https://orcid.org/0000-0001-8607-7129

Kristina Daunoraviciene: https://orcid.org/0000-0003-0898-4860

Jurgita Ziziene: https://orcid.org/ /0000-0002-2624-4375

Darius Plonis, Edas Kalinauskas, Andrius Katkevičius, Audrius Krukonis
Non-Invasive Analysis of the Bioelectrical Impedance of a Human Forearm

DOI 10.2478/ama-2024-0053

NON-INVASIVE ANALYSIS OF THE BIOELECTRICAL IMPEDANCE
OF A HUMAN FOREARM

Darius PLONIS* , Edas KALINAUSKAS** , Andrius KATKEVIČIUS* , Audrius KRUKONIS*

*Vilnius Gediminas Technical University, Department of Electronic Systems, Vilnius, Lithuania,
**JSC Kongsberg NanoAvionics, Vilnius, Lithuania

darius.plonis@vilniustech.lt, edas.kalinauskas@nanoavionics.com, andrius.katkevicius@vilniustech.lt, audrius.krukonis@vilniustech.lt

Abstract: This study explores the practical application and impact of bioimpedance analysis in mobile devices for monitoring human health. The objective of the study is to propose a feasible application of non-invasive bioimpedance analysis by using the tetrapolar electrode connection method and the Cole–Cole model. Bioimpedance measurements and the calculation of electrical parameters are performed using ANSYS HFSS software for theoretical calculations and digital signal processing technology for real-time measurements using hardware devices. The study focuses on a model of the front arm, including tissues such as bone, fat, muscles, arteries and skin, with glucose concentrations as test cases. The simulated characteristic impedance with the ANSYS HFSS software package at 125 kHz varied from 315.8 Ω to 312.6 Ω, and the measured forearm characteristic impedance with hardware varied from 150.1 Ω to 151.3 Ω. The measured characteristic impedance when the heart is in systole and diastole also differed, with a difference of about 0.85% of the maximum impedance measured. The study demonstrates the potential of non-invasive bioimpedance analysis to address health issues such as obesity and heart disease. It also highlights its usefulness as a non-invasive alternative for measuring glucose concentration in diabetic patients to reduce the risk of infection. The findings indicate the feasibility of using bioimpedance analysis in mobile devices for health monitoring purposes.

Key words: non-invasive, bioimpedance, digital signal processing, human tissues

1. INTRODUCTION

Bioimpedance measurements are widely used in various fields [1–3]. Bioimpedance refers to the tissue's resistance to the flow of alternating electrical current. When an alternating electrical current passes through the human body, the Z complex impedance of the tissues consists of two components: the imaginary part (reactive resistance, X) and the real part (active resistance, R).

Bioimpedance measurement is based on the fact that the complex impedance of conductive materials depends on the shape of the conductor. In a cylindrical conductor, impedance is directly proportional to the length of the conductor and inversely proportional to the cross-sectional area of the conductor [4].

Different tissues in the body respond to electrical current such as conductors, semiconductors or dielectrics. Lean body tissues respond to electrical current as conductors because they contain a high amount of water and electrolytes. Fatty tissues and bones, on the other hand, act as dielectrics. Since electrical current will primarily flow through the most conductive tissues in the human body, such as interstitial fluids and muscles, these tissues will have the greatest influence on the overall body impedance.

Since the human body does not have a regular cylindrical shape, a model is used when measuring body impedance, in which different body parts are represented as separate cylindrical components: two arms, the human back and two legs [5].

When performing bioimpedance analysis and calculating body impedance, challenges arise because the human body consists of both active and reactive components of impedance. The reactive component of impedance in the human body is due to cell membranes, while the active component is influenced by intracellular and extracellular fluids. It can be inferred that cell membranes in the body act as capacitors, while intracellular and extracellular fluids function as resistors [6].

To calculate the reactive and active components of body impedance, several steps are involved. Firstly, the effective voltage and current values are calculated. Then, power and effective power are determined. Finally, phase shift, reactive impedance and active impedance can be calculated [7].

Currently, the most commonly used methods for electrode placement are bipolar and tetrapolar configurations. In bipolar electrode placement, two electrodes are connected to the human body to supply the tissue with alternating current and measure the voltage between the two electrodes. In tetrapolar electrode placement, four electrodes are used. Two electrodes are responsible for supplying the body with alternating current, while the other two measure the voltage [8].

When measuring the whole-body impedance, there are several methods for electrode placement on the body. These methods differ based on electrode attachment to different body locations, allowing the alternating current to flow through different segments of the human body. The most commonly used methods are hand-to-hand [9], foot-to-foot [10] and hand-to-foot [11] methods. Currently, the most popular method is the hand-to-foot method.

Segmental impedance analysis involves measuring the impedance of individual body segments. By measuring the impedance of specific body regions using this method, it is possible to

calculate the mass of different tissues in the entire body and monitor parameters such as heart activity, the relationship between impedance and blood glucose levels, and the respiratory rhythm of an individual [12,13].

In this method of measuring bioelectrical impedance, a constant frequency current of 50 kHz is commonly used. This current flows through electrodes attached to the individual's hand and foot [14]. By conducting measurements with a current of 50 kHz, it is possible to determine the fat-free mass and total body water content, but it is not possible to assess the specific quantities of intracellular and extracellular fluids since the 50 kHz current cannot penetrate cell membranes, which act as capacitors. Therefore, this method is most suitable for healthy individuals with a normal body water content and who are not severely dehydrated [15].

The limitations mentioned earlier can be overcome by the method of bioimpedance analysis using multiple frequency measurements, typically ranging from 0 kHz to 500 kHz (0, 1, 5, 50, 100, 200 and 500). This method is based on the principle that different tissues have different electrical conductivities at different frequencies. By applying currents at different frequencies, it is possible to estimate fat-free mass, total body water and the quantities of extracellular and intracellular fluids. Research has shown that using variable-frequency electrical currents allows for more accurate estimation of extracellular fluids [16]. However, a previous study [17] has indicated that this method is not suitable for determining the distribution of water between intracellular and extracellular compartments in the bodies of elderly individuals.

Another application area is in the measurement of fat-free mass. The initial methods for calculating fat-free mass were solely based on measurements of height and impedance. To improve measurement accuracy, formulas were later derived that required additional individual data such as weight, age, sex, reactance and anthropometric [18]. In this method, a constant current of 50 kHz, 800 µA is used to calculate bioimpedance. Kotler et al. [19] developed a formula specifically for calculating fat-free mass in healthy individuals. This method utilises a fixed-frequency 50 kHz electrical current for impedance measurement.

In the human body, body fluids consist of intracellular (ICW) and extracellular (ECW) fluids. Studies have shown that bioelectrical impedance analysis at varying frequencies is more suitable for measuring body fluids as it yields lower measurement errors [20].

This method involves measuring impedance at varying frequencies. A current with frequencies ranging from 0 kHz to 100 kHz was passed through the human body. Impedance was measured at specific frequencies: 1 kHz, 5 kHz, 50 kHz and 100 kHz. The results showed that the most accurate determination of the total body fluid volume was obtained using an electrical current of 100 kHz, while the most accurate data for extracellular fluid volume were obtained using an electrical current of 1 kHz. Kushner and Schoeller [21] derived a formula for calculating the total body fluid volume.

Constant frequency bioelectrical impedance analysis is also used to determine the total body skeletal muscle mass [22].

Equally important research has been conducted in a previous study [23], stating that under different pressures in the arteries, their diameter can vary from 2.274 mm to 2.519 mm. Another study [24] presents an electrical model of arterial impedance variation due to the heart pulse. It is also crucial to note that blood conductivity and dielectric permittivity parameters vary with changes in blood glucose levels. In a study by Li et al. [25], the

calculation was made on how glucose affects the electrical parameters of deionised water at glucose concentrations ranging from 0 mmol/L to 225 mmol/L.

In addition, in this study, it was observed that dielectric permittivity parameters, conductivity and dielectric relaxation time have a linear dependence on the glucose concentration in deionised water solution. Based on this finding, formulas were derived to calculate the electrical parameters of the solution at different glucose concentrations [22]. The observation that glucose affects the conductivity of a non-conductive solution leads to the conclusion that different glucose concentrations also affect the electrical conductivity of blood. Therefore, by knowing the electrical parameters of blood, the diameter and length of the artery at the measurement site, it is possible to calculate bioimpedance [26].

To measure bioimpedance, a tetrapolar 4-electrode connection method was used. This method was chosen because the shape of electrodes, and contact impedance has less influence on measurement accuracy than the bipolar measurement method. In the tetrapolar configuration, a current is applied through two electrodes, while the other two electrodes are used for voltage measurement. No current flows through the voltage measurement electrodes. The entire electrical current passes only through the tested tissue. Consequently, there is no voltage drop due to contact impedance between the voltage measurement electrodes. Additionally, when performing measurements using the tetrapolar connection method, the body position has less impact than the bipolar measurement method. Finally, the tetrapolar electrode configuration method is less sensitive to electrode placement than the bipolar method.

This study will be conducted on the human wrist. This location was chosen because it is convenient to measure the variation in bioimpedance due to the presence of the radial and ulnar arteries in the wrist. The radial and ulnar arteries are among the main arteries in the human body.

2. INVESTIGATED MODELS OF THE FOREARM

The described study consists of two parts: theoretical and experimental. In the theoretical part, computer modelling of bioimpedance analysis is performed using the ANSYS HFSS software package. In the experimental part, bioimpedance measurements are taken on the human wrist using a specially prepared measurement setup.

2.1. Computer-based model of the forearm using ANSYS HFSS

The electrical model of the forearm was created using the ANSYS HFSS software package, considering six tissues: the skin layer, adipose tissue, arteries, muscles, cortical bone and bone marrow (Fig. 1).

The length of the electrical spatial model is 25 cm, the width is 5.05 cm and the height is 3.05 cm. The proportions and thickness of the selected tissues in the study were the same as those used in the study conducted by Yu et al. [1] Dimensions of the model are provided in Fig. 2. In the model (Fig. 1), the skin layer is depicted in orange, adipose tissue in yellow, arteries in red, muscles in pink, cortical bone in grey and bone marrow in white. Green colour represents the electrodes used to connect to the tissues.

Darius Plonis, Edas Kalinauskas, Andrius Katkevičius, Audrius Krukonis
Non-Invasive Analysis of the Bioelectrical Impedance of a Human Forearm

DOI 10.2478/ama-2024-0053

The electrodes located on the outer surface of the model are current electrodes, while the electrodes at the centre are voltage electrodes. During the development of the model, separate parameters were defined, allowing for easy modification of the length, width, height, artery diameter and electrode layout of the model. In the centre of the model, a space was created to visualise the distribution of current within the model.

The power source for bioelectrical impedance measurements in the ANSYS HFSS software package was implemented using current sources. The amplitude of the current was chosen to be 0.1 mA. The selected maximum current value is safe and complies with the IEC 60601-2-47:2012 safety standard for medical electrical equipment – Part 2-47: Particular requirements for the basic safety and essential performance of ambulatory electrocardiographic systems. This choice was made in consideration of the capabilities of the experimental equipment that will be used in subsequent experiments.
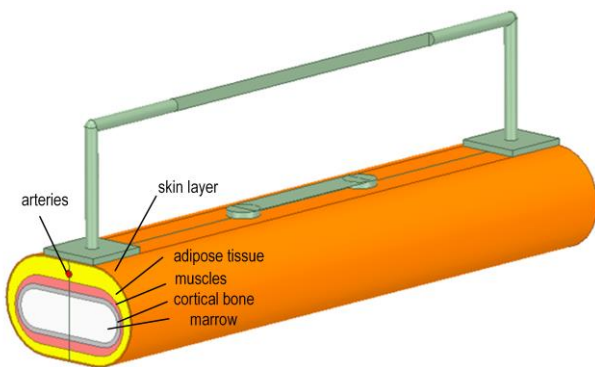
Fig. 1. Computer model of the human forearm, created using the ANSYS HFSS software package, consists of the following components: the skin layer in orange, adipose tissue in yellow, arteries in red, muscles in pink, cortical bone in grey and bone marrow in white
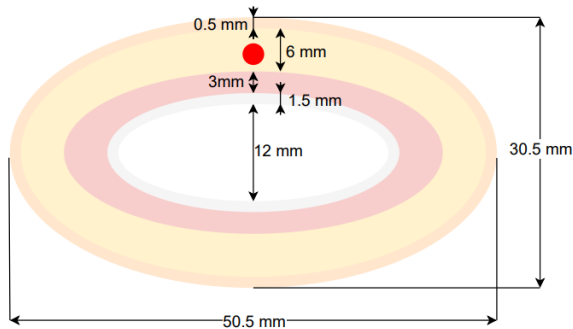
Fig. 2. Dimensions of the tissues in the model
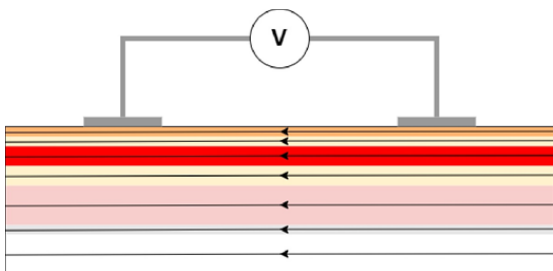
Fig. 3. Distribution of current through the tissues in the forearm electrical model [15]

The ANSYS HFSS software package cannot automatically calculate the bioelectrical impedance using current sources. Therefore, it needs to be calculated using the field calculator function. To do this, firstly, the voltage drop between the voltage measurement electrodes needs to be calculated. This can be done by knowing the electric field and the distance between two points.

The voltage drop was calculated using the field calculator tool in the ANSYS HFSS software package. The expression of Ohm's law was used to calculate the bioelectrical impedance.

The current flow in each tissue depends on the geometry and electrical parameters of the tissues. Since the geometry of the tissues in the thigh is irregular and complex, it is challenging to create an accurate electrical model. For this reason, the model will be composed of stacked tissues. The forearm model is shown in Fig. 3.

By creating such a model, it can be concluded that the current distributes through all the tissues of the forearm. Since the tissues in the model are connected in parallel, this model is equivalent to an electrical model of six resistors connected in parallel, and their total resistance can be calculated using formula [15]:

$$\frac{1}{Z_{fa}} = \frac{1}{Z_{bm}} + \frac{1}{Z_{cb}} + \frac{1}{Z_m} + \frac{1}{Z_f} + \frac{1}{Z_b} + \frac{1}{Z_s}, \tag{1}$$

where Z_{fa} is the characteristic impedance of the forearm, Z_{bm} is the characteristic impedance of the bone marrow, Z_{cb} is the characteristic impedance of cortical bone, Z_m is the characteristic impedance of muscle, Z_f is the characteristic impedance of fat, Z_b is the characteristic impedance of blood and Z_s is the characteristic impedance of the skin.

To calculate the resistance of an individual tissue, the tissue's electrical conductivity needs to be known. With this value, the tissue resistance can be calculated using formula [15]:

$$R_{tiss} = \frac{l}{\sigma_{tiss} \cdot A_{tiss}}, \tag{2}$$

where R_{tiss} is the resistance of the tissue, L is the distance, σ_{tiss} is the electrical conductivity of the tissue and A_{tiss} is the cross-sectional area of the tissue.

During the computer modelling of the human wrist, three parameters were varied: frequency, radial artery diameter and different glucose concentrations in the blood. The radial artery diameter of the human wrist was limited to two different values ranging from 2.274 mm to 2.519 mm. Four different glucose concentrations in the blood were used to calculate the characteristic resistances of the simulated human wrist model in ANSYS HFSS software. In total, calculations were performed at 48 different frequency points. The calculated results of the model are presented on a logarithmic frequency scale to visualise the variation in the characteristic resistances, especially at lower frequencies.

2.2. Physical forearm measurement model

Bioelectrical impedance was measured using the MAX30001 integrated circuit. The MAX30001 circuit is a complete solution for wearable applications and serves as a biopotential and bioimpedance front analogue. It provides high performance and extremely low power consumption to extend battery life in clinical and fitness applications. In a single biopotential channel, it provides a form of pulse in the electrocardiogram (ECG), heart rate and pacemaker edge detection. In addition, there is a single bioimpedance channel to measure breathing.

The bioimpedance channel is equipped with integrated programmable current drives that support common electrodes and provide flexibility for measuring two or four electrodes.

During the measurements, the maximum allowable frequency of MAX3001 was used, which is 128 kHz. A 1 kHz high-frequency filter was applied to the voltage measurement electrode inputs. This filter is designed to eliminate high-frequency noise from bioelectrical impedance measurements. Two MAX3001 integrated operational amplifiers were used during the measurements. The first operational amplifier was set to operate in a low-noise mode. The gain coefficient of the second operational amplifier was set to 20 V/V. The analogue-to-digital converter sampling frequency was set to the maximum possible value of 64 Hz. The integrated digital filters of the MAX3001 prototype were not used since all signal processing will be performed using the MATLAB software package. The amplitude of the alternating electrical current was set to the maximum allowable value of 0.096 mA.

The real physical experiment setup, using the MAX3001 prototype with four electrodes, is shown in Fig. 4. The four electrodes were arranged equidistantly on the human forearm. Real physical experiments were conducted on a healthy individual without diabetes. The glucose concentration in the blood was measured using a commercial glucometer – CONTOUR PLUS ELITE.
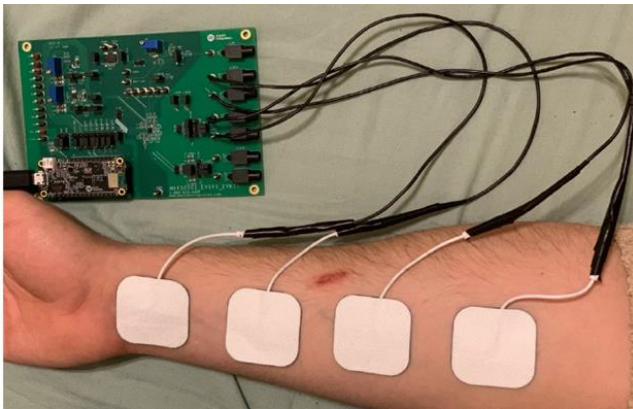
Fig. 4. Physical measurements of the characteristic impedance based on the MAX30001 integrated circuit

The MAX30001 integrated circuit is connected to an external microcontroller, which is used to collect and process information. Samples are recorded in 20-bit resolution. This sensor calculates the characteristic impedance based on the following formula [6]:

$$Z(\Omega) = ADC \cdot \frac{V_{ref}}{2^{19} \cdot CG_{MAG} \cdot GAIN},$$ (4)

where *ADC* is the voltage reading from the analogue-to-digital converter of the sensor, V_{ref} is the reference voltage, CG_{MAG} is the magnitude of the current excitation and *GAIN* is the internal gain applied to the voltage measurement.

This integrated circuit was chosen for its small size, low cost and ease of operation. Due to these characteristics, this sensor is suitable for use not only in scientific research but also in commercial products.

The tested individual should fasting for 8 h. The person was tested when in a lying position, fully relaxed. Initially, the blood glucose level was measured before glucose consumption. After measuring the blood glucose concentration, the subject consumed 75 g of glucose. The next measurement was taken 30 min after glucose intake, and subsequent measurements were taken every 30 min. The test lasted a total of 90 min. A total of four measurements of glucose and characteristic impedance were taken. Each measurement of characteristic impedance lasted 5 min to capture multiple readings of impedance changes and increase the reliability of the test. The testing equipment and its setup are shown in Fig. 4.

After conducting the measurements of characteristic impedance, the previously discussed signal processing was performed to remove noise caused by human respiration and electrical interference.

3. SIGNAL PROCESSING METHODS

The planned measurements are expected to be quite noisy, so it is important to discuss possible signal processing methods. The most common sources of noise include human movements, respiration and electrical interference. The initial unprocessed measurement results with evident low-frequency noise influenced by human respiration are presented in Fig. 5. Fig. 5 shows a 30 s data window with characteristic impedance calculations from raw data.
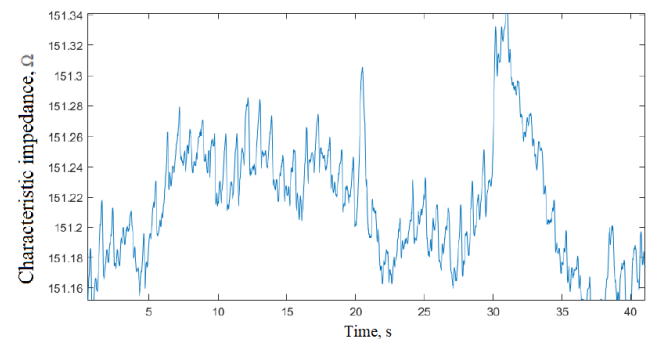
Fig. 5. Example of characteristic impedance calculated from raw data

The aforementioned filtering helps to eliminate the previously mentioned sources of noise, such as respiration and electrical interference. From the filtered signal, peak values of the signal can be extracted. Each signal maximum and minimum indicate the impedance during systole and diastole of the heart, respectively (Fig. 6).
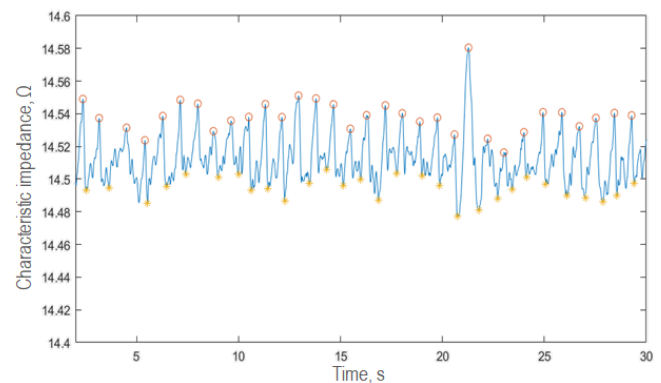
Fig. 6. Calculated characteristic impedance from raw data with the applied Chebyshev bandpass filter

sciendo

Darius Plonis, Edas Kalinauskas, Andrius Katkevičius, Audrius Krukonis
Non-Invasive Analysis of the Bioelectrical Impedance of a Human Forearm

DOI 10.2478/ama-2024-0053

A hundredth-order Chebyshev bandpass filter was designed for the initial data filtering by using the MATLAB software package. The sampling frequency was chosen to be 130 Hz, justified by the hardware used. The MAX3001 sampling rate was 64 Hz. The passband of the bandpass filter was set from 0.7 Hz to 40 Hz. Frequencies below 0.7 Hz were filtered out to eliminate the influence of respiration on the measurements [27]. The heart rate in the human body ranges from 50 beats/min to 180 beats/min, corresponding to frequencies from 0.8 Hz to 3 Hz. This filter ensures that such heart rate frequencies are captured. The upper cutoff frequency of the passband was chosen to be 40 Hz to filter out electrical interference present in the measurements. The filter has a stopband attenuation of 80 dB at the filtered frequencies.

4. RESULTS

In this section, computer modelling results using the ANSYS HFSS software package and experimental research results under real conditions are presented to validate the hardware on the human wrist.

4.1. Characteristic impedance results of computer-based modelling

Computer modelling using the ANSYS HFSS software package yielded results of the wrist computer model's characteristic impedance dependence on the frequency and blood glucose level. Computer modelling was performed in the frequency range from 10 kHz to 2 MHz. During the experiments, the blood glucose level was varied within the range of 4 mmol/L to 33 mmol/L.

The results of the variation in the characteristic impedance when the blood glucose concentration is 4 mmol/L are shown in Fig. 7. The graph displays three curves. The orange curve represents the characteristic impedance when the heart is in the systolic state. The blue curve represents the characteristic impedance when the heart is in the diastolic state. The grey curve shows the difference in the characteristic impedance between different heart states as a function of frequency. The initial analysis of the results indicates that the model's characteristic impedance is inversely proportional to the frequency of the alternating current.

As the artery fills with blood during systole, the pressure and diameter of the artery increase. At this moment, there is more electrically conductive blood in the artery, resulting in a lower value of characteristic impedance than the diastolic state when there is less blood in the artery and the diameter is reduced. The characteristic impedance during systole, as the frequency increases, varies from 330.86 Ω to 229.13 Ω in our examined frequency range from 10 KHz till 2 MHz. During diastole, the characteristic impedance varies from 332.99 Ω to 229.53 Ω.

The difference between the characteristic impedances is not directly dependent on the frequency of the alternating current. As shown in Fig. 8, initially, there is an increasing difference at lower frequencies. The difference reaches its maximum at around 150 kHz and then starts to decrease again. Throughout the frequency range, the difference varies from 0.39 Ω to 2.13 Ω. It can be concluded that at 150 kHz, the difference between the characteristic impedance of blood and the overall model impedance is the greatest, indicating the highest concentration of current in the artery. The current distribution in the other layers of the model is minimal. Since the highest current concentration is in the artery at

150 kHz, the largest difference between the characteristic impedances is observed when the heart is in systole and diastole. As the frequency increases, the characteristic impedance of the other model tissues decreases faster than the blood impedance, resulting in a smaller variation in characteristic impedance for the different cardiac states.
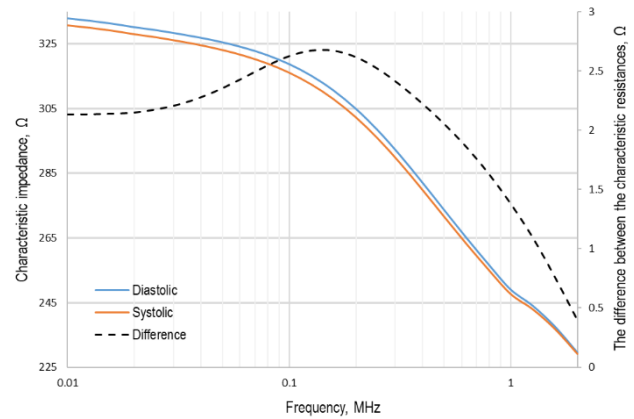
Fig. 7. Change in characteristic impedance with respect to the artery diameter and frequency, when the glucose concentration in the blood is 4 mmol/l
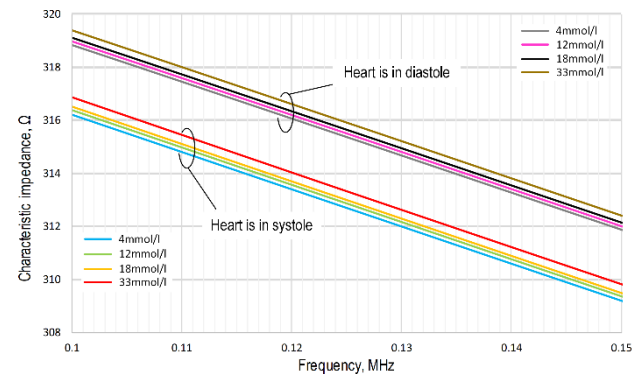
Fig. 8. Change in characteristic impedance as the artery diameter, frequency and glucose concentration vary

Since the overall model impedance is more dependent on frequency than on blood glucose concentration to better visualise the dependence of the characteristic impedance on glucose concentration and frequency, the results are presented in a narrower frequency range from 100 kHz to 150 kHz. This dependence is shown in Fig. 8. In Fig. 8, two groups of curves can be distinguished. The red, orange, green and blue curves represent the results of the characteristic impedance when the heart is in systole, while the brown, black, pink and grey curves represent the results when the heart is in diastole. In systole, the blood glucose concentrations were chosen as follows: red – 33 mmol/l, orange – 18 mmol/l, green – 12 mmol/l and blue – 4 mmol/l. In diastole, the blood glucose concentrations were chosen as follows: brown – 33 mmol/l, black – 18 mmol/l, pink – 12 mmol/l and grey – 4 mmol/l. As seen in the graph, different blood glucose concentrations lead to changes in the overall model's characteristic impedance. With varying glucose concentrations and different cardiac states (systole and diastole), the characteristic impedance varies throughout the frequency range. Thus, the blood glucose concentration is directly proportional to the characteristic impedance of the entire model of the human wrist current.

To investigate the dependence of the difference between the characteristic impedance of the heart in systole and diastole on glucose concentration, the values of characteristic impedance were examined at 125 kHz. This frequency was chosen because it is the closest frequency that our used hardware is capable of measuring, and it is close to the frequency of 150 kHz, where the largest difference between the characteristic impedance of the wrist in systole and diastole is observed. This dependence is shown in Fig. 9.

From the graph, it can be observed that the difference in the characteristic impedance of the whole model is inversely proportional to the glucose concentration in the blood. When the glucose concentration is 4 mmol/l, the difference in characteristic impedance is 2.674 Ω. When the glucose concentration is 12 mmol/l, the difference between the characteristic impedances in the model, in systole and diastole, is 2.645 Ω. When the glucose concentration is 18 mmol/l, the difference in impedances is 2.636 Ω. When the glucose concentration is 33 mmol/l, the difference in impedances is 2.578 Ω. These calculated results confirm the previous assumption that the difference in characteristic impedances between systole and diastole decreases with increasing glucose concentration in the blood.
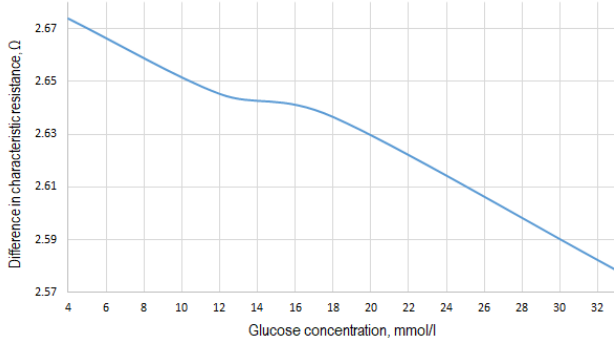
Fig. 9. Dependence of the difference between the characteristic impedance of the wrist in systole and diastole on glucose concentration, when the frequency is 125 kHz

4.2. Characteristic impedance results of the physical experiment

The equipment for measuring the characteristic impedance and the procedure of measurements and the entire experimental study are described in the previous section. For each 5-min measurement, the values of the characteristic impedance in systole were subtracted from the values in diastole. The average difference was then calculated. The measured glucose concentration values and the average values of the impedance differences are presented in Fig. 10.

The blue curve (Fig. 10) represents the measured glucose concentration using the commercial glucometer CONTOUR PLUS ELITE. The orange curve represents the measured difference in characteristic impedance between systole and diastole. From the graph in Fig. 11, it can be observed that as the glucose concentration in the blood increases, the difference in characteristic impedance decreases. Conversely, as the glucose concentration decreases, the difference in characteristic impedance increases. At the initial time, the measured glucose concentration was 4.5 mmol/l, and the difference in characteristic impedance was 67

mΩ. After 30 min of consuming 75 g of glucose, the glucose concentration in the blood increased to 8.1 mmol/l, and the difference in characteristic impedance at that time was 43.4 mΩ. After 60 min from the start of the test, the glucose concentration in the blood decreased to 6.2 mmol/l, and the average difference in characteristic impedance increased to 60.7 mΩ. At the final time, after 90 min from glucose consumption, the glucose concentration in the blood dropped to 4.8 mmol/l, and the difference in characteristic impedance increased to 64.4 mΩ. The measured values of the difference in characteristic impedance and glucose concentration are presented in Tab. 1.

From the conducted test, it is evident that the difference in characteristic impedance between systole and diastole has an opposite relationship to the difference in characteristic impedance. The dependence of the difference in characteristic impedance on glucose is depicted in Fig. 11.
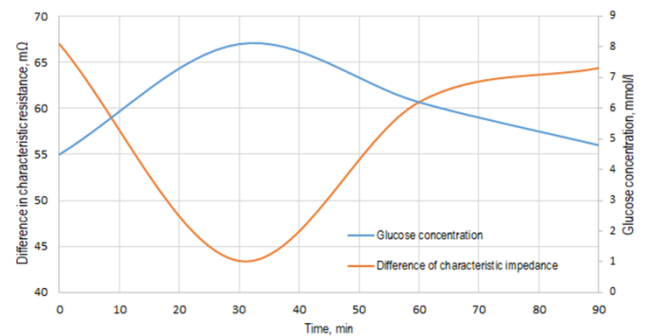
Fig. 10. Measurement of glucose concentration in the blood and difference in bioelectric impedance

Tab. 1. Values of measured glucose concentration in blood and the differences in characteristic impedance in systolic and diastolic states

Time (min)	Glucose concentration (mmol/l)	Differences in characteristic impedance (mΩ)
0	4.5	67
30	8.1	43.4
60	6.2	60.7
90	4.8	64.4
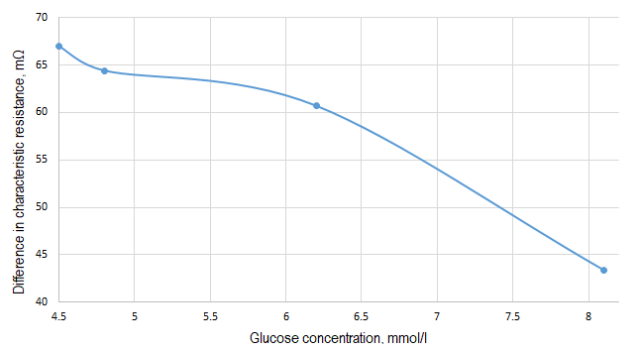
Fig. 11. Dependencies of the differences in the characteristic impedance when the heart is in systolic and diastolic states on blood glucose concentration

The (Fig. 11) shows that the relationship between the difference in characteristic impedance and glucose concentration is not linear. Due to this nonlinearity, it may be challenging to accurately

sciendo

Darius Plonis, Edas Kalinauskas, Andrius Katkevičius, Audrius Krukonis
Non-Invasive Analysis of the Bioelectrical Impedance of a Human Forearm

DOI 10.2478/ama-2024-0053

calculate the exact glucose concentration in the blood based on the measurement of the difference in characteristic impedance between systole and diastole.

5. DISCUSSION

In the study, the characteristic impedance of the entire forearm was measured under real conditions and by creating a forearm computer model and measuring its impedance using the ANSYS HFSS software package. Despite using the same electrode size and arrangement in both the software package and the real measurements, a significant difference in characteristic impedance was observed. Through calculations with the software package, the characteristic impedance of the entire forearm varied from 315.8 Ω to 312.6 Ω at a frequency of 125 kHz, depending on the different states of the heart and different glucose concentrations in the blood. When measuring the characteristic impedance under real conditions, the impedance of the entire forearm ranged from 150.1 Ω to 151.3 Ω due to low-frequency noise caused by respiration and different glucose concentrations in the blood, as well as different states of the cardiac cycle.

As can be seen, the measurements differ significantly. This difference arises because not all tissues present in the human forearm were included in the model. The forearm contains tissues that are more conductive to electrical current than fat, bone and skin. Due to the presence of these more conductive tissues, the actual characteristic impedance of the entire forearm is lower than what was calculated using the software package. Tissues such as tendons, arterial walls, extracellular fluids and other tissues were not represented in the model.

The measurements also differ because the proportions of muscle, fat, artery, bone and skin tissues vary between the measured individual and the model. In reality, the individual may have more blood vessels, which are more conductive to electrical current, than what was included in the designed model. Additionally, the forearm not only contains the radial artery but also the ulnar artery, which is smaller in size but still a good conductor of electricity.

When performing calculations with the ANSYS HFSS software package, it was observed that the model's characteristic impedance varied significantly when the heart was in the systolic and diastolic states, with a maximum change occurring at a blood glucose concentration of 4 mmol/l. At this glucose concentration, the model's characteristic impedance ranged from 312.57 Ω to 315.25 Ω. The difference between these two impedances was 2.67 Ω, corresponding to a 0.85% variation from the maximum characteristic impedance value.

During real measurements using the MAX30001 hardware, it was observed that the characteristic impedance varied, on average, by about 67 mΩ at a glucose concentration of 4.5 mmol/l. The average measured characteristic impedance over a 5-min period was 151.21 mΩ at a glucose concentration of 4.5 mmol/l. This variation in characteristic impedance corresponds to a 0.04% change from the average measured characteristic impedance.

When measuring the results under real conditions, much smaller variations in characteristic impedance were obtained than those in the calculations using ANSYS HFSS software. Although the variations in characteristic impedance due to the heart rate differ significantly, both the theoretical calculations and the actual measurements showed the same dependency of characteristic impedance on the heart rate. The significant discrepancy between

the results obtained in the software and the measurements under real conditions could be attributed to the fact that not all tissues present in the forearm were included in the model. Tissues such as tendons, arterial walls, extracellular fluids and veins were not included in the model. These tissues are more conductive to electrical current than skin, fat or bones, but their characteristic impedance does not depend on the heart rate. Therefore, in real conditions, the electrical current distributes through more electrically conductive tissues, and the variation in characteristic impedance of the radial artery has a smaller impact on the overall characteristic impedance of the forearm.

The results obtained from the calculations using ANSYS HFSS software showed that the variation in the model's characteristic impedance, when the heart is in systole and diastole, is inversely proportional to the glucose concentration in the blood. When the measurement frequency of the characteristic impedance was set to 125 kHz, and the glucose concentration was 4 mmol/l, the variation in the characteristic impedance due to the heartbeat was 2.67 Ω. When the glucose concentration reached 12 mmol/l, the variation in the characteristic impedance was 2.65 Ω. For a glucose concentration of 18 mmol/l, the impedance variation was 2.64 Ω. When the glucose concentration reached its highest level of 33 mmol/l, the impedance variation was 2.58 Ω.

During the measurements under real conditions, although the measured glucose concentrations differed from the concentrations used in software, the same dependence of the characteristic impedance variation was observed as in the calculations with software. When the measured glucose concentration was 4.5 mmol/l, the variation in the characteristic impedance was 0.067 mΩ. At a glucose concentration of 4.8 mmol/l, the measured variation in the characteristic impedance due to the heartbeat was 64.6 mΩ. With an increase in glucose concentration to 6.2 mmol/l, the variation in the characteristic impedance was 60.7 mΩ, and when the maximum glucose level of 8.1 mmol/l was reached, the variation in the characteristic impedance was 64.4 mΩ.

6. CONCLUSIONS

After conducting a literature review on the characteristic impedance of the human body, the electrical model of the human body was analysed. The main parameters of the body's characteristic impedance were reviewed, and an overview of different measurement methods for the characteristic impedance was conducted.

In the scope of this study, a computer model of the human wrist was developed using the ANSYS HFSS software package. Additionally, measurements of the bioelectrical characteristic impedance were performed with different glucose concentrations in the blood. The characteristic impedance of the wrist was also measured using the MAX30001 integrated circuit at various glucose concentrations in the blood, and the measured characteristic impedance signals were processed programmatically.

The measurements using the ANSYS HFSS software package revealed that the characteristic impedance of the wrist is inversely proportional to the glucose concentration in the blood. Using the software package, the characteristic impedance of the wrist was also calculated when the heart is in systole and diastole, resulting in changes in the diameter of the radial artery in the wrist. The difference in characteristic impedance between the maximum and minimum diameter of the radial artery was calculated, and the dependence of this difference on the glucose con-

centration in the blood was examined. The results showed that the difference in characteristic impedance is inversely dependent on the glucose concentration in the blood.

The measurements conducted using the MAX30001 integrated circuit confirmed the results obtained from the computer simulation. The largest difference in impedance between systole and diastole was observed at 150 kHz. By performing calculations with different glucose concentrations in the blood, it was observed that the difference in characteristic impedance is inversely proportional to the glucose concentration in the blood. All measurements were performed in the frequency range from 10 kHz to 2 MHz. The glucose concentration in the model artery varied from 4 mmol/l to 33 mmol/l.

The measured values of the model and real wrist's characteristic impedance revealed a significant difference, which arose due to the absence of all tissues in the computer model created with ANSYS HFSS. In future work, there are plans to improve the computer model of the human wrist by including as many tissues as possible that are actually present in the human wrist.

The study conducted using ANSYS HFSS software showed that the wrist's characteristic impedance is inversely proportional to the frequency of the alternating current, and the model's characteristic impedance is higher when the heart is in the diastolic state.

In this feasibility study on non-invasive bioelectric impedance analysis, we recognised that changes in physical parameters of the human hand can lead to changes in measurements. However, since the main objective of the study was to determine the technical effectiveness of the method, a complete study of these variations was not part of the scope. Recognising the importance of individual differences in hand anatomy, we believe that in the future, more comprehensive research will take these parameters into account to ensure robust and accurate application of non-invasive biomechanical analysis.

REFERENCES

1. Yu Y, Anand G, Lowe A, Zhang H, Kalra A. Towards estimating arterial diameter using bioimpedance spectroscopy: a computational simulation and tissue phantom analysis. Sensors 2022;22(13):4736. https://doi.org/10.3390/s22134736

2. Jacob J, Ishaan B, Onyezili F. Diabetes induction with streptozotocin and insulin action on blood glucose levels in albino rats. International Journal of Modern Science and Technology. 2021;3(10):208–212.

3. Eyth E, Basit H, Swift CJ. Glucose tolerance test. StatPearls [Internet] 2023 Apr [cited 2023 Jul 18].
Available from: https://www.ncbi.nlm.nih.gov/books/NBK532915/

4. Kushner RF. Bioelectrical Impedance Analysis: A Review of Principles and Applications. The American Journal of Clinical Nutrition. 1992;11(2):199-209.
https://doi.org/10.1080/07315724.1992.12098245

5. Patterson R. Body Fluid Determinations Using Multiple Impedance Measurements. IEEE Engineering in Medicine and Biology Magazine. 1989;8(1):16–18. https://doi.org/10.1109/51.32399

6. Anamika P, Mukesh R. Bioimpedance analysis of vascular tissue and fluid flow in human and plant body: A review. Biosystems Engineering. 2020; 197: 170–187.
https://doi.org/10.1016/j.biosystemseng.2020.06.006

7. Cpindean R, Holonec R, Dragan F, Muresan C. Method for Body Impedance Measuremen. 6th International Conference on Advancements of Medicine and Health Care through Technology. 2018:17–20. Cluj-Napoca, Romania. IFMBE Proceedings 71. Springer, Singapore. https://doi.org/10.1007/978-981-13-6207-1_13

8. Critcher S, Freeborn TJ. Multi-Site Impedance Measurement System based on MAX30001 Integrated-Circuit. 2020 IEEE 63rd International Midwest Symposium on Circuits and Systems (MWSCAS). Springfield, MA, USA. 2020:381–386.
https://doi.org/10.1109/MWSCAS48704.2020.9184451

9. Ghosh S, Meister D, Cowen S, Hannan JW, Ferguson A. Body composition at the bedside. Eur. J. Gastroenterol. Hepatol. 1997; 9(8): 783–788. https://doi.org/10.1097/00042737-199708000-00009

10. Nuñez C, Gallagher D, Visser M, Pi-Sunyer FX, Wang Z, Heymsfield SB. Bioimpedance analysis: Evaluation of leg-to-leg system based on pressure contact footpad electrodes. Med Sci Sports Exerc. 1997; 29(4):524–531. https://doi.org/10.1097/00005768-199704000-00015

11. Hoffer EC, Meador CK, Simpson DC. Correlation of whole-body impedance with total body water volume. J. Appl. Physiol. 1969; 27(4):531–534. https://doi.org/10.1152/jappl.1969.27.4.531

12. Bera TK. Bioelectrical Impedance Methods for Noninvasive Health Monitoring: A Review, Journal of medical engineering. 2014:381251. https://doi.org/10.1155/2014/381251

13. Moonen HPFX, Van Zanten ARH. Bioelectric impedance analysis for body composition measurement and other potential clinical applications in critical illness. Current Opinion in Critical Care. 2021 27(4):344–353. https://doi.org/10.1097/MCC.0000000000000840

14. Utter AC, Nieman DC, Ward AN, Butterworth DE. Use of the leg-to-leg bioelectrical impedance method in assessing body-composition change in obese women. The American Journal of Clinical Nutrition. 1999;69(4):603–607. https://doi.org/10.1093/ajcn/69.4.603

15. Heymsfield SB, Nuñez C, Testolin C, Gallagher D. Anthropometry and methods of body composition measurement for research and field application in the elderly. European Journal of Clinical Nutrition. 2000;54(3):526–532. https://doi.org/10.1038/sj.ejcn.1601022

16. Patel RV, Peterson EL, Silverman N, Zarowitz BJ. Estimation of total body and extracellular water in post-coronary artery bypass surgical patients using single and multiple frequency bioimpedance. Critical Care Medicine Journal. 1996;24(11):1824–1828.
https://doi.org/10.1097/00003246-199611000-00011

17. Olde Rikkert MG, Deurenberg P, Jansen RW, Van't Hof MA, Hoefnagels WH. Validation of multifrequency bioelectrical impedance analysis in detecting changes in geriatric patients. Journal of the American Geriatrics Society. 1997;45(11):1345–1351.
https://doi.org/10.1111/j.1532-5415.1997.tb02934.x

18. Kyle UG, Genton L, Karsegard L, Slosman DO, Pichard C. Single prediction equation for bioelectrical impedance analysis in adults aged 20–94 years. Nutrition. 2001;17(3):248–253.
https://doi.org/10.1016/s0899-9007(00)00553-0

19. Kotler DP, Burastero S, Wang J, Pierson RNJr. Prediction of body cell mass, fat-free mass, and total body water with bioelectrical impedance analysis: effects of race, sex, and disease. The American Journal of Clinical Nutrition. 1996;64(3):489–497.
https://doi.org/10.1093/ajcn/64.3.489s

20. Deurenberg P, Tagliabue A, Schouten FJ. Multi-frequency impedance for the prediction of extracellular water and total body water. British Journal of Nutrition. 1995;73(3):349–358.
https://doi.org/10.1079/bjn19950038

21. Kushner RF, Schoeller DA. Estimation of total body water by bioelectrical impedance analysis. The American Journal of Clinical Nutrition. 1986;44(3):417–424. https://doi.org/10.1093/ajcn/44.3.417

22. Kyle UG, Bosaeus I, De Lorenzo AD, Deurenberg P, Elia M, Gómez JM, Heitmann BL, Kent-Smith L, Melchior JC, Pirlich M, Scharfetter H, Schols AM, Pichard C. Composition of the ESPEN Working Group. Bioelectrical impedance analysis--part I: review of principles and methods. Clinical Nutrition. 2004;23(5):1226–1243.
https://doi.org/10.1016/j.clnu.2004.06.004

23. Chamiot-Clerc P, Copie X, Renaud JF, Safar M, Girerd X. Comparative reactivity and mechanical properties of human isolated internal mammary and radial arteries. Cardiovascular Research. 1998;37(3): 811–819. https://doi.org/10.1016/S0008-6363(97)00267-8

24. Cho MC, Kim JY, Cho SH. A bio-impedance measurement system for portable monitoring of heart rate and pulse wave velocity using small body area, 2009 IEEE International Symposium on Circuits and Systems (ISCAS). 2009;3106–3109.
http://dx.doi.org/10.1109/ISCAS.2009.5118460

25. Li J, Igbe T, Liu Y, Nie Z, Qin W, Wang L, Hao Y. An approach for noninvasive blood glucose monitoring based on bioimpedance difference considering blood volume pulsation. IEEE Access. 2018;6: 51119–51129. https://doi.org/10.1109/ACCESS.2018.2866601

26. Karacolak T, Moreland RC, Topsakal R. Cole-Cole model for glucose-dependent dielectric properties of blood plasma for continuous glucose monitoring. Microwave and Optical Technology Letters. 2013;55(5):1160–1164. https://doi.org/10.1002/mop.27515

27. Bailon R, Sornmo L, Laguna P. ECG derived respiratory frequency estimation. Advanced Methods and Tools for ECG Data Analysis. 2006; 215–243.

Darius Plonis: https://orcid.org/0000-0001-6579-1526

Edas Kalinauskas: https://orcid.org/0009-0006-1608-4749

Andrius Katkevičius: https://orcid.org/0000-0002-1623-5643

Audrius Krukonis: https://orcid.org/0000-0002-4694-781X

APPLICATION OF CONVOLUTIONAL GATED RECURRENT UNITS U-NET
FOR DISTINGUISHING BETWEEN RETINITIS PIGMENTOSA AND CONE–ROD DYSTROPHY

Maria SKUBLEWSKA-PASZKOWSKA[*], **Pawel POWROZNIK**[*], **Robert REJDAK**[**], **Katarzyna NOWOMIEJSKA**[**]

*Faculty of Electrical Engineering and Computer Science, Department of Computer Science,
Lublin University of Technology, Nadbystrzycka 38D, 20-618 Lublin, Poland
**Faculty of Medicine, Chair and Department of General and Pediatric Ophthalmology,
Medical University of Lublin, Chmielna 1, 20-079, Lublin, Poland

maria.paszkowska@pollub.pl, p.powroznik@pollub.pl, robertrejdak@yahoo.com, katarzyna.nowomiejska@umlub.pl

Abstract: Artificial Intelligence (AI) has gained a prominent role in the medical industry. The rapid development of the computer science field has caused AI to become a meaningful part of modern healthcare. Image-based analysis involving neural networks is a very important part of eye diagnoses. In this study, a new approach using Convolutional Gated Recurrent Units (GRU) U-Net was proposed for the classifying healthy cases and cases with retinitis pigmentosa (RP) and cone–rod dystrophy (CORD). The basis for the classification was the location of pigmentary changes within the retina and fundus autofluorescence (FAF) pattern, as the posterior pole or the periphery of the retina may be affected. The dataset, gathered in the Chair and Department of General and Pediatric Ophthalmology of Medical University in Lublin, consisted of 230 ultra-widefield pseudocolour (UWFP) and ultra-widefield FAF images, obtained using the Optos 200TX device (Optos PLC). The data were divided into three categories: healthy subjects (50 images), patients with CORD (48 images) and patients with RP (132 images). For applying deep learning classification, which rely on a large amount of data, the dataset was artificially enlarged using augmentation involving image manipulations. The final dataset contained 744 images. The proposed Convolutional GRU U-Net network was evaluated taking account of the following measures: accuracy, precision, sensitivity, specificity and F1. The proposed tool achieved high accuracy in a range of 91.00%–97.90%. The developed solution has a great potential in RP diagnoses as a supporting tool.

Key words: retinitis pigmentosa, convolutional GRU U-Net, classification, UWFP, UWFAF, deep learning

1. INTRODUCTION

Retinitis pigmentosa (RP) is a group of inherited retinal diseases characterised by the progressive dysfunction of rod and cone photoreceptors in the retina. The majority of cases are inherited in Mendelian patterns, namely autosomal dominant (30%–40% of cases), autosomal recessive (50%–60%) or X-linked (5%–15%) inheritance. Depending on the type of cell primarily affected, RP can be categorised into rod-dominant (classical RP) and cone-dominant cone–rod dystrophy (CORD) [5]. According to RetNet [22], more than 300 causative genes have been identified for RP so far, of which 58 genes have been identified in families with autosomal recessive RP. It is the most common inherited retinal dystrophy globally, affecting 1 in 4,000 individuals [8, 13]. In the past two decades, extensive genetic studies on RP not only led to the identification of the molecular basis in at least 60% of families but also set the basis for gene-based therapy. Although new treatments for RP such as gene therapy are being developed [25], current practice mainly involves care for residual visual function and surgery or medical therapy for complications. Thus, an appropriate clinical evaluation and estimation method for residual visual function and structure in patients with RP should be established. Identifying potential anatomical biomarkers for disease progression in RP is highly relevant for assessing treatment end points in RP in clinical trials.

Artificial Intelligence (AI) has gained a prominent role in the medical industry. The rapid development of the computer science field has caused that AI to become a meaningful part of modern healthcare. Image-based analysis involving neural networks is a very important part of eye diagnoses [18]. These modern approaches allow us to find and localise the patterns that are characteristic of the particular disease. The neural network methods are the perfect tool to distinguish between healthy people and patients suffering from various disorders. Moreover, not only particular diseases but also their place of occurrence may be recognised.

The methods and system involving AI are additional support indicating the proper diseases or alerting when the specific features are located [18, 19, 21]. These methods may increase the detection of diseases, especially rare diseases, such as retina problems, in the areas where there are no specialists. The patients would gain the knowledge from AI. Moreover, AI will definitely shorten the consultation time.

Applying AI algorithms reduces the diagnosis time, supports decision-making and avoids misdiagnosis, and reduces the cost of treatment or provides the treatment recommendation [16].

Despite the undoubted advantages of existing solutions, there is still a niche in the quick and effective diagnosis of eye diseases, especially rare or genetic diseases. The main motivation of this study is to apply AI for the classification of eye changes caused by RP disease. It is very important to distinguish between RP and CORD.

The main aim of this study is to classify RP. The Convolutional Gated Recurrent Units (GRU) U-Net was created for this purpose. The dataset, gathered in the Chair and Department of General and Pediatric Ophthalmology of Medical University in Lublin, consisted of 230 ultra-widefield fundus photography (UWFP) and ultra-widefield FAF (UWFAF) images. For applying deep learning classification, which rely on a large amount of data, the dataset was artificially enlarged using augmentation involving image manipulations. Three classes were defined: healthy cases, cases with CORD and cases with RP on the periphery of an eye.

This study proposes that Convolutional GRU U-Net is an adequate tool for the classification of retina dystrophy and healthy cases.

The following research questions have been formulated:
- RQ1: Is it possible to achieve satisfactory classifier accuracy for the proposed the Convolutional GRU U-Net model?
- RQ2: Is it possible to distinguish RP from CORD and healthy cases from Optos images?
- RQ3: Is it possible to achieve satisfactory classifier accuracy while reducing the raining time?

2. RELATED WORKS

A great number of studies concern application of AI for medical purposes, including U-Net solutions.

2.1. RP

Patients with RP typically experience night blindness and progressive loss of peripheral vision in both eyes during their first decades of life [36]. This leads to the loss of central vision and progression to total blindness. Typical form of RP is a long-lasting disease that usually evolves over several decades. However, there are extreme cases with a rapid evolution over two decades or a slow progression that never leads to blindness. This long-lasting disease is classified into three stages: early, middle and end stages. In the early stage, night blindness is the main symptom. In the mid stage, visual field testing shows mild periphery scotomas that tend to enlarge towards extreme periphery and macular areas. In the end stage, patients can no longer move autonomously, as a result of peripheral vision loss (classical tunnel vision), with few degrees of remaining visual field around this the main diagnostic tool to assess the visual function.

Variable expression in fundus changes may be present in different patients with different causes, in different patients caused by mutations in the same gene, in different patients from the same family and probably even in different eyes of the same patient.

The classic ophthalmoscopic findings in RP are typically described as a triad: retinal blood vessel attenuation, waxy pallor of the optic disc and intraretinal pigmentation in a bone–spicule pattern in the periphery of the retina. Abnormal retinal pigmentation occurs when pigment from disintegrating retinal pigment epithelial cells migrates into the superficial ("inner") retina as a response to photoreceptor cell death [22]. Initially, the pigmentary abnormality appears as a fine dusting extending from the mid to the far peripheral retina. Later, "bone-spicules" form throughout the mid and the far retinal periphery due to accumulations of pigment surrounding retinal blood vessels. In advanced RP, atrophy of the choriocapillaris may expose the large choroidal vessels beneath. CORD presents with visual acuity loss, photophobia, and dyschromatopsia, showing abnormal cone photopic electroretinography (ERG) response with no or mild rod involvement and pigmentary changes within the posterior pole [12].

The diagnosis and monitoring of RP require a comprehensive approach, encompassing patient history, fundus examination with fundus wide-field photos, wide-field fundus autofluorescence (FAF), visual field and full-field ERG. These diagnostic tools play a crucial role in detecting early signs of RP in at-risk individuals and monitoring the progression of the disease.

There is a progress towards preventing the loss of or restoring the function of rod photoreceptors in retinal dystrophies. There is also a need for a standardised device for quantifying the loss of the structure and function of the retina in RP. The emergence of molecular therapies for inherited retinal degenerations has highlighted the need for imaging modalities that can sensitively identify structural disease progression and robust methods to analyse disease progression.

Both UWFP and UWFAF imaging using a scanning laser ophthalmoscope are imaging techniques that enable clinicians to obtain fundus images with a 200° angle of view easily and noninvasively [24]. FAF is a noninvasive method that evaluates retinal disease and reflects retinal pigment epithelium functions by visualising the accumulation of lipofuscin. FAF images of patients with RP show hyperfluorescence in the early disease stages, whereas hypofluorescence corresponds to lesions in later stages. In typical RP, an autofluorescence ring, which represents a hyperfluorescent ring in FAF images, might be observed at the border separating the dysfunctional retina from the functional retina [27, 30].

2.2. Tools for RP classification

Various appliances for improving the modern eye healthcare have been developed using AI. The new models have been created to classify various eye diseases, including RP, as a tool supporting diagnoses. Modern systems have also been developed for these purposes. Many types of images have been taken into consideration.

The identification of normal and selected diseases based on ultra-wide field imaging is described in a previous study [35]. Three deep learning networks, namely EfficientNet-B7, DenseNet and ResNet-101 were applied to classify diabetic retinopathy, retinal vein occlusion, pathologic myopia, retinal detachment, RP, age-related macular degeneration, vitreous opacity and optic neuropathy. The best accuracy was obtained for the EfficientNet-B7 network.

Image-based diagnosis using the Deep Neural Network (DNN) model Visual geometry group-16 (VGG-16) has been applied to classify patients suffering from RP disease and healthy patients [21]. Two types of images were involved: 373 UWFP and 373 UWFAF. The obtained results proved that the proposed model is efficient for RP recognition (UWFP: sensitivity 99.3%, specificity 99.1%, and UWFAF: sensitivity 100%, specificity 99.5%). It was also stated that further studies involving AI are required for RP diagnosis.

Glaucoma, maculopathy, pathological myopia and RP were classified using three deep learning approaches, namely MobileNetV2, InceptionV3 and AlexNet, in a previous study [14]. Due to small amount of data, transfer learning with ImageNet was applied. The Kaggle dataset was used for the studies. The best results obtained MobileNetV2: up to 98.4% for accuracy, 96% for sensitivity and 99% for specificity.

A previous study has used Convolutional Neural Network (CNN) for automatic recognition of healthy people and people with the diseased retina based on retinal fundus images [18]. The images were obtained from the Friedrich-Alexander University machine learning data repository and from the eye hospital in Bangalore, India. Both data were augmented to obtain the proper number of images for classification. The average accuracy was reached between 96.5 % and 99.7%.

Four pretrained neural networks, namely AlexNet, DenseNet-161, ResNet-50 and ResNet-152, were applied for infrared (IR) images, optical coherence tomography (OCT) images and the two combined [20]. The highest precision was obtained for the combination of IR and OCT images. The studies were performed on the Johns Hopkins University (JHU) dataset.

The end-to-end deep learning approach consisting of Dense-Net, Recurrent Neural Network and Fully Connected layer was applied for predicting treatment requirements for the management of neovascular age-related macular degeneration [28]. Three classes were defined as low, intermediate and high. The obtained accuracy was up to 72%, sensitivity was 82% and specificity was 71%.

The RP classification based on OCT images using pixel-based and component-based methods was described in a previous study [8]. Four classifiers, namely Random Forest pixel-based, AdaBoost.M1 pixel-based, Random Forest component-based and AdaBoost.M1 component-based, were applied with great success. The obtained accuracy exceeded 99%. The precision was in the range of 42.45%–68.40%, the recall between 52.2% and 79.4%, and F1 between 46.76% and 64.69%.

Classification methods were also applied for developing expert systems for various eye disease. An expert system, dedicated to diabetic retinopathy and RP, consisting of extracting colour channels, red and green, was used [10]. Then, the Sobel method for edge detection was applied for the red channel, and the region of interest was determined for the green channel. In the next step, pixels with high intensity from the red channel and low-intensity pixels from the green channel were merged. Finally, preprocessing included the masking and removal of unwanted regions, and pixels were interpolated to the original image. The data were obtained using a CANON CF-60UVi camera with a 60o FOV. The achieved accuracy was 85.45%.

Another expert system for RP classification was used in a previous study [7]. It aimed at the automatic segmentation of pigment deposits in the fundus of the eye images. OCT images and those obtained from fundus camera were preprocessed. Then, the watershed transformation was applied to obtain homogeneous regions. This made the proposed method highly resistant to high variabilities in pigment deposits in terms of colour and shape. Even small pigment deposits were recognised. The system accuracy, precision, recall and F1 score were 97.90%, 74.43%, 98.44% and 59.04%, respectively.

After analysing the studies concerning applying AI for RP diagnosis, the authors noticed that there are no studies involving the location of retinal changes based on Optos images. Therefore, the aim of this study is to propose a Convolutional GRU U-NET classifier to distinguish between a healthy eye and an eye with pigmentary changes located in the centre (CORD) or on the periphery of the retina (RP).

3. MATERIALS AND METHODS

3.1. Dataset

UWFP and UWFAF images have been obtained using an Optos 200TX device (Optos PLC). Medical records of patients with cone–rod and cone–rod dystrophies at the Chair and Department of General and Pediatric Ophthalmology of the Medical University of Lublin were retrospectively reviewed. Patients were included if they had UWFAF and clinical characterisation of their retinal dystrophy. UWFP and UWFAF imaging were performed after pupil dilation with topical 0.5% tropicamide. All patients underwent visual acuity testing, slit lamp biomicroscopy and dilated fundus examination. All patients had clinical findings consistent with RP or CORD. UW-FAF characteristics were analysed qualitatively by two reviewers. The study included 69 eyes of 35 patients. All patients were identified and diagnosed clinically by an inherited retinal disease specialist. The control group consisted of 15 healthy subjects (30 eyes). A total of 230 optomap retinal images were obtained. The data were divided into three categories: healthy subjects (50 images), patients with CORD (48 images) and patients with RP on the periphery of the retina (132 images). For applying deep learning classification, which requires a large amount of data, the dataset was artificially enlarged using augmentation involving image manipulations [38]. The images were rotated by 10°. Additionally, they were converted to greyscale. Previous study involving CNN and their various modifications, has proven how important role plays a properly selected and sufficiently large dataset. However, in studies using medical data, there is almost always a problem of shortage of training data due to various difficulties in obtaining data. Data augmentation is commonly used in such cases. As numerous studies indicate [1,2,6,32,37] that data augmentation improves the quality of classification, regardless of the measure used. In this study, the augmented dataset was exclusively utilised for training the model.

Deep learning models were employed for acquiring RP features and associated classifications from the complete input images. Automatically, irrelevant regions, including those beyond the circular boundary of the main object, were eliminated. These unimportant areas may arise from artefacts introduced during image acquisition. Subsequent to the removal of irrelevant portions, the images were resized to 300 × 375 pixels, with pixel values ranging from 0 to 1.

The final dataset, used in this study, contained the following number of images: 264 for healthy persons, 240 for patients with RP in the centre of the retina and 240 for patients with RP on the periphery of the retina. The images representing RP and CORD are shown in Fig. 1. In case of RP, the most typical changes in the fundus of the eye are narrow vessels and pigmented changes in the periphery of the fundus. The macular area is characterised by shiny reflexes. In case of CORD, the changes are located in the centre of the retina.

The final dataset, used in this study, contained the following number of images: 264 for healthy persons, 240 for patients with RP in the centre of the retina and 240 for patients with RP on the periphery of the retina. The images representing RP and CORD are shown in Fig. 1. In case of RP, the most typical changes in the fundus of the eye are narrow vessels and pigmented changes in the periphery of the fundus. The macular area is characterised by shiny reflexes. In case of CORD, the changes are located in the centre of the retina.
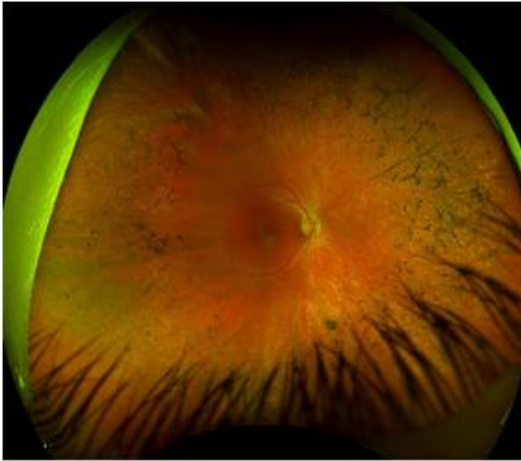
RP
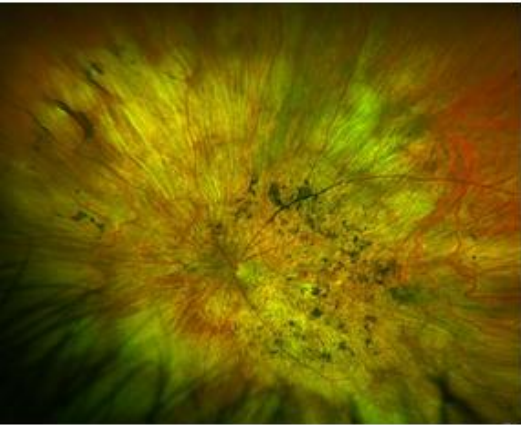
CORD

Fig. 1. Example images from dataset representing RP and CORD dystrophy. Wide-field fundus photography obtained using Optos device

3.2. GRU

The GRU model, widely recognised as a prominent variation of Long Short-Term Memory (LSTM) [13], combines the forgetting gate and input gate into a single update gate while also merging the cell state and hidden state. Consequently, the resulting GRU model is simpler and faster than the conventional LSTM model. This characteristic proves particularly advantageous when handling large datasets, allowing for substantial time savings with only minimal performance discrepancies compared to the standard LSTM model. Both LSTM and GRU models excel at preserving crucial features by using various gates, ensuring these features remain intact even during long-term transmissions. At time t, the new state can be calculated using Eq. (1) [9]:

$$u_t = (1 - z_t) \circ u_{t-1} + a_g \circ \widetilde{u}_t \tag{1}$$

where u_t is the new GRU state, u_{t-1} is the previous GRU state, a_g is the update gate and $\widetilde{u}_t$ is the current candidate sate with a new collection of information.

The update gate plays a crucial role in determining the balance between retaining past information and incorporating new information. Its primary function is to regulate the extent to which information from the previous state influences the current state. By adjusting the value of a_g, the degree to which the information from the previous state is integrated into the current state is con-

trolled. A higher value of a_g signifies a greater incorporation of information from the previous state. Eq. (2) defines the process of update gate modification [9]:

$$a_g = \delta(W_z x_t + R_z u_{t-1} + b_z) \tag{2}$$

where x_t is the data vector at time t, W_z is update gate weights at time t, R_z is update gate weights at time $t - 1$ and b_z is bias.

The current candidate state is calculated using Eq. (3) [9]:

$$\widetilde{u}_t = \tanh(W_h x_t + r_t \circ R_z u_{t-1}) \, b_z \tag{3}$$

where r_t is a reset gate at time t. Its role is to control the flow of information between current and previous states.

In numerous modelling tasks, having access to both past and future states proves advantageous. However, Conventional GRU networks process sequences in a temporal order, disregarding any future state. To address this limitation, Bidirectional GRU networks are introduced, which expand unidirectional GRU networks by incorporating a second layer. In this additional layer, the hidden-to-hidden connections flow in the opposite temporal order, enabling the model to capture the future state alongside the traditional past state. This bidirectional approach enhances the network's ability to understand and leverage information from both temporal directions.

3.3. Proposed methodology

This network is inspired from previous studies [3, 15, 29, 34]. The network utilises the strengths of Bidirectional GRU networks as well as densely connected convolutions. The individual components are described in detail in the following sections. The whole structure of it is presented in Fig. 2.

3.3.1. Encoding

The modified U-Net's contracting path involves a series of four steps. Each step comprises two convolutional filters of a size of 3×3, followed by a 2×2 max pooling function and the Rectified Linear Unit (ReLU) activation. At each step, the number of feature maps is doubled. The contracting path progressively extracts image representations while increasing the dimensionality of these representations layer by layer. Ultimately, the final layer in the encoding path produces a high-dimensional image representation with rich semantic information.

In the original U-Net, the last step of the encoding path consists of a sequence of convolutional layers. This enables the network to learn diverse types of features. However, this approach may result in the learning of redundant features through successive convolutions. To address this issue, densely connected convolutions were introduced [15]. This enhancement allows the network to improve its performance by leveraging the concept of "collective knowledge", which involves concatenating feature maps learned from all previous convolutional layers with the feature map obtained from the current layer. These concatenated feature maps are then forwarded as input to the subsequent convolutional layer. This strategy promotes the reuse of valuable feature maps throughout the network, mitigating the risk of learning redundant features.

The idea of densely connected convolutions has some advantages over regular convolutions [15]. First of all, it helps the network to learn a diverse set of feature maps, instead of redun-

dant features. Moreover, this idea improves the network's representational power by allowing information flow through the network and reusing features. Furthermore, densely connected convolutions can benefit from all the produced features before it, which prompts the network to avoid the risk of exploding or vanishing gradients. In addition, the gradients are sent to their respective places in the network more quickly in the backward path. The

idea of densely connected convolutions was implemented in the proposed network, due to which two consecutive convolutions are introduced as one block. There is a sequence of N blocks in the last convolutional layer of the encoding path, as shown in Fig. 3. These blocks are densely connected.
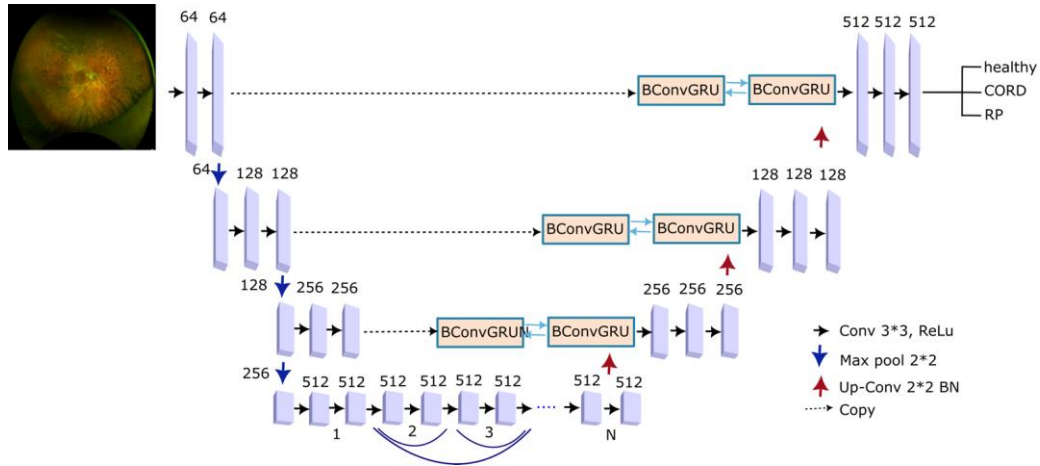
Fig. 2. The architecture of Convolutional GRU U-Net segment-based
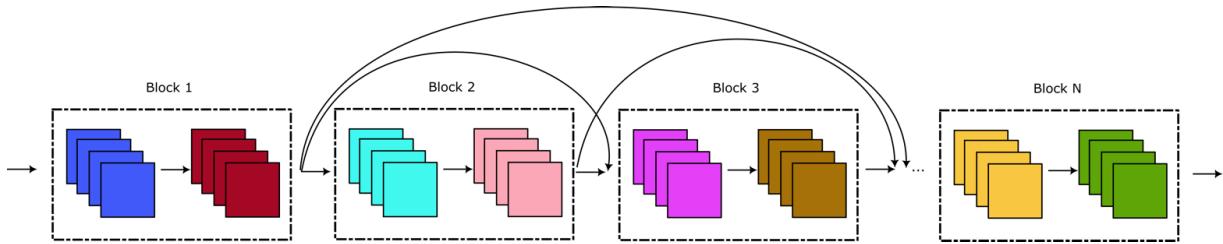
Fig. 3. U-Net dense layer

3.3.2.   Decoding

The decoding path begins by applying an up-sampling function to the output of the previous layer. In the traditional U-Net model, the corresponding feature maps from the contracting path are cropped and copied to the decoding path. These feature maps are then concatenated with the result of the up-sampling function. However, in this approach, the Bidirectional Convolutional GRU was utilised to process these two types of feature maps in a more

intricate manner (Fig. 4).

The set of feature maps copied from the encoding part is initially passed through an up-convolutional layer. This layer applies an up-sampling function, followed by a 2×2 convolution operation, effectively doubling the size of each feature map and halving the number of feature channels. Consequently, the resulting feature maps progressively increase in size throughout the expanding path, layer by layer, until they reach the original size of the input image after the final layer.
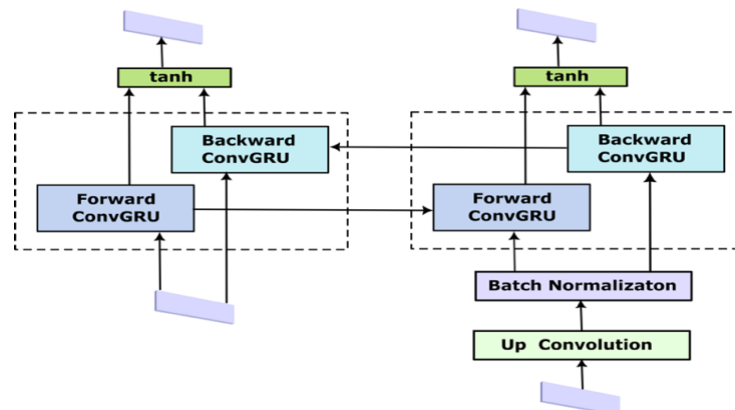
Fig. 4. Bidirectional ConvGRU

3.3.3. Batch normalisation

Following the up-sampling, the feature map is processed by a batch normalisation (BN) operation. During the training phase, intermediate layers often encounter a challenge where the distribution of activations tends to vary. This issue can significantly slow down the training process as each layer needs to adapt to new distributions in every training step. To address this problem, BN [17] is employed to enhance the stability of the neural network. BN standardises the inputs to a layer by subtracting the batch mean and dividing it by the batch standard deviation (SD). This normalisation process effectively accelerates the training speed of the neural network. Additionally, in certain cases, BN provides a regularisation effect, which can further improve the model's performance.

3.3.4. Bidirectional Convolutional GRU

The output of the BN step is fed to the Bidirectional Convolutional GRU layer. The main disadvantage of the standard GRU is that these networks does not take account of the spatial correlation since these models use full connections in input-to-state and state-to-state transitions. To solve this problem, bidirectional convolutional GRU was proposed, which exploited convolution operations. The idea of this layer based on LSTM was presented in a previous study [31]. It consists of an input vector x_t, an output vector u_t, a reset gate r_t, update gate a_g and candidate activation vector $\widetilde{u}_t$. The update rule for the input vector x_t and the previous output u_{t-1} is given by the following equations (for convenience, the subscript and superscript from the parameters were removed):

$$r = \sigma(W_{r*n}[u_{t-1}; x_t] + b_r) \tag{4}$$

$$a = \sigma(W_{u*n}[u_{t-1}; x_t] + b_u) \tag{5}$$

$$c = \rho(W_{c*n}[x_t; r \odot u_{t-1}] + b_c) \tag{6}$$

$$u_t = a \odot u_{t-1} + (1-a) \odot c \tag{7}$$

where σ and ρ are sigmoidal and ReLU functions, respectively; $*n$ denotes a convolution kernel of size $n \times n$; $\odot$ indicates Hadamard product operation. Brackets are presented to indicate a feature concatenation; and r, a, c and u_t denote the typical elements (reset, update, current memory content and a new state) of GRU.

4. RESULTS

4.1. Classifier evaluation

The proposed Convolutional GRU U-Net segmentation-based network was evaluated taking account of the following measures [4]: accuracy (Eq. 8), precision (Eq. 9), sensitivity (Eq. 10), specificity (Eq. 11) and F1 score (Eq. 12). A series of experiments were performed considering a random split of the data into training and testing sets: 80% and 20%, respectively. To reduce the randomness of the results, 10 iterations were performed independently. To assess the effectiveness of the classifier, commonly used measures were used: accuracy, precision, sensitivity, specificity and the F1 score [23,33].

Tab. 1. Accuracy results obtained using Convolutional GRU U-NET

Class	Mean (%)	Max (%)	Min (%)	±SD (%)
All	94.05	97.90	91.00	3.97

Tab. 2. Accuracy results of healthy, RP and CORD cases obtained using Convolutional GRU U-NET

Class	Mean (%)	Max (%)	Min (%)	±SD (%)
Healthy	95.48	97.91	92.14	3.86
CORD	94.87	97.98	92.07	4.89
RP	95.15	97.92	92.09	2.90

The accuracy results for the proposed Convolutional GRU U-Net for RP classification based on Optus images are provided in Tab. 1. This measure gives the information about the ability to distinguish between healthy cases and the cases with RP disease. The obtained values confirm that the segmentation method involving deep learning is a suitable tool for RP recognition purposes. The mean accuracy exceeds 94%, which proves that the classification RP has been developed with great success. The detailed accuracy results for three defined classes are presented in Tab 2. The proposed tool recognises heathy eyes, cases CORD and pigmentary changes localised on the periphery of the retina. The obtained minimum accuracy was greater than 92%. The maximum measures almost reached 98%.

Precision results for the developed tool is presented in Tab. 3. The mean value exceeds 95% for each class and ranges between 90% and 98%. The high value of this measure means that the network does many correct classifications or only few misclassifications.

Tab. 3. Precision results of healthy, RP and CORD cases obtained using Convolutional GRU U-NET

Class	Mean (%)	Max (%)	Min (%)	±SD (%)
Healthy	95.85	98.36	91.90	3.74
CORD	95.47	98.36	91.90	4.70
RP	95.32	98.41	90.00	3.90

Tab. 4. Sensitivity results of healthy, RP and CORD cases obtained using Convolutional GRU U-NET

Class	Mean (%)	Max (%)	Min (%)	±SD (%)
Healthy	94.71	97.56	91.54	3.47
CORD	93.41	96.82	90.27	3.82
RP	95.11	97.91	91.38	2.75

Sensitivity measures how correctly the model can classify positive instances. The obtained results, presented in Tab. 4, indicate that the developed tool has a great performance. The mean sensitivity exceeded 93%, while the values were in the range of 90.27%–97.91%.

Specificity measures the amount of real negative data that are classified as negative ones. The obtained values are summarised in Tab. 5. As it can be seen, the mean measure exceeds 92%, while the whole calculated indicator is between 89.64% and 99.17%. The higher specificity value was obtained for cases with RP diseases located in the centre of the retina.

Tab. 5. Specificity results of healthy, RP and CORD cases obtained using Convolutional GRU U-NET

Class	Mean (%)	Max (%)	Min (%)	±SD (%)
Healthy	95.42	98.36	92.28	3.61
CORD	98.46	99.17	95.36	3.27
RP	92.94	97.64	89.64	2.50

The harmonic mean of F1 scores was calculated based on precision and specificity. Based on the previous values of precision and specificity, it can be assumed that this measure also gained high values. The mean value exceeded 94%. The proposed tool recognises the cases with RP located in the central of the retina the best. Slightly worse results were obtained for the cases with RP located in the periphery of the retina and healthy ones. However, these differences are very small, up to 2.79%. The SD in all measures (Tabs. 1–6) is very low, which means that the observations are gathered around the mean and thus repeatable.

The learning parameters of the proposed tool are presented in Figs. 5 and 6. As can be seen from Fig. 5, after reaching 40 epochs, the accuracy, both for training and test sets, is on the same level (high). Another element indicating how well the neural network works is the loss value. As can be observed in Fig. 6, 50 epochs are enough to minimise the loss. These two parameters prove that the learning process was enough to obtain high accuracy values.
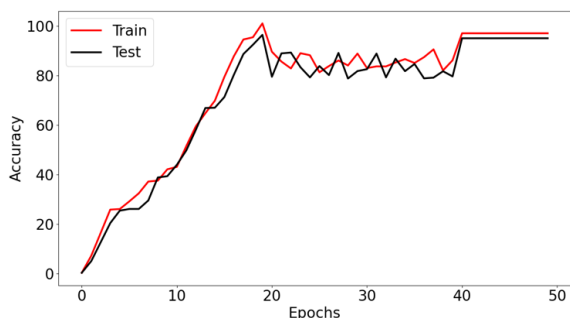
Fig. 5. Accuracy obtained for training and test data within 50 epochs. At least 40 epochs are required to obtain the accuracy exceeding 90%
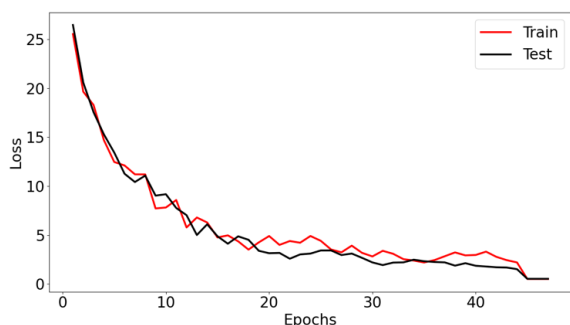
Fig. 6. Values of loss function while training and testing the model. It can be observed that almost 50 epochs are required to achieve the acceptable error level

The confusion matrix, depicted in Fig. 7, presents the classes that were misclassified. As it can be seen, two types of RP disease are confused, as well as the RP cases with healthy cases. The level of misclassification is very small, up to 5.14%. This

matrix also is also suitable for the neural network. To ensure the accuracy of the developed model, the Leave-One-Out Cross-Validation (LOOCV) was performed. Although this procedure is computationally intensive, it provides reliable and unbiased insights into the model's performance. By employing LOOCV, the root mean squared error (RMSE) for n tests was calculated. The obtained results for LOOCV were as follows: RMSE: 5.83% and SD ±3.78%.

Tab. 6. F1 score results for healthy, RP and CORD cases obtained using Convolutional GRU U-NET

Class	Mean (%)	Max (%)	Min (%)	±SD (%)
Healthy	95.63	98.36	92.10	3.65
CORD	96.94	98.76	93.99	3.35
RP	94.11	98.02	89.82	3.11
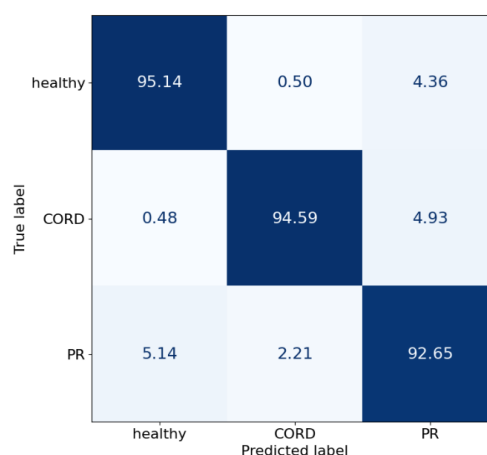
Fig. 7. Confusion matrix for three classes: healthy, CORD and RP cases. It can be seen that the proposed classifier has a little problem in distinguishing between healthy cases and RP, and between RP and CORD. CORD is classified correctly in many cases.

5. DISCUSSION

AI has gained a lot of interest recently, achieving success in various fields of study. In the medical field, many neural network models have been applied for solving challenging tasks concerning retinal diseases [7, 8, 10, 14, 20, 28, 35]. The accuracies depended on various classifiers and datasets consisting of different types of images.

In this study, the Convolutional GRU U-NET model was proposed for the classification of rare eye diseases of rod-dominant (classical RP) and cone-dominant CORD. These two groups were distinguished from healthy cases. A set of experiments was performed, which yielded satisfactory results, exceeding an accuracy of 90%, reaching up to 97.90%. These results of both the proposed model and individual classes confirm that the answer for RQ1 is positive.

The obtained performance results of the proposed the Convolutional GRU U-Net model for RP disorders were compared with the similar outcomes for various eye disease classifications using various networks. In Tab. 7, seven studies were gathered, including RP issues, that used different types of images, datasets and neural approaches. Various types of eye diseases were analysed. The highest accuracy was achieved by MobileNetV2, ImageNet,

Maria Skublewska-Paszkowska, Pawel Powroznik, Robert Rejdak, Katarzyna Nowomiejska
Application of Convolutional Gated Recurrent Units U-Net for Distinguishing between Retinitis Pigmentosa and Cone-Rod Dystrophy

DOI 10.2478/ama-2024-0054

InceptionV3 and AlexNet classifiers, while the lowest was achieved by a deep learning approach consisting of DenseNet, Recurrent Neural Network and Fully Connected layer. The accuracy results, obtained in this research, is at the top of the other accuracy values. This indicates that the proposed tool is adequate for the classification of RP disorders. Moreover, based on the state-of-the-art, it can be easily seen that this study made the first attempt to classify RP utilising the proposed neural network.

Analysing the results gathered in Tabs. 2–6, it can be indicated that the proposed classifier can recognise two types of retina dystrophies. The answer for RQ2is definitely affirmative. The proposed Convolutional GRU U-NET model achieved satisfactory results by reaching only 40 epochs, which also gives a positive answer to RQ3. The obtained results clearly indicate that the hypothesis is true and that the proposed Convolutional GRU U-Net can be successfully applied for rare eye diseases.

Tab. 7. Machine learning studies in retinitis pigmentosa (RP) recognition. OCT-optical coherence tomography, CORD-cone-rod dystrophy, UWFP- ultra-widefield pseudocolour, UWFAF-ultra widefield fundus autofluorescence

Classifier	Dataset	Diseases	Accuracy (%)	References
EfficientNet-B7, DenseNet, ResNet-101	574 UWFI images	Diabetic retinopathy, retinal vein occlusion, pathologic myopia, retinal detachment, RP, age-related macular degeneration, vitreous opacity and optic neuropathy	96.44–99.32	[35]
EfficientNet-B7	High-Resolution Fundus	Diabetic retinopathy and RP	85.45	[10]
Own system	Fundus images, OCT	RP	97.90	[7]
Random Forests pixel-based, AdaBoost.M1 pixel-based, Random Forests component-based, AdaBoost.M1 component-based	RIPS, OCT	RP	98.99–99.45	[8]
LCDNet	Fundus camera imaging	Retinal eye diseases	96.52–99.70	[20]
MobileNetV2, ImageNet, InceptionV3 and AlexNet	Kaggle	Glaucoma, maculopathy, pathological myopia and RP	100	[14]
Deep learning approach consists of DenseNet, Recurrent Neural Network and Fully Connected layer	OCT imaging	Neovascular age-related macular degeneration	72	[28]
Convolutional GRU U-Net	UWFP and UWFAF	Healthy, CORD, RP	91.00–97.90	Own

6. CONCLUSIONS AND FUTURE WORKS

Nowadays, the machine learning models have been applied for medical purposes. In this study, a new model was proposed to automatically classify the RP based on Optos images. Three classes were distinguished: healthy, CORD and RP. The Convolutional GRU U-Net was applied for this purpose. The verification of this tool proved that it is a very good solution to classify this rare disease. Each of the mean measures, namely accuracy, precision, sensitivity, specificity and F1 score, produced very high results, above 94%. The learning parameters showed that the dataset was enough to obtain high quality. Additionally, the proposed neural tool was verified using the LOOCV. The obtained results proved the high performance of the developed model. Moreover, according to the authors' knowledge, this is the first type of tool to classify the location of the RP in the retina. The obtained results proved the hypothesis.

The developed Convolutional GRU U-Net shows great potential in practical software implementation to diagnose RP diseases as a supporting tool. It may help to decrease the time of diagnoses. It also may give the recommendations for further treatment. However, this model also has some limitations. It was trained and verified on dataset containing Optors images. Neither temporal nor spatial relationships were taken into consideration. The model requires a high hardware memory capacity. Future studied may extend the classes to perform a more detailed analysis of this rare disease. The proposed neural model may be used for the as-

sessment of the level of the RP. The experiments will be performed focusing on improving the accuracy of the proposed model.

REFERENCES

1. Abeysinghe A, Tohmuang S, Davy J L, Fard M. Data augmentation on convolutional neural networks to classify mechanical noise. Appl. Acoust. 2023:203:109209.
2. Alomar K, Aysel H I, Cai X. Data augmentation in classification and segmentation: A survey and new strategies. J. Imaging. 2023; 9(2): 46.
3. Azad R, Asadi-Aghbolaghi M, Fathy M, Escalera S. Bi-directional ConvLSTM U-Net with densley connected convolutions. 2019. Proc - IEEE/CVF international conference on computer vision workshops.
4. Baratloo A, Hosseini M, Negida A, El Ashal, G. Part 1: Simple definition and calculation of accuracy, sensitivity and specificity. Emergency. 2015;3(2):48-49.
5. Berger W, Kloeckener-Gruissem B, Neidhardt J. The molecular basis of human retinal and vitreoretinal diseases. Prog Retin Eye Res. 2010;29(5):335–75.
6. Bonnici E, Arn P. The impact of Data Augmentation on classification accuracy and training time in Handwritten Character Recognition. Kth Royal Institute of Technology. 2021.
7. Brancati N, Frucci M, Gragnaniello D, Riccio D, Di Iorio V, Di Perna L. Automatic segmentation of pigment deposits in retinal fundus images of Retinitis Pigmentosa. Comput. Med. Imag. Graph. 2018;66:73-81.

8. Brancati N, Frucci M, Gragnaniello D, Riccio D, Di Iorio V, Di Perna L, Simonelli F. Learning-based approach to segment pigment signs in fundus images for retinitis pigmentosa analysis. Neurocomputing. 2018,308:159-171.

9. Chen JX, Jiang DM, Zhang YN, A hierarchical bidirectional GRU model with attention for EEG-based emotion classification, IEEE Access. 2019;7:118530-118540.

10. Das H, Saha A, Deb S. An expert system to distinguish a defective eye from a normal eye. Proc - 2014 International Conference on Issues and Challenges in Intelligent Computing Techniques (ICICT). IEEE. 2014:155-158.

11. Fahim AT, Daiger SP, Weleber RG. Nonsyndromic retinitis pigmentosa overview. 2017: Adam MP, Ardinger HH, Pagon RA, Wallace SE, Bean LJH, Stephens K, Amemiya A, eds.Gene Reviews. Seattle: University of Washington.

12. Gill JS, Georgiou M, Kalitzeos A, Moore AT, Michaelides M. Progressive cone and cone-rod dystrophies: Clinical features, molecular genetics and prospects for therapy. Br. J. Ophthalmol. 2019;103(5): 711-720.

13. Graves A, Mohamed AR, Hinton G. Speech recognition with deep recurrent neural networks. Proc - IEEE Int. Conf. Acoust., Speech Signal Process. 2013;38:6645–6649.

14. Guo C, Yu M, Li J. Prediction of different eye diseases based on fundus photography via deep transfer learning. J. Clin. Med. 2021;10(23):5481.

15. Huang G, Liu Z, Van Der Maaten L, Weinberger KQ. Densely connected convolutional networks. Proc - IEEE conference on computer vision and pattern recognition. 2017:4700-4708.

16. Hartong DT, Berson EL, Dryja TP. Retinitis pigmentosa. Lancet. 200618;368(9549):1795-809. doi: 10.1016/S0140-6736(06)69740-7

17. Ioffe S, Szegedy C. Batch normalization: Accelerating deep network training by reducing internal covariate shift. Proc - International conference on machine learning. 2015:448-456.

18. Jain L, Murthy H S, Patel C, Bansal D. Retinal eye disease detection using deep learning. Proc - Fourteenth International Conference on Information Processing (ICINPRO). IEEE. 2018:1-6.

19. LeCun Y, Bengio Y, Hinton G. Deep learning. Nature. 2015;521(7553):436-444.

20. Liu T Y A, Ling C, Hahn L, Jones C K, Boon C J, Singh M S. Prediction of visual impairment in retinitis pigmentosa using deep learning and multimodal fundus images. Br. J. Ophthalmol. 2022.

21. Masumot H, Tabuchi H, Nakakura S, Ohsugi H, Enno H, Ishitobi, N., Ohsugi E, Mitamura Y. Accuracy of a deep convolutional neural network in detection of retinitis pigmentosa on ultrawide-field images. PeerJ. 2019;7:e6900.

22. Merin S, Auerbach E. Retinitis pigmentosa. Surv. of Ophthalmol. 1976; 20(5):303-46. doi: 10.1016/s0039-6257(96)90001-6

23. Monaghan T F, Rahman S N, Agudelo C W, Wein A J, Lazar J M, Everaert K, Dmochowski R R. Foundational statistical principles in medical research: sensitivity, specificity, positive predictive value, and negative predictive value. Medicina. 2021;57(5):503.

24. Oishi A, Miyata M, Numa S, Otsuka Y, Oishi M, Tsujikawa A. Widefield fundus autofluorescence imaging in patients with hereditary retinal degeneration: a literature review. Int. J. of Retina Vitr. 2019;12(5) (Suppl 1):23. https://doi.org/10.1186/s40942-019-0173-z.

25. Piri N, Grodsky JD, Kaplan HJ. Gene therapy for retinitis pigmentosa. Taiwan J. Ophthalmol. 2021;11(4):348-351.

26. RetNet Retinal Information Network. https://sph.uth.edu/retnet/ [6.06.2023]

27. Robson AG, Egan CA, Luong VA, Bird AC, Holder GE, Fitzke FW. Comparison of FAF with photopic and scotopic fine-matrix mapping in patients with retinitis pigmentosa and normal visual acuity. Invest. Ophthalmol. Vis. Sci. 2004;45(11):4119-4125.

28. Romo-Bucheli D, Erfurth U S, Bogunović, H. End-to-end deep learning model for predicting treatment requirements in neovascular AMD from longitudinal retinal OCT imaging. IEEE J. Biomed. Health Inform. 2020;24(12):3456-3465.

29. Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation. Proc - 18th International Conference, Munich, Germany. October 5-9. Proceedings. Part III. 2015; 18:234-241. Springer International Publishing.

30. Schmitz-Valckenberg S, Holz FG, Bird AC, Spaide RF. Fundus autofluorescence imaging: review and perspectives. Retina. 2008;28(3):385-409.

31. Shi X, Chen Z, Wang H, Yeung DY, Wong WK, Woo WC. Convolutional LSTM network: A machine learning approach for precipitation nowcasting. Adv. Neural. Inf. Process. Syst. 2015;28.

32. Shorten C, Khoshgoftaar T M. A survey on image data augmentation for deep learning. J. Big Data. 2019;6(1):1-48.

33. Skublewska-Paszkowska M, Powroznik P. Temporal Pattern Attention for Multivariate Time Series of Tennis Strokes Classification. Sensors. 2023;23(5):2422.

34. Song H, Wang W, Zhao S, Shen J, Lam KM. Pyramid dilated deeper convlstm for video salient object detection. Proc - European conference on computer vision (ECCV). 2018:715-731.

35. Sun G, Wang X. Xu L. Li C. Wang W. Yi, Z., Luo H, Su Y, Zheng J, Li Z, Chen Z, Zheng H, Chen, C. Deep learning for the detection of multiple fundus diseases using ultra-widefield images. Ophthalmol. Ther. 2023;12(2):895-907.
https://doi.org/10.1007/s40123-022-00627-3

36. Tee JJ, Smith AJ, Hardcastle AJ, Michaelides M. RPGR-associated retinopathy: clinical features, molecular genetics, animal models and therapeutic options. Br J Ophthalmolol. 2016;100(8):1022-7. doi: 10.1136/bjophthalmol-2015-307698

37. Wong S C, Gatt A, Stamatescu V, McDonnell M D. Understanding data augmentation for classification: when to warp? Proc - International conference on digital image computing: techniques and applications. IEEE. 2016:1-6.

38. Yang S, Xiao W, Zhang M, Guo S, Zhao J, Shen F. Image data augmentation for deep learning: A survey. 2022. arXiv preprint arXiv:2204.08610.

Maria Skublewska-Paszkowska: https://orcid.org/0000-0002-0760-7126

Pawel Powroznik: https://orcid.org/0000-0002-5705-4785

Robert Rejdak: https://orcid.org/0000-0003-3321-2723

Katarzyna Nowomiejska: https://orcid.org/0000-0002-5805-8761

EXPERIMENTAL INVESTIGATION INTO THE TENSILE STRENGTH POST-REPAIR ON DAMAGED ALUMINIUM 2024 -T3 PLATES USING HYBRID BONDING/RIVETING

Abdelkrim MERAH*⬤**, Amin HOUARI**/*****⬤**, Kouider MADANI****⬤**, Mohamed BELHOUARI****⬤**,
Salah AMROUNE******⬤**, Ahmed CHELLIL*****⬤**, Cherif Zineelabidine YAHIA***⬤**, Raul D.S.G. CAMPILHO****/*******⬤

*Research Unit: Materials, Processes and Environment (UR/MPE), M'Hamed Bougara University of Boumerdes,
Cité Frantz Fanon, 35000 Boumerdes, Algeria
** Laboratory of Mechanics of Structures and Solids, LMSS, Mechanical Engineering Department,
Djillali Liabes University of Sidi-Bel-Abbes, BP 89 Sidi Bel Abbes 22000, Algeria
***Laboratory of Motor Dynamics and Vibroacoustic (LDMV), Mechanical Engineering Department,
M'Hamed Bougara University of Boumerdes, , Cité Frantz Fanon, 35000, Boumerdes, Algeria
****Laboratory of Materials and Structural Mechanics (LMMS), Mechanical Engineering Department,
Mohamed Boudiaf University-M'sila, BP 166 M'sila 28000, Algeria
****ISEP – School of Engineering, Polytechnic of Porto, Rua Dr. António Bernardino de Almeida, 431, 4200-072 Porto, Portugal
*****INEGI – Pólo FEUP, Rua Dr. Roberto Frias, s/n, 4200-465 Porto, Portugal

abdelkrimerah@univ-boumerdes.dz, a.houari@univ-boumerdes.dz, koumad10@yahoo.fr, belhouari@yahoo.com
salah.amroune@univ-msila.dz, a.chellil@univ-boumerdes.dz, yahiacherifzine@gmail.com, raulcampilho@gmail.com*

Abstract: Since the implementation of repair processes by composite patch bonding, this process has consistently demonstrated high performance across various industrial sectors, especially in the fields of aeronautics, aerospace and civil engineering. Consequently, there are situations in which the riveting process becomes the sole solution, particularly when the structure is subjected to severe mechanical or thermo-mechanical stresses, since adhesives have low mechanical strength after aging. Each method has its own set of advantages and disadvantages. The current trend is to combine these two processes to minimise their drawbacks as much as possible. The objective of this work is to present an experimental study on the repair of an aluminium plate AL2024-T3 with a central circular notch using a patch of different nature (metal or composite), under tensile loading conditions. The repair composite considered is a carbon/epoxide. The results of the tensile tests showed that the repair by the combination of the two processes improves the mechanical strength of the damaged structure. A comparison of the results of the experimental curves obtained on riveted, bonded and hybrid assemblies has been taken into consideration.

Key words: composite patch, aluminium AL2024-T3, bonding, riveting, hybrid process

1. INTRODUCTION

Throughout their service life, aeronautical structures are subjected to various mechanical stresses. When a geometric discontinuity, such as a crack or notch, is present, these mechanical loads can lead to a high concentration of stress, which may precipitate the premature failure of the structure. Currently, new techniques are being developed with the aim of delaying the onset of cracks and, in most cases, reducing the rate of crack propagation, thus increasing the lifespan of structures.

In the repair process, we can cite repair by welding [1], riveting [2], bolting [3] or other methods as in the study by Zhen-Yu Chen [4], and more recently, Elyas Ghafoor [5] used wire arc additive manufacturing (WAAM) to strengthen cracked steel components under fatigue loading using innovative techniques. The experimental results showed that this technique increases in the number of rupture cycles without crack propagation and minimises the effect of stress in the vicinity of the damage. The technology for repairing damaged structures by bonding a composite patch has progressed considerably and is currently being widely used,

particularly in aeronautics, due to the advantages it provides. Maintenance inspectors can recommend structural repair depending on the extent of damage, which should be the simplest and least expensive option for restoring the strength of the structure. The repair must not only provide resistance to ultimate loads but also ensure a long service life. Repairs (temporary or permanent) are carried out using bolted or riveted metal reinforcements or with bonded metal or composite reinforcements. This last solution is used in particular for laminated composite plates. Therefore, repairs are one of the issues that are the subject of studies on composite–metallic or composite–composite assemblies. The repair of a cracked structure can also be carried out by bonding an external patch to the structure to stop or slow down the propagation of the crack.

Composite materials are used in many fields because of their low weight, fatigue strength, corrosion resistance and enhanced damage tolerance [6]. The composite material patch plays an important role in the repair process due to its useful properties [7], and it is conventionally used in aeronautics for the repair of metal structures for damage such as cracks, notches or impacts.

The use of composite patches for the repair of damaged struc-

tures has interested several researchers, such as Baker et al. [8–9]. In an experimental study, Hosseini-Toudeshky et al. [10] investigated the growth of fatigue cracks on aluminium plates with a central mode I crack repaired using composite patches on one side only. On the other hand, Khalili et al. [11] conducted an experimental study on the effect of a composite patch applied to one side for repairing edge cracks in aluminium specimens using the Charpy impact test. Additionally, Maleki et al. [12] have studied the effect of applying a bonded composite patch to a cracked 2024-T3 aluminium plate under mixed mode loading. Basaid et al. [13] employed a fibreglass/epoxy patch for the maintenance of damaged plates in Air Algérie Maintenance workshops. Similarly, Gu et al. [14] used basic fibreglass/epoxy patches to repair aluminium plates with cracks and studied the development of cracks in the plate and the delamination behaviour in the patches under static stresses. Benkheira et al. [15] conducted an experimental analysis under tensile loading to investigate the effect of repair by using single and double patches of boron/epoxy-laminated composite plates with a central circular notch and also presented a comparison between these two repair modes. Nadia et al. [16] analysed the mechanical and failure behaviours of a damaged structure repaired by using a composite patch for defects in the adhesive layer. Madani et al. [17] carried out both experimental and numerical studies on the mechanical behaviour of several plates with notches of various shapes by using tensile tests. Aldeen, A et al. [18] studied the effect of isothermal and isochronous aging to investigate precipitate evolution and recrystallisation of zirconium alloy N36 after β-quenching.

Rivallant et al. [19] introduced a discrete 3D finite element method (FEM) that uses cohesive elements to simulate both inter-lamina delamination and intra-lamina matrix cracking. Similarly, R. Rashnooie et al. [20] successfully simulated crack growth in composite plates using an element method modelling approach extended finishes (XFEM); they took into consideration the propagation of damage in the adhesive layer, the different layers of the composite and the delamination of the metal–FRP interface. The proposed XFEM model simulates the fatigue behaviour of FRP-reinforced metal plates.

Ait Kaci et al. [21] have shown that a hybrid patch combining carbon fibre/epoxy and aramid/epoxy plies can reduce the stress in the damaged area and thus ensure the structure a long service life. Horn et al. [22] have shown that it is necessary to optimise the length and thickness of the repair patch and that these dimensions are important to increase the tensile strength of the repaired structure. The calculation of the stress distributions in the structure is therefore an important aspect in proposing an appropriate reinforcement solution.

Analysing stresses in the adhesive joint is essential to avoid deterioration of this layer as its mechanical properties are weaker than the plate and the patch. Madani et al. [23] analysed the stress distribution in an aluminium alloy 2024-T3 plate in the presence of a notch, repaired using a composite patch, through the FEM. The authors showed that the composite patch repair method greatly reduces the high stress concentration. Rezgani et al. [24] conducted experimental tests to analyse the effect of hydrothermal aging of the patch and the adhesive on the fatigue behaviour of a damaged 2024-T3 aluminium plate repaired by a carbon composite/epoxy patch. Wahrhaftig et al. [25] have proposed an equivalent stiffness system for calculating the minimum bending moment for concrete slender columns. Al-Abboodi et al. [26] have produced a device featuring a three-point curve test to evaluate the mechanical properties of a metallic glass alloy sample (Fe49.7 Cr17.1 Mn1.9 Mo7.4 W1.6 B15.2 C3.8 Si2.4) prepared by high-speed spark plasma sintering (SPS).

The weak point in reinforcing composite materials lies in the adhesive responsible for ensuring the adhesion of the reinforcement. According to reference [27], 53% of the observed failures in aeronautical structures repaired are due to the adhesive layer. These failures are essentially due to the transfer of loads from the adherend to the composite patch. This load transfer zone, in fact, results in a shear stress peak near the free edge of the composite patch. On the other hand, the nature of the adhesive joint has shown its effectiveness in absorbing and transferring the load from the damaged area. The adhesive used to bond the repair patch and the cracked plate together should also be prepared beforehand [28]. Maleki et al. [29] studied the failure of cracked aluminium plates repaired by one-sided glass/epoxy composite patches under fatigue loading. The acoustic emission technique was employed to monitor the effect of damage progression in the repair patch. Rivet patch repair involves placing a plate (metal or composite) over a damaged area and riveting the patch to the plate [30]. Riveting requires creating holes not only in the repair plate but also in the damaged structures. However, as composite materials are highly brittle, this operation introduces damage. Zitoune and Collombet [31] have shown that this damage can occur at the entrance, exit and periphery of the hole, creating delamination, fibre breakage and matrix degradation.

In the light of the previously mentioned studies, the novelty of our research is in its purely experimental approach to analyse tensile tests on damaged and repaired 2024-T3 aluminium plates. Employing various repair methods (riveting, bonding and hybrid repair) with three distinct types of patches, the study aims to compare their performance. It highlights that the presence of a patch does not always guarantee a significant improvement in the resistance of the damaged plate. Furthermore, it underscores specific findings, such as the potential undesirability of hybrid repair under certain conditions, and recommends the preferential use of a metallic patch in such scenarios. .

2. MATERIALS AND METHODS

In the present study, an analysis was conducted on an AL2024-T3 plate with a central circular notch. The composite patch includes two types: carbon/epoxy and fibreglass/epoxy (Fig. 1).

Fig. 1. Fibreglass and carbon fibre composite patches made under vacuum

The metal patch is of the same nature as the plate to be repaired. The composite materials used in this study were obtained from Air Algeria. They are mainly intended for aircraft repair, and the choice of the type of matrix and reinforcement is made according to the requirements of international aeronautical regulations. The composite plates and patches were fabricated from an eight-ply 300 gsm fabric. Polymer matrix composite materials are increasingly used in aeronautics due to their low mass. All the laminates were made up of eight plies (0°/90°/0°/90°/0°/90°/0°/90°) and had a nominal thickness of 1.86 mm. The fibre volume fraction was chosen according to ISO 1268-2 standards, with the range between 30% and 45%.

Fig. 2. (a) Resin EPOCAST 50-A1. (b) HARDENER 946

The adhesive consists of a homogeneous mixture of resin (EPOCAST 50-A1) and hardener (HARDENER 946 US), which is presented as a crosslinking agent (Fig. 2).

Epocast 50-A1 is a thermosetting resin. This matrix can be used for the manufacture or repair of composite structures in aeronautics. The product complies with the BMS 8-201 standard (Boeing Material Specification). This epoxy resin is of bisphenol type A (Fig. 2a) and is combined with a low-reactivity amine hardener (Fig. 2b). These resins pass successively from the liquid state to the gel state and then to the solid state. This characteristic process of thermosetting resins is called crosslinking.

3. EXPERIMENTAL STUDY

In the first part, experimental studies on metal patch repairs are presented (Fig. 3).
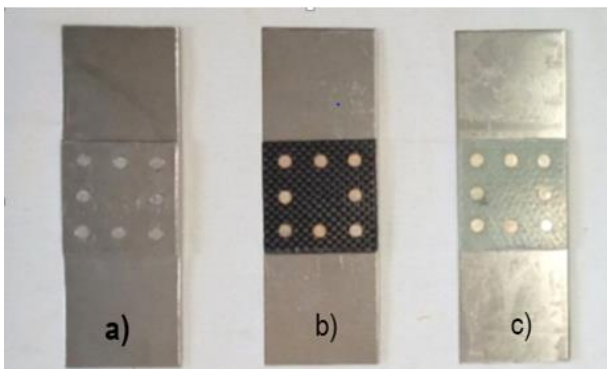
Fig. 3. Representation of specimens prepared for tensile tests: (a) With metal patch, (b,c) with composite patch

To assess the repair performance, static tensile tests were carried out on intact specimens (without the presence of geometric defects), then on specimens with a central circular notch (without repair) and finally on specimens repaired with riveted or bonded metal patches (Fig. 3a). In the second part, the damaged plate was repaired with a composite patch of the carbon/epoxy type or of the glass/epoxy type (Fig. 3b). Subsequently, the effect of an initial crack emanating from the notch on the degradation of the mechanical properties of the plate was assessed to repair this geometric defect. The considered crack has an initial length of a = 5 mm, and it was repaired using a bonded/riveted hybrid patch.

4. TEST SAMPLE CHARACTERISATION

The relevant parameters to consider in the repair process are the thickness of the adhesive (ta), length, width and shape of the patch. However, due to the difficulty of highlighting these parameters during the study, ta and the rivet diameter are kept constant. However, it is possible to remove one of these elements (adhesive or rivet) to assess the separate influence of the adhesive or the rivet.

The dimensions of the patch are fixed during this analysis. However, its nature is variable (aluminium patch, carbon/epoxy composite patch and glass/epoxy patch) to optimise the compatibility of the patch material with the number of rivets required to maintain the load transfer efficiency.
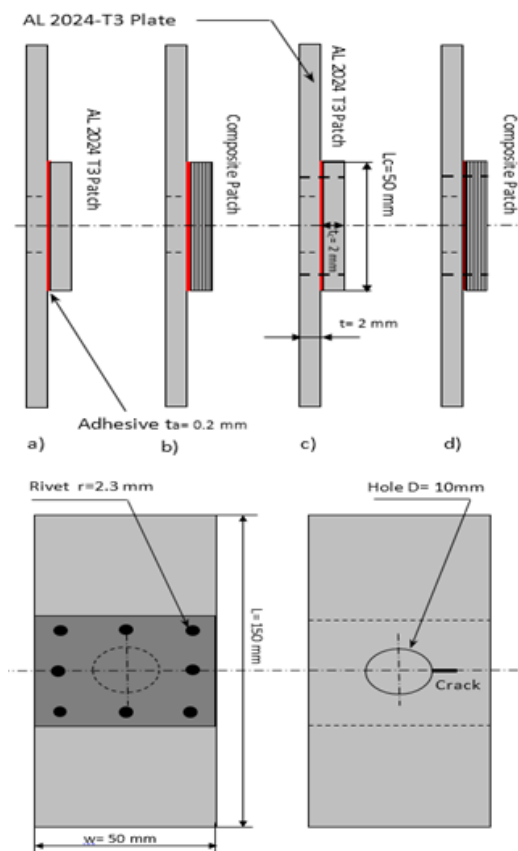
Fig. 4. Schematic representation of a plate repaired with (a) Metal patch by adhesive, (b) Composite patch by adhesive, (c) Metal patch by rivet/adhesive, (d) Composite patch by rivet/adhesive and (e) Plate with crack emanating from notch

Fig. 4 shows the geometry of the repaired structure with the different natures of the repair patch depending on the bonding or hybrid mode (bonding + riveting). It is considered that the plate has a central circular hole with a diameter of d = 10 mm (Fig. 4a). A crack emanating from a notch is considered in the fixed length study (Fig. 4e). The width w of the sample is also considered, which is important to assess the influence of the number of rivets on the stiffness and strength of a hybrid repair. The length of the adherend (the metal plate) is defined as L is shown in Tab. 2. Lc is the length of the patch, ta is the adhesive thickness, t is the adherend thickness, D is the hole diameter and d is the diameter of the fastener head of the fastener inside the hole.

In practice, to determine the diameter of the rivet (d_{rivet}) depending on the thickness of the sheets (Fig. 5), the following formula is used:

$$d_{rivet} = \frac{45H}{15+H} \ (mm)$$

where H represents the thickness of the repaired set [mm]. The length of the rivet (l) is defined as shown in Fig. 5.

The parameters of the samples with the NAS1399 C4-4 type rivet in the case of a plate repair with a laminated patch are shown in Fig. 5 and in Tab. 1. This type of rivet is used for all the experimental tests, due to the compatibility of its parameters with the repair conditions (Fig. 6).
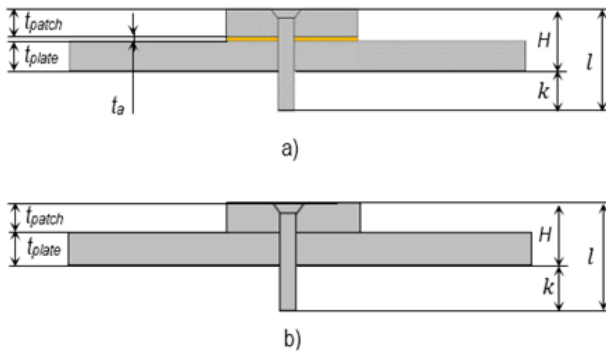
Fig. 5. Countersunk rivet dimensions in patch repair with/without bonding

Tab. 1. Dimensional parameters of a riveted sample

Rivet type	Rivet dimensions [mm]
Countersunk rivet	$l = 1.1.H + 0.6.d$
Stem free length	$k = (0.7–1.3).d$
Rod length	$l = \sum t + k$
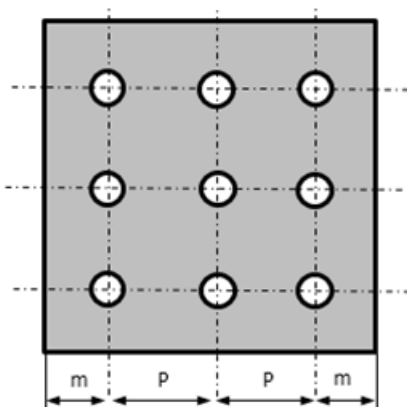
Fig. 6. Arrangement of rivet holes in the patch

In this work, we used the chain arrangement, in which the rivets are ordered and aligned with a respected spacing of the P value of 16.66 mm (Fig. 6).
− The pitch (P) of the rivet is defined as the distance between the centre of a rivet and the centre of the adjacent rivet in the same row (P = 5d).
− The margin (m) is the distance between the edge of the plate and the axis of the rivets of the nearest row (m = P/2).

However, in the case of bonded repair, the required dimensions of bonded specimens are shown in Fig. 5 and Tab. 3. The thickness of the adhesive is kept within 0.234 ± 0.025 mm.

Tab. 2. Dimensional parameters of a riveted joint

Adherend thickness t (mm)	4.15 ± 0.15
Lap length (mm)	50 ± 0.05
Sample width (mm)	50 ± 0.05
Rivet shank diameter d (mm)	3.2
Rivet head diameter D = 2d (mm)	6.2
Rivet shank length l (mm)	5 ± 0.25

Tab. 3. Dimensional parameters of a bonded joint

Adherend thickness t (mm)	4.28 ± 0.15
Overlap length (mm)	50 ± 0.25
Overlap width (mm)	50 ± 0.25
Adhesive thickness t_a (mm)	0.234 ± 0.025

5. PREPARING FOR PATCH REPAIR

The repair of the circular notch is carried out using an external patch (Fig. 7). Three repair modes were addressed, namely bonded patch, riveted patch and hybrid patch (Fig. 7). For the bonded patch, cleaning of the surface with acetone and the duration of crosslinking of the adhesive were taken carefully. For the hybrid repair, once the adhesive was put on the bonded area, the rivets were quickly put in their positions to avoid having hardened adhesive in the rivet holes. The crosslinking time was the same for both processes.
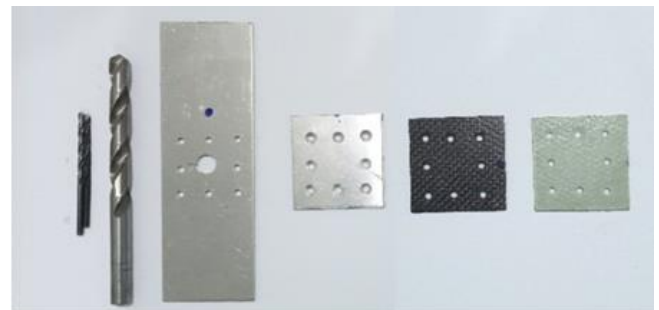
Fig. 7. Preparing for patch repair

The main step in preparing a repair using the riveting process consists of piercing the materials without damaging it for the aluminium and without inducing delamination for the composite.

6. DIFFERENT REPAIR CONFIGURATIONS

The performance of different repairs was analysed in this work. For this purpose, the samples used in the different repair modes have the same dimensions. For each repair mode (bonded, riveted and hybrid), three types of patches were considered (aluminium patch, fibreglass composite patch and carbon fibre composite patch) (Fig. 8). The three types of configuration are manufactured using the same materials, to obtain the most consistent results from one test to the next. The samples by repair type are shown in following Tab. 4 (Fig. 8). The three modes of repair are also considered for the specimens with a crack emanating from the notch.
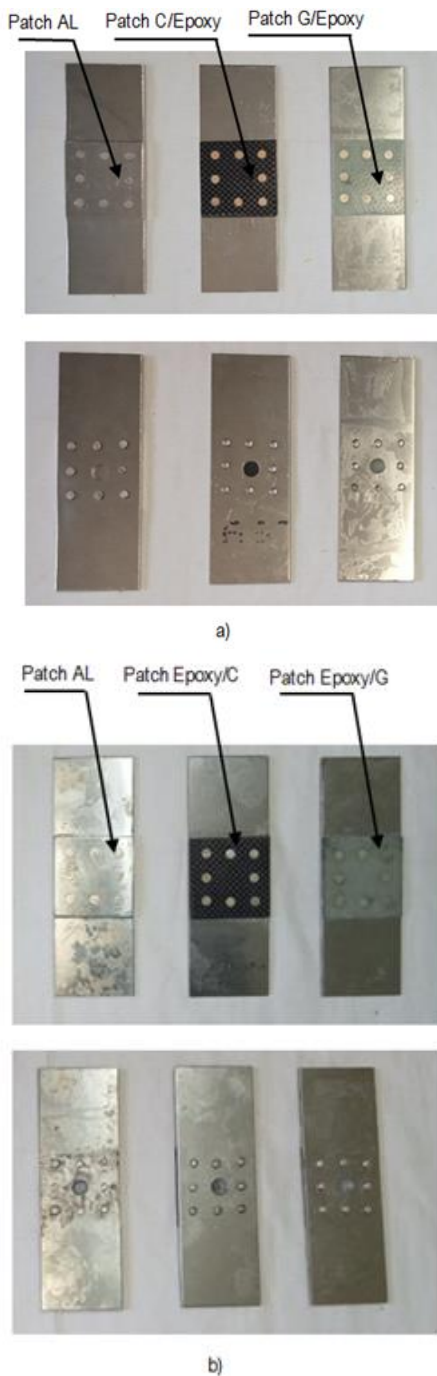
Fig. 8. Presentation of different configurations of repair. (a) riveted mode (b) hybrid mode

Tab. 4. Configuration of samples for tensile tests

Configuration	Plate	Damage type	Patch material nature	Repair mode
1	AL	Notch	Aluminium	Bonded
2	AL	Notch	Aluminium	Riveted
3	AL	Notch	Aluminium	Hybrid
4	AL	Notch	Composite carbon/epoxy	Bonded
5	AL	Notch	Composite carbon/epoxy	Riveted
6	AL	Notch	Composite carbon/epoxy	Hybrid
7	AL	Notch	Composite glass/epoxy	Bonded
8	AL	Notch	Composite glass/epoxy	Riveted
9	AL	Notch	Composite glass/epoxy	Hybrid
10	AL	Crack manating from notch	Aluminium	Hybrid
11	AL	Crack manating from notch	Composite carbon/epoxy	Hybrid
12	AL	Cracke manating from notch	Composite carbon/epoxy	Hybrid

The tensile tests were carried out at Coatings, Materials and Environment Laboratory (CMEL) at the University of M'hamed Bougara in Boumerdes, Algeria using a ZWITCK Z010 tensile machine (Fig. 9). The testing conditions were maintained at a temperature 23 ± 3°C and a relative humidity of 30 ± 10%, respectively, according to the ASTM D3039 and ASTM D3165 standards. (Fig. 9). The tensile machine was equipped with a 50 kN load cell and a crosshead drive system powered by an electric motor. The machine was controlled by software, which allowed the results of the loads and displacements to be recorded. The tests were conducted at a crosshead travel speed of 1 mm/min. A 25-mm-gauge length extensometer was used to obtain the displacement on the samples. Fig. 9 shows an example of extensometer positioning.

Fig. 9. Illustration of the test bench for tensile tests

It was necessary to carry out three tensile tests for each configuration. The various load–displacement curves obtained were processed to have reproducibility to better estimate the behaviour of each structure.

The main mechanical characteristics of the materials used (aluminium plates, composite patch and EPOCAST 50-A1/946 adhesive) for the numerical model are obtained from tensile tests (Fig. 10) and are grouped in Tab. 5.
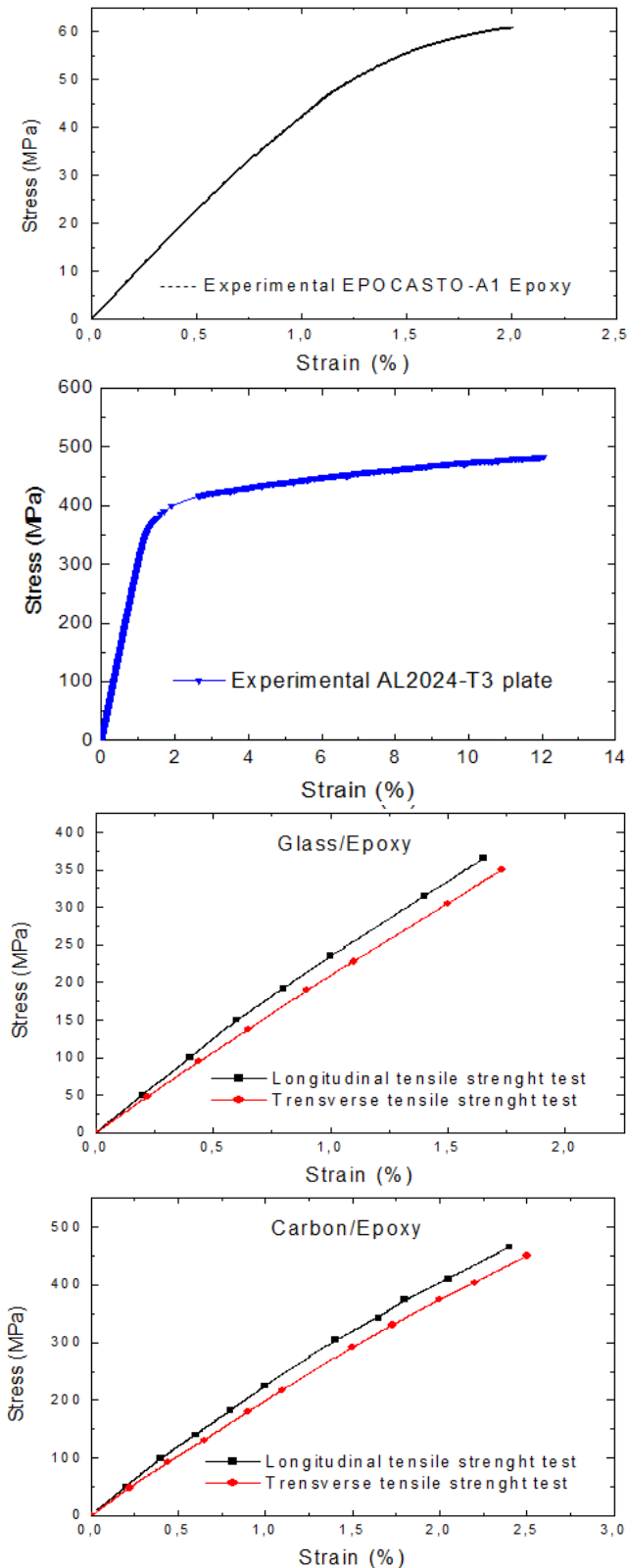
Fig. 10. Traction curves carried out on (a) EPOCASTO 50 -A1 adhesive, (b) aluminium plates, (c) glass/epoxy composite and (d) carbon/epoxy composite

Tab. 5. Mechanical properties of AL2024-T3, composite and resin EPOCASTO 50 -A1

Materials	AL 2024T	Ply CRFP	Ply GRFP	Resin
Young Modulus	74,70	20,40	17,90	34,60
Tensile strength	462	450	335	63.6
Yield stress (MPa)	311	-	-	-
Poison/ratio	0.33	0.3	0.3	0.3
Elongation (A%)	12 %	2.35	1.86	1.97 %

7. RESULTS

7.1. Effect on the nature of patches for different repair processes

The plates were subjected to a tensile load, and the load transferred by the rivets as well as by the adhesive was estimated. It is important to note that various factors can influence the result, including the bond quality, friction between the components (between the rivet head and the composite, and the deformed part of the rivet and the rivet shank and the interior of the hole) and the interaction between the rivet and the edge of the notch. The presence of notches in plate components leads to a reduction in their strength compared to the unnotched plate. It was found that these notches can significantly influence the expansion of damaged regions, especially when the rivet holes are located near the free edge of the plate. The tensile test results of notched and unnotched specimens are shown in Fig. 11, which combines the load–displacement curves for notched and unnotched specimens.
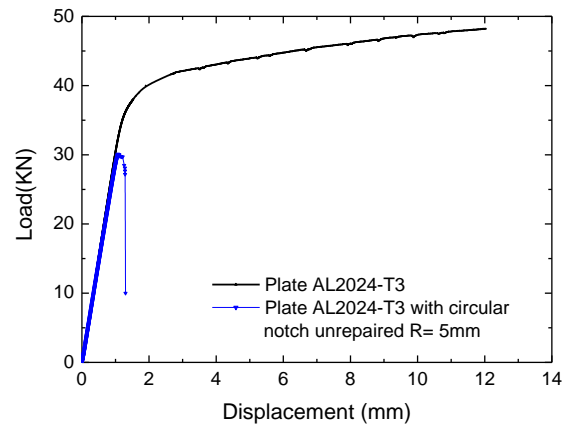
Fig. 11. Load–displacement curves for notched and unnotched specimens

The presence of a central notch with a radius of 5 mm weakens the material, leading to an approximately 30% reduction in its tensile strength. It is worth noting that the tensile curve of the plate without a hole comprises two parts. The first part corresponds to the elastic phase and exhibits a linear relationship. Subsequently, the second segment is nonlinear, displaying an alteration in the slope. This section is distinguished by a maximum stress featuring a plasticity threshold, succeeded by a phase of plastic flow with a very high tensile strength. The extensive plastic domain is a result of the material's ductility. However, the presence of a central circular notch reduces the nonlinear portion of the tensile curve.

The aluminium test specimen includes a 10-mm-diameter circular notch at the middle, followed by repairing using a single aluminium patch. The patch is riveted to the plate by eight aluminium rivets. Fig. 12 depicts the experimentally measured load variation with the applied displacement for a plate repaired with a single riveted patch, alongside an unrepaired plate. The advantageous impact of the patch is evident, with the loads at failure for the repaired specimens are notably higher than those at the unrepaired counterparts. A direct comparison with the unrepaired structure reveals that single patch repair techniques can enhance tensile strength by approximately 6%. On the other hand, for the plate with a central notch, the effect of the repair by riveting is almost negligible due to the presence of additional rivet holes adjacent to the main notch.
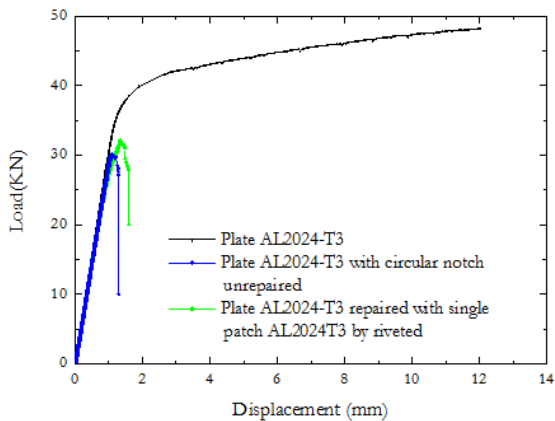
Fig. 12. Load–displacement plot for a plate with single riveted aluminium patch and an unrepaired plate

Currently, the aeronautical industry is interested in repairing structures by bonding external patches. The advantages of this method are related to the nature of the adherend, where the edge effects resulting from the drilling of the plates in the bolted or riveted repair prove to be very harmful to its mechanical strength. The predominant technique involves the repair of notched or cracked plates by affixing a bonded patch onto the affected region. In this context, numerous research efforts have been undertaken since the 1980s to explore the application of patches for the restoration of damaged structures. The external bonding patch repair technique entails adhesively bonding a damaged plate to composite or aluminium patches.

Fig. 13 depicts the experimentally measured load variation in relation to displacement for a plate repaired with a single bonded patch, as well as an unrepaired plate. The bonded patch applied to the notched plate diminishes stresses at the notch's edge, thereby enhancing the strength of the plate. This disparity arises from the transfer of load from the damaged zone to the repair patch through the adhesive layer. The maximum difference in tensile strength between specimens repaired with a bonded patch and unrepaired plates does not exceed 15%. Fig. 14 presents a comparison between the different repair techniques such as bonding, riveting and hybrid joining (riveting + bonding). The obtained curves are irrespectively of the repair technique used (whether bonded or riveted). The behaviour is mostly linear up to failure, although a minor slope reduction is visible for the repaired specimens. The findings unequivocally demonstrate that the bonded patch yields a more favourable effect compared to the riveted patch given the size of the repaired surface. On the other hand,

repair by the combination of the two processes (bonded + riveted) in the repair generates a higher strength of the notched plate by approximately 20%.
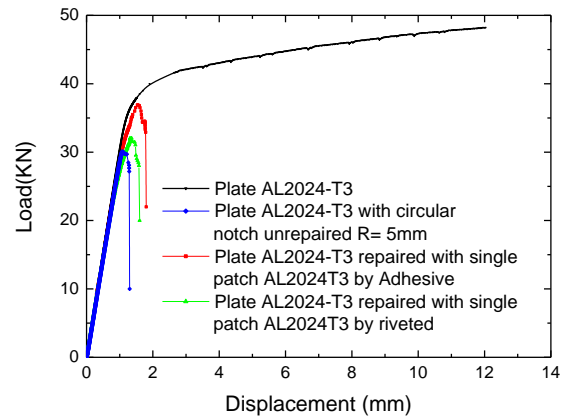
Fig. 13. Load–displacement plot for of the repaired plate using a bonded aluminium patch and unrepaired plate
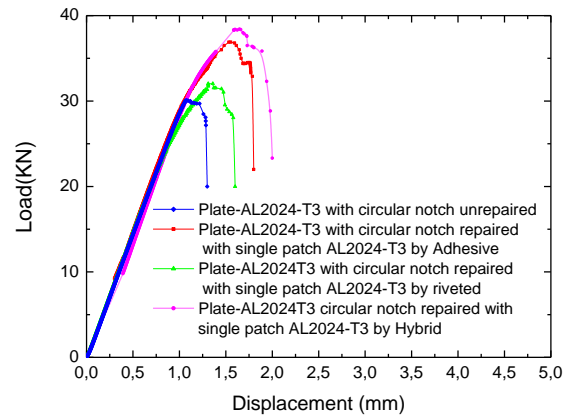
Fig. 14. Comparison of load–displacement curves for the different repair techniques by aluminium patch (riveting, bonding and hybrid joining)

The comparison between load–displacement curves shows that the maximum load for the undamaged plate drops considerably if the plate contains a notch (Fig. 15).

Fig. 15. Maximal load for the different plates (0 denotes repaired by hybrid process, 1 repaired by the bonded aluminium patch and 2 repaired by the riveted patch)

Patch repairing using the different techniques results in an increase in the strength of the damaged plate, which depends on the type of repair. Repair using the hybrid process produces a considerable increase in maximum tensile strength.

During the tensile test of the notched plates, cracks are initiated at the level of the central notch, which then propagate towards the free edge of the plate, as shown in Fig. 16. A deviation of the propagation path of the crack is observed in the case of a hybrid repair (specimen 6 in Fig. 16). The rivet holes create additional stress concentrations. The hybrid patch absorbs better some of the stress concentrations that are localised at the central notch compared to other repair techniques.
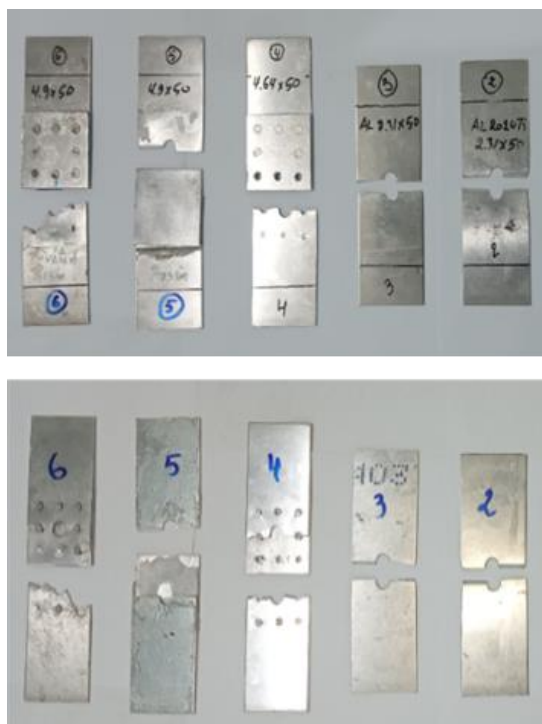
Fig. 16. Unrepaired and repaired aluminium plates with different processes after failure

The mechanical strength and life of a restored structure are markedly contingent on several factors, with particular emphasis on the mechanical and geometric attributes of the structure, the adhesive properties and the quality of the repair patch employed. It is crucial to underscore that the patch stands out as a primary component, directly influencing the performance of the repaired structure. Numerous studies have been conducted to enhance the effectiveness and longevity of composite patch repairs, ultimately striving to extend the service life of the restored structure. In this context, two types of composite patch are used, namely carbon/epoxy and glass/epoxy, aiming to assess the patch influence on the repair ability of a plate with notch and subsequent repair by composite patches. In this case, the patches are adhesively bonded to the plate, riveted or applied by a hybrid process (riveting + bonding). The results of the tensile tests are shown in Figs 17 and 18.

The obtained results show that the glass/epoxy composite patch does not have a major effect on the repair performance. A reliable improvement in the tensile strength is observed in Fig. 17.

The behaviour of the structure repaired by the glass/epoxy composite patch is practically the same as for the structure re-

paired by a metal patch. It is also noted that the improvement in the structural strength is low for the two repair methods. On the other hand, it is observed in Fig. 18 that the strength of the damaged plate repaired by the composite carbon/epoxy patch has slightly improved compared to the previous cases (repair by glass/epoxy patch) since the mechanical properties of the carbon/epoxy patch and the glass/epoxy patch are quite distinct.
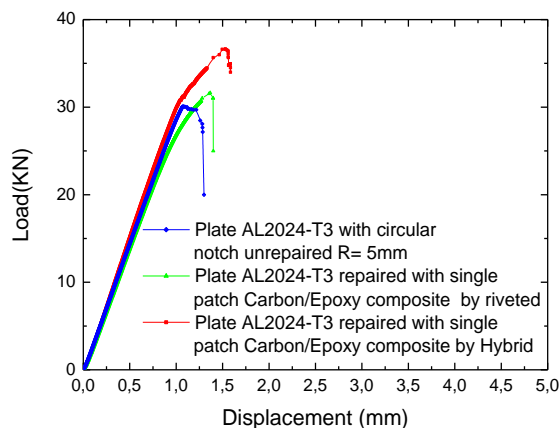
Fig. 17. Load–displacement plot for the plate repaired by the single glass/epoxy composite patch
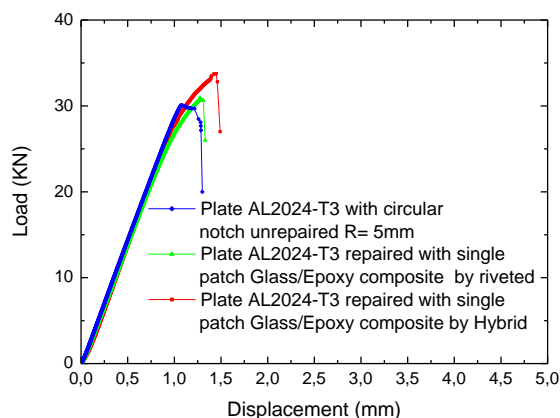
Fig. 18. Load–displacement plot for the plate repaired by the single carbon/epoxy composite patch, considering different repair processes
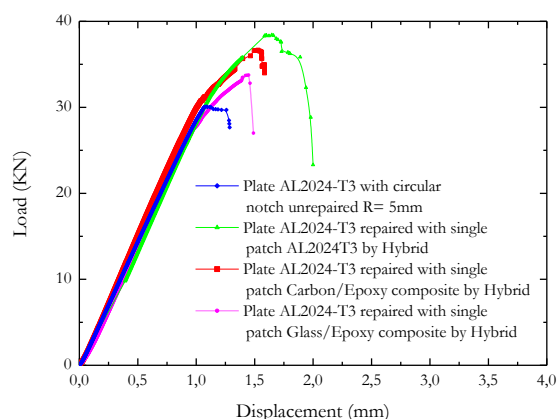
Fig. 19. Comparison of the load–displacement curves of a notched plate with a drilled plate repaired with different types of patches using hybrid bonding/riveting

Improving the strength of the structure repaired by the hybrid repair process (bonding and riveting) has shown effectiveness in improving the strength of the structure, even if the nature of the repair patch is varied. In plates repaired by the hybrid patch, and for all types of patch, the rivets and the adhesive layer together increase the structural strength.

Since the hybrid process (bonding + riveting) offers the best performance for increasing the strength of the damaged plate, a comparison of this process using different types of repair patch is shown in Fig. 19.

The hybrid repair process (bonding and riveting) was effective in improving the strength of the repaired structure, even when the nature of the repair patch was varied (Fig. 19).

Fig. 20. Damage to notched plates repaired by different types of patch using hybrid repair processes (bonding/riveting)
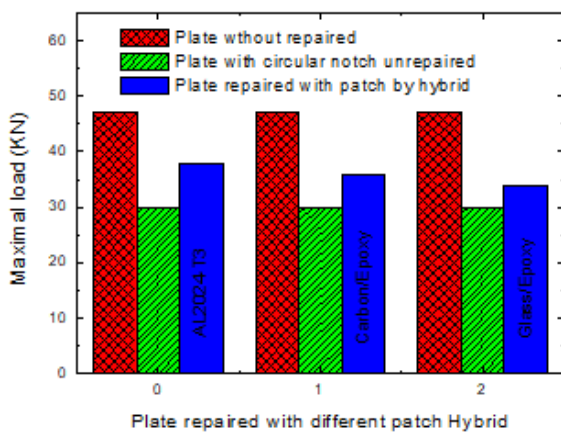
Fig. 21. Maximal load for different plates (0 denotes repaired by the hybrid process, 1 repaired by the bonded aluminium patch and 2 repaired by the riveted patch)

In the hybrid repair, regardless of the nature of the patch, it was found that the rivet and adhesive work together to transfer loads. This is a very important aspect to consider in the present study to achieve the objective of improving the strength of the damaged structure. It is preferable to reduce the number of rivets or eliminate those in the vicinity of the notch across the half-width of the plate. To achieve this, it is important to ensure that the rivet

and adhesive work together and to look for another more effective type of adhesive.

Damage to plates repaired by the hybrid process (bonding and riveting) has the same appearance (Fig. 20): the crack initially starts at the notch and propagates along the half-width of the plate. Once in the vicinity of the rivet notch, the crack propagates slowly until the plate breaks completely.

The maximum force of the damaged plate in the presence of the hybrid process improves considerably and has a slightly higher value than that of the damaged plate without repair. Even in the presence of a repair patch, the maximum force always remains lower than that of the undamaged plate, whatever the nature of the patch (Fig. 21).

7.2. Effect of presence of a crack emanating from a notch

The presence of a crack of length a = 5 mm emanating from a notch (Fig. 22) considerably reduces the tensile strength of the plate by up to 50% compared to the continuous plate without notch. It was observed that the presence of the crack emanating from notch in the plate has a considerable effect on the tensile strength. On the other hand, its effect on the stiffness of the plate is almost negligible.
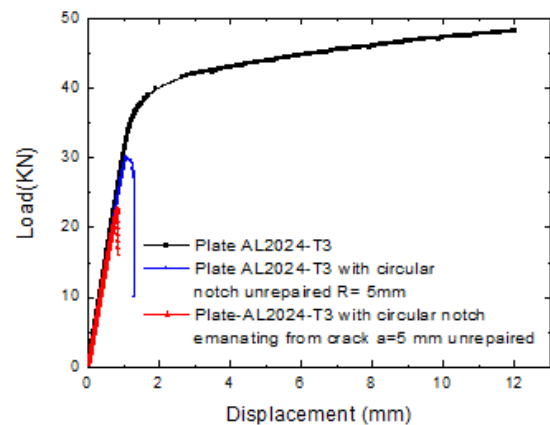
Fig. 22. Effect of the presence of a crack on the variation of the load–displacement diagram

7.3. Patch effect in plate repair in the presence of a notch crack

It is widely acknowledged that structures in service are often subjected to mechanical stresses, which could cause damage. In this case, one alternative to repairing these structures is to prevent them before a crack appears. Currently, new techniques are being developed to reduce the crack propagation and consequently increase the service time of structures. The most used technique consists of repairing the structure by placing a bonded or riveted patch part in the area damaged defined by the notch. For this purpose, the plate in the presence of a crack emanating from the notch was repaired by the three methods (Fig. 23).

The presence of a metal repair patch improves the mechanical strength of the damaged plate by increasing its maximum tensile force. The value of this force depends on the type of repair. The riveted patch only slightly improves the mechanical resistance

of the damaged plate given the presence of additional holes. However, the patch bonded by the hybrid process further improves the resistance of the plate.

The presence of a crack emanating from a notch creates a strong concentration of stress; in this case, patch repair according to the three methods only slightly improves the resistance of the plate. The damage plate in the presence of a notch crack is shown in Fig. 24, where it can be seen that plate failure rapidly occurs in the presence of the crack and that its propagation is rapid towards the free edge of the plate, more precisely towards the rivet hole located at the mid-width of the plate.
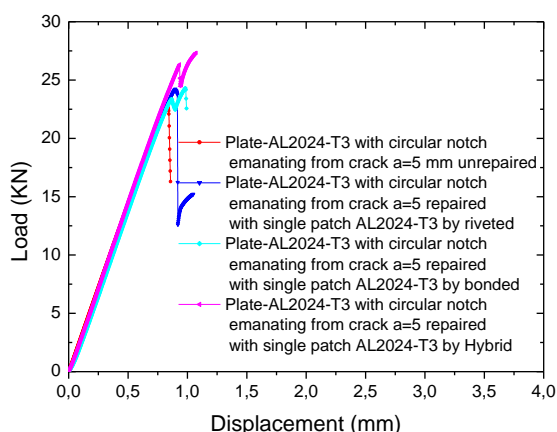
Fig. 23. Load–displacement plot for the plate with a crack emanating from notch repaired by aluminium patch with different repair processes

Fig. 24. Damage to the notched plates repaired by an aluminium patch using hybrid repair processes (bonding/riveting)

The maximum strength of the plate with repair varies according to the type of repair (riveting, bonding and hybrid) (Fig. 25). The presence of a riveted patch in a plate in the presence of a crack emanating from a notch does not improve the mechanical strength of the plate since the rivet hole is in the vicinity of the crack and the width of the plate will be reduced. However, the presence of a bonded patch slightly improves the value of the plate's strength, while the presence of a hybrid repair further improves the strength of the damaged plate and presents the highest load.

An attempt was made to determine the effect of the patch type in increasing the strength of the damaged plate in the presence of a notch crack using the hybrid repair process (bonding + riveting) since this is the most effective process compared with the other two tested ones (bonding and riveting) (Fig. 26).

It is clear that the aluminium patch performs better than the other two patch types (glass/epoxy and carbon/epoxy). This is because the presence of rivet holes creates defects in the composite patches (possible delamination) and stress absorption is not effective.

Plate damage in the presence of a notch crack is shown in Fig. 27, where it can be seen that plate failure occurs rapidly in the presence of the crack and that its propagation is rapid towards the free edge of the plate, more precisely towards the rivet hole located at the mid-width of the plate, regardless of the nature of the patch.
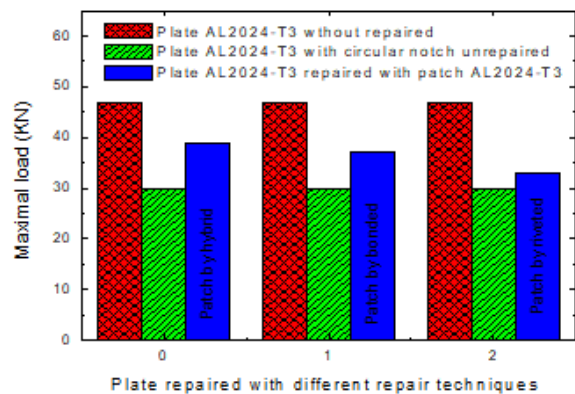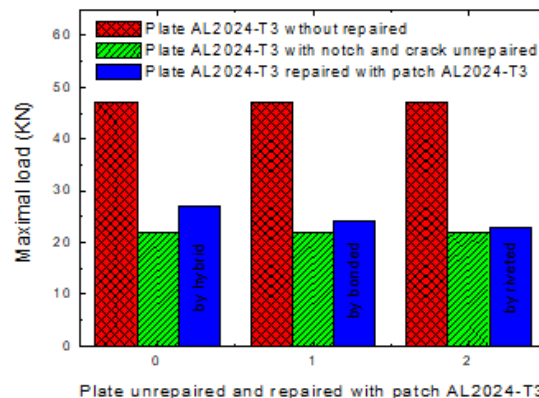

Fig. 25. Maximal load for the different plates (0 denotes repaired by the hybrid process, 1 repaired by the bonded aluminium patch and 2 repaired by the riveted patch) in the presence of a crack emanating from a notch
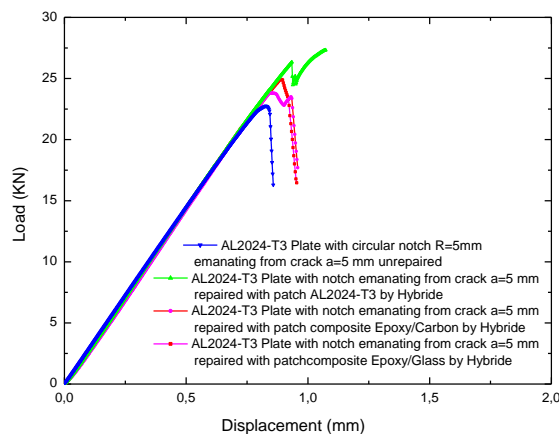
Fig. 26. Comparison of the load–displacement curves of the plate in the presence of crack emanating from a notch and repaired with different types of patches using hybrid bonding/riveting

Fig. 27. Damage to the notched and cracked plates repaired by different types of patches using hybrid repair processes (bonding/riveting)

The strength of the plate repaired by the hybrid process (riveting + bonding) varies according to the type of the repair patch (aluminium, glass/epoxide and carbon/epoxide) (Fig. 28). The presence of a glass/epoxide repair patch in a plate with a crack

emanating from a notch only slightly improves the mechanical strength of the plate. However, the presence of an aluminium patch improves the strength further.
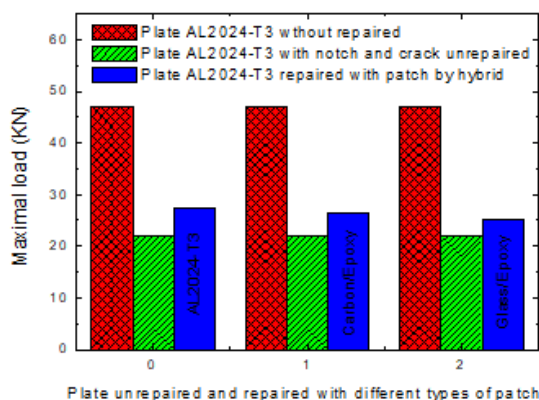
Fig. 28. Maximal load for different plates (0 denotes repaired by an aluminium patch, 1 repaired by bonded carbon/epoxide and 2 repaired by glass/epoxide) in the presence of a crack emanating from a notch repaired by hybrid repair

8. CONCLUSION

This work focused on the influence of different types of repair (bonding, riveting and hybrid) using patches of different natures (aluminium, glass/epoxide and carbon/epoxide) on a damaged aluminium plate subjected to tensile loading. In conclusion, improving the strength of the plate must take into account both the nature of the patch and the adhesive. In the riveting or hybrid repair process, the arrangement and number of rivets must also be analysed in detail.

It was found that the composite patch only slightly absorbs the stress concentration of the damaged area once it is pierced and that the presence of rivet holes in the composite leads to further damage because of the possible creation of delamination of the different layers.

− The aluminium patch works best in the presence of a hybrid repair.
− The presence of a crack at the notch considerably reduces the strength of the plate, and in the presence of a repair, the strength of the plate improves only slightly as the crack propagates rapidly towards the rivet hole.

REFERENCES

1. Phyo AM, Hirohito K, Mikihito H. Fatigue-performance improvement of patch-plate welding via PWHT with induction heating. Journal of Constructional Steel Research. 2019;160(3):280–288.
https://doi:10.1016/j.jcsr.2019.05.047
2. D'Aniello M, Portioli F, Fiorino L, Landolfo R. Experimental investigation on shear behaviour of riveted connections in steel structures. Eng. Struct. 2011; 33(2):516–531.
https://doi:10.1016/j.engstruct.2010.11.010
3. Ishikawa T, Ikeda T. Patch Plate Repair Method for Steel Structures Combining Adhesives and Stud Bolts. International Journal of Steel Structures. 2018;18:1410–1419.
https://doi.org/10.1007/s13296-018-0149-0
4. Chen ZY, Gu XL, Zhao XL, Ghafoori E. Fatigue Tests on Fe-SMA Strengthened Steel Plates Considering Thermal Effects.

5. Ghafoori E, Dahaghin H, Diao Ch, Pichler N, Li L, Ding J, Ganguly S, Williams S. Metal 3D-Printing for Repair of Steel Structures. Proceedings in civil engineering. 2022;796-801.
https://doi.org/10.1002/cepa.2285.
6. Tolga D, Costas S. Recent developments in advanced aircraft aluminium alloys. Materials & Design. 2014;56(1):862–871.
https://doi:10.1016/j.matdes.2013.12.002.
7. Nayak NV. Composite materials in aerospace design. Mater. Des. 2014; 4(9): 1–10.
8. Baker A. Bonded composite repair of fatigue-cracked primary aircraft structure. Compos. Struct. 1999;47(1):431-443.
https://doi: 10.1016/S0263-8223(00)00011-8
9. Baker A, Rose A, L. R. F and Jones R. Advances in the Bonded Composite Repair of Metallic Aircraft Structur. 1ère ed. Netherlands. Elsevier Science. 2002. ISBN: 0-08-042699-9.
10. Hosseini-Toudeshky H, Sadeghi G, Daghyani HR. Experimental fatigue crack growth and crack-front shape analysis of asymmetric repaired aluminium panels with glass/epoxy composite patches. 2005; 71(3-4): 401–406. http://doi:10.1016/j.compstruct.2005.09.032.
11. Khalili SMR, Ghadjar R, Sadeghinia M, Mittal RK. An experimental study on the Charpy impact response of cracked aluminum plates repaired with GFRP or CFRP composite patches. Composite Structures.2008; 489(2): 270-274.
http://doi:10.1016/j.compstruct.2008.07.032
12. Maleki HN, Chakherlou TN. Investigation of the Effect of Bonded Composite Patch on the Mixed-Mode Fracture Strength and Stress Intensity Factors for an Edge Crack in Aluminum Alloy 2024-T3 Plates. Journal of Reinforced Plastics and Composites.2017; 36(15): 1074-1091. http://doi:10.1177/0731684417702001
13. Basaid D, Benmounah A, Aribi Ch, May A. Experimental study of repair of aircraft structures by adhesive patches based on epoxy and fiberglass. Journal of Materials and Engineering Structures. 2019; 6(3):409–426.
14. Gu J-U, Yoon H-S, Choi N-S. Caractérisation de l'émission acoustique d'une plaque d'aluminium crantée réparée avec un patch en fibre composite. Composites Part A: Applied Science and Manufacturing. 2012;43(12):2211–2220.
http://doi:10.1016/j.compositesa.2012.07.018
15. Benkheira A, Belhouari M, Benbarek S. Comparison of Double- and Single-Bonded Repairs to Symmetrical Composite Structures. Journal of Failure Analysis and Prevention.2018.
http://doi:10.1007/s11668-018-0557-7
16. Kaddouri N, Madani K, Rezgani L, Mokhtari M, Feaugas X. Analysis of the effect of modifying the thickness of a damaged and repaired plate by composite patch on the J-Integral; effect of bonding defects. Journal of the Brazilian Society of Mechanical Sciences and Engineering. 2020;42(8). http://doi:10.1007/s40430-020-02515-y
17. Madani K, Touzain S, Feaugas X, Cohendouz S, Ratwani M. Experimental and numerical study of repair techniques for panels with geometrical discontinuities. Computational Materials Science. 2010; 48(1):83–93. http://doi:10.1016/j.commatsci.2009.12.005
18. Aldeen A, Mahdi D, Zhongwei C, Disher I, Mohamad B. Effect of isothermal and isochronal aging on the microstructure and precipitate evolution in beta-quenched n36 Zirconium alloy. Facta Universitatis-Series Mechanical Engineering.2023.
https://doi:10.22190/FUME230 405019A
19. Rivallant S, Bouvet C, Hongkarnjanakul N . Failure analysis of CFRP laminates subjected to compression after impact simulation using discrete interface elements.Compos. Part A: Appl. Sci. Manuf. 2013.55:83-93.
https://doi.org/10.1016/j.compositesa.20 13.08.003
20. Rashnooie R, Zeinoddini M, Ahmadpour F, Beheshti Aval SB, Chen T. A coupled XFEM fatigue modelling of crack growth, delamination and bridging in FRP strengthened metallic plates. Engineering Fracture Mechanics. 2023.279(17):200-230.
https://doi.org/10.1016/j.engfracmech.2022.109017

Publication: Journal of Structural Engineering. 2022;149(3).
https://doi.org/10.1061/JSENDH.STENG-11694

21. Ait Kaci, K Madani, M Mokhtari, X Feaugas, S Touzain. Impact of composite patch on the J-Integral in adhesive layer for repaired Aluminum plate. Advances in Aircraft and Spacecraft Science. 2017; 4(6): 679-699. https://doi.org/10.12989/aas.2017.4.6.679.

22. Bernhard Horn, Johannes Neumayer and Klaus Drechsler. Influence of patch length and thickness on strength and stiffness of patched laminates. Journal of Composite Materials. 2018;52(16):2199–2212. https://doi.org/10.1177/0021998317740413

23. K Madani, S Touzain, X Feaugas, M Benguediab, M Ratwani. Stress distribution in a 2024-T3 aluminum plate with a circular notch, repaired by a graphite/epoxy composite patch. International Journal of Adhesion and Adhesives. 2009; 29: 225-233.
https://doi:org/10 .1016/ j.ijadhadh.2008.05.004

24. Rezgani L, Madani K, Feaugas X, Touzain S, Cohendoz S, Valette J. Influence of water ingress onto the crack propagation rate in a AA2024-T3 plate repaired by a carbon/epoxy patch. Aerospace Science and Technology.2016;55:359–365.
https://doi:10.1016/j.ast.2016.06.010

25. Wahrhaftig AM, Plevris V, Mohamad B A, Pereira D L .Minimum design bending moment for systems of equivalent stiffness. Structures.2022;57:105224.
https://doi.org/10.1016/j.istruc.2023.105224

26. Al-Abboodi H, Fan H, Al-Bahrani M, Abdelhussien A, Mohamad B. Mechanical characteristics of nano-crystalline material in metallic glass formers. Facta Universitatis-Series Mechanical Engineering. 2023. https://doi:10.22190/FUME230128016A

27. Davis M, Bond D. Principles and practices of adhesive bonded structural joints and repairs. International Journal of Adhesion and Adhesives.1999;19:91–105.
https://doi.org/10.1016/S0143-7496(98)00026-8

28. Xi J, Yu Z. Toughening mechanism of rubber reinforced epoxy composites by thermal and microwave curing. J. Appl. Polym. Sci. 2017;135(5): 45767–45775. https://doi.org/10.1002/app.45767

29. Maleki A, Saeedifar M, Najafabadi MA, Zarouchas D. The Fatigue Failure Study of Repaired Aluminum Plates by Composite Patches using Acoustic Emission. Engineering Fracture Mechanics.2017; 210(1):300-311.https://doi.org/10.1016/j.engfracmech.2017.12.034

30. Seidl AL. Repair Aspects of Composite and Adhesively Bonded Aircraft Structures. Handbook of Composites. Chapter 39. Springer. 1998;857-882.

31. Zitoune R, Collombet F. Numerical Prediction of the Thrust Force Responsible of Delamination During the Drilling of the Long-fibre Composite Structures. Composites Part A: Applied Science and Manufacturing.2007;38(3):858–866.
https://doi.org/10.1016/j.com positesa.2006.07.009

Abdelkrim Merah: https://orcid.org/0000-0003-1376-5400

Amin Houari: https://orcid.org/0009-0004-2617-2182

Kouider Madani: https://orcid.org/0000-0003-3277-1187

Mohamed Belhouari: https://orcid.org/0000-0001-7863-1222

Salah Amroune: https://orcid.org/0000-0002-9565-1935

Ahmed Chellil: https://orcid.org/0000-0001-9467-4214

Cherif Zineelabidine Yahia: https://orcid.org/0009-0006-3155-5238

Raul D.S.G.Campilho: https://orcid.org/0000-0003-4167-4434

A REVIEW OF THE RELAXATION MODELS FOR PHASE TRANSITION FLOWS CENTERED ON THE TOPOLOGICAL ASPECTS OF THE NONEQUILIBRIUM MASS TRANSFER MODELLING

Wojciech ANGIELCZYK*

*Faculty of Mechanical Engineering, Białystok University of Technology, ul. Wiejska 45C, 15-351 Białystok, Poland

w.angielczyk@pb.edu.pl

Abstract: The first part of this work is a brief (application-oriented) review of the different classes of multiphase flow models. The review starts with the most generic approaches and descends to the class of Homogeneous Relaxation Models (HRM) of two-phase flow. Subsequently, this work presents a detailed review of the developed relaxation equations describing nonequilibrium mass transfer in two-phase flows. Some of the reviewed equations (in particular, the closure equations of HRMs) have quite simple mathematical structures but there are indications that they should be, in a specific way, more complex. Consequently, the main aim of this article is to bring attention to this problem and expose its nature and practical importance. The analyses conducted in this study reveal that relaxation closure equations formulated as advection equations may disrupt the phase space structure of the model, whereas equations formulated as phasic mass conservation do not pose such an issue. This distinction arises from the presence of a greater number of gradients in the conservation equations (a minimum of two, compared to potentially just one in an advection equation), rendering the conservation equations mathematically more complex.

Key words: nonequilibrium mass transfer, relaxation equation, Homogeneous Relaxation Model, Delayed Equilibrium Model

1. INTRODUCTION

Over the past few decades, significant strides have been made in advancing the mathematical and numerical modelling of phase transition flows. These advancements have substantially enhanced our ability to describe intricate phenomena, particularly nonequilibrium processes occurring within and between flowing phases. A prevalent method for characterizing nonequilibrium involves treating it as a superposition of distinct disequilibrium processes [1]. These processes are incorporated into the model as relaxation terms within the mass, momentum, and energy conservation laws or via dedicated closure relaxation equations.

Each type of a multi-phase flow is a flow of separate fluids. There always are interfaces (or an interface) that distinguish one phase (fluid) from another [2]. Consequently, every multi-phase flow can be modelled using the Navier-Stokes equations providing that the boundary conditions for each of the phases and at the moving interfaces are specified (direct simulation) [3]. The computational cost of such an approach is enormous, and accounting for the nonequilibrium processes is problematic. Besides, practically important is the number and the surface area of these interfaces in the considered element of the flowing fluid. Taking the ratio of the mentioned quantities into consideration allows us to make appropriate simplifications concerning boundary conditions, fluid and flow variable averaging procedures and finally, the type and number of governing equations. Implementation of those simplifications in a mathematical form gives a flow model specialized for a given two-phase flow structure or range of structures. In this way a variety of specialized models were developed. One of the

most generic of them is the two-fluid model in which a system of conservation equations for mass, momentum, and energy is solved for each phase/fluid [3]. As a consequence of the space and/or time averaging of the phases' properties, the detailed structure of the interphase is lost. However, this six-equation model can account for the difference in pressures, temperatures, and velocities of the phases [4]. Moreover, source terms or closure equations describing interphase mass, momentum, and energy transfer can account for the nonequilibrium effects [4].

The other broad application range model is a seven-equation two-phase flow model of Baer and Nunziato (BN) [5] that can account for velocity, pressure, and temperature disequilibrium between the phases. Originally, the BN model was developed to describe a deflagration-to-detonation transition in granular explosive materials (thus, it is a two-phase solid-gas flow model). However, after certain modifications, it turned out to be capable of describing multiphase flows with ongoing phase transitions accompanied by various non-equilibrium effects [6].

The following multiphase and multidimensional version of the BN model is possibly the most generic:

$$\frac{\partial(\alpha_k \rho_k)}{\partial t} + \nabla(\alpha_k \rho_k u_k) = C_k, \tag{1}$$

$$\frac{\partial(\alpha_k \rho_k u_k)}{\partial t} + \nabla(\alpha_k \rho_k u_k \otimes u_k + \alpha_k p_k \underline{I} - \alpha_k \underline{\tau}_k) = p_{k*} \nabla \alpha_k + M_k - \underline{\tau}_I \nabla \alpha_k + \alpha_k \rho_k g_k, \tag{2}$$

$$\frac{\partial(\alpha_k \rho_k E_k)}{\partial t} + \nabla[\alpha_k(\rho_k E_k + p_k)u_k - u_k \alpha_k \underline{\tau}_k] = u_I p_{k*} \nabla \alpha_k - p_I F_k + \epsilon_k - u_I(\underline{\tau}_I \nabla \alpha_k) + q_k + J_k + \alpha_k \rho_k(u_k g_k), \tag{3}$$

$$\frac{\partial \rho_k}{\partial t} + \nabla(\rho_k u_k) = -\frac{\rho_k}{\alpha_k} F_k, \tag{4}$$

where: α - the volume fraction, ϵ - the interphase energy transfer rate per infinitesimal volume (piv) [W/m3], ρ - the density [kg/m3], C – the interphase mass transfer rate piv [kg/(s m3)], E - the specific total energy [J/kg] (the potential energy is commonly neglected), F - the compaction rate (mimics the effects of micro-structural forces) [1/s] [7], g - the acceleration vector caused by the net external body force [m/s2], J - the external heat transfer rate piv [W/m3], M – vector of the interphase momentum transfer rate piv [N/m3], p - the pressure [Pa], q – the conduction heat transfer rate piv [W/m3], t – time [s], $u = [u_x, u_y, u_z]^T$ - the velocity vector [m/s], τ - the viscous stress tensor (frequently calculated by the Newtonian approximation [4]) [Pa], $\underline{I}$- the identity tensor, k - subscript indicating the phase ($k = 1, ..., n$), k^* - subscript indicating the conjugate phase of the phase k [6], n - number of phases, I – subscript denoting interface quantity.

The blue terms in Eq. (1-4) are viscosity, heat conduction, and external energy source terms, respectively, recently added to the BN model by Zhang et al. [6]. However, the interphase mass transfer C_k was omitted in [6], consequently, the model can accurately describe flows without interphase mass exchange.

In the presented model formulation (Eq. 1-4) this limitation was removed (by addition of the first green term). Also, the terms for including external body forces effects (two last green terms) were added.

The presented partial differential equations express the conservation of mass (Eq. 1), momentum (Eq. 2), and energy (Eq. 3) and they must be formulated for each of n phases, while the compaction dynamics equation (Eq. 4) is required only for $(n - 1)$ phases [7].

Taking $n = 2$ and replacing the green and blue terms with zeros yields the multidimensional version of the original (seven-equations) BN model [7] wherein for k=1, k*=2 (or for k=2, k*=1), moreover:

$$C_2 = -C_1, \quad \epsilon_2 = -\epsilon_1, \quad M_2 = -M_1, \quad F_2 = -F_1. \tag{5}$$

Let us further assume that the mixture is a vapor-liquid system and admit k=1 for the vapor and k=2 for the liquid, and introduce $C = C_1 = |C_2|$, and $F = F_1 = |F_2|$, then the previously described compaction rate and the source terms can be expressed as follows [8]:

$$F = r_p(p_2 - p_1), \tag{6}$$

$$M = Cu_1 + r_u(u_1 - u_2), \tag{7}$$

$$\epsilon = C\left(e_2 + \frac{1}{2}u_2^2\right) + r_u u_2(u_1 - u_2) \\ + r_T(T_1 - T_2), \tag{8}$$

where: T denotes absolute temperature [K] and e stands for specific internal energy [J/kg]. More on the physical meaning of those source terms can be found in [8]. For the purpose of this work, it is only important to understand that those source terms characterize phase interactions and that some of their components (relaxation terms) decide how fast the difference in a given phasic variable decreases (relaxes) with time. In this meaning, one can say that the drag source $r_u(u_1 - u_2)$ present in the momentum (Eq. 2) and energy (Eq. 3) equations, equilibrates velocities. The heat transfer term $r_T(T_1 - T_2)$ in the energy equations (Eq. 3) equilibrates temperatures. Finally, the whole compaction dynamics equation (Eq. 4, with the relaxation term F)

is a relaxation law that drives the phases toward pressure equilibrium.

The relaxation rates r_p, r_u, r_T are all positive and the higher they are the faster a given variable equilibrates.

In contrast to the two-fluid model (dedicated mainly to fully separated flows), BN-type models assume that the flowing mixture is homogeneous enough to be treated as a continuum with phenomenologically appropriate parameters such as density, velocity, temperature, or sound speed.

1.1. Hierarchy of the BN-type relaxation models

A narrower-scope model can be derived from BN model (Eq. 1-4) by introducing infinite-rate relaxation for a selected property. In such a way a hierarchy of relaxation models can be established [9].

Applying an infinite relaxation rate of velocity reduces the model to the six-equation single-velocity model [9] that can account for pressure, temperature, and chemical potential disequilibrium between the phases. Thus, the resulting model is simpler (easier to be solved [9]) but unable to fully describe a possible mechanical disequilibrium in the flow (both phases still could have different pressures but must have the same velocity). This kind of model can accurately describe cavitating, flashing and condensing flows providing that one of the phases is quite uniformly distributed in the other phases (homogeneous flow). Moreover, it can be used for modelling of the interfacial flows as the same type equations apply to the direct numerical simulation of boiling flows at sub-bubble scale.

Subsequently, assuming instantaneous pressure relaxation the five-equation mechanical equilibrium model (single-velocity and single-pressure) is obtained [9]. The formulation of a such model was done by Kapila et al. (2001) [8] and then it was adjusted for cavitating flows by Saurel et al. (2008) [10] and Le Martelot et al. (2013) [11]. This class of models accurately describes flashing and condensing homogeneous flows. It can be used for predicting the cavitating flows in cases when the pressure difference between phases does not play a significant role [11].

In the limit of instantaneous mechanical and thermal relaxation we obtain a four-equation, single-velocity, single-pressure, and single-temperature two-phase flow model.

The first of such models relayed on the relaxation of the chemical potential [12]. The later model proposed simultaneous relaxation of three different fluid properties (including the Gibbs free energy) [13]. However, in the subsequent models [14], the interphase mass transfer is driven by a difference in only the Gibbs free energy of the phases. In [15], a fast solver for such models is proposed.

Considering the equality of pressures and temperatures of the phases, a non-zero difference in Gibbs free energy is possible due to the application of the van der Waals Equation of State (EOS) for each phase [16]. However, a physically consistent EOS leads to the same Gibbs free energy for vapor and liquid at the same temperature and pressure. In turn, an assumption that both phases are in metastable states introduces inequality of at least one of the mentioned intensive properties. For those reasons, this class of models can be treated as physically inconsistent.

Finally, in the limit of full instantaneous thermodynamic equilibrium, the three-equation Homogeneous Equilibrium Model (HEM) [17] is obtained. The HEM can accurately describe homogeneous two-phase flows without strong pressure changes and

rater with subsonic velocities as it does not account for any of the nonequilibrium effects, and it underestimates the sound speed for flows characterized by low vapor mass fraction.

There is a four-equation (single-velocity, single-pressure) two-phase flow model that is often treated as a BN-type model however it cannot be obtained as the above-mentioned models. It is a Homogeneous Relaxation Model (HRM) [18] in which the interphase mass transfer is driven by a thermal disequilibrium between the metastable liquid and the saturated vapor. An evolution of vapor mass fraction in the HRM is described by a rate equation [19]. Introducing of a third phase into the HRM leads to the Delayed Equilibrium Model (DEM). This additional phase is a saturated liquid that has the same pressure p as the remaining phases and the same temperature as the saturated vapour $T_{sat}(p)$. The temperature of the metastable liquid T is higher than the temperature of the saturated phases.

1.2. General remarks and historical background on modelling of the relaxation phenomenon in fluids

Relaxation is a process of return of a disturbed system to equilibrium. Relaxation time θ characterizes the speed of the return. If the state of the system is characterized by the property ψ, then its relaxation changes are described as follows [20]:

$$\frac{D\psi}{Dt} = \frac{\partial \psi}{\partial t} + u\nabla\psi = -\frac{\psi - \psi_{eq}}{\theta}, \tag{9}$$

where ψ_{eq} is the value of the property ψ at equilibrium. The physical sense of the relaxation time θ is such that after it has elapsed the deviation from the equilibrium state $(\psi - \psi_{eq})$ decreases e times (e - the base of the natural logarithm).

The idea of relaxation time has been successfully used already by Einstein (1920) [21], to analyse sound propagation in partially dissociated gases, and Mandelshtam and Leontovich (1937) [22], in the analysis of sound absorption in liquids. Nevertheless, probably the first application in the field of two-phase flows is attributed to Bauer (1976) [23], who replaced ψ in Eq. (9) with the vapour mass fraction x to describe nonequilibrium mass transfer between phases of a gas-liquid mixture. Then, the concept presented by Eq. (9) was adopted in the seven-equation two-phase flow model of Baer and Nunziato (1986) [5]. However, it was not until 1990 that Bilicki et al. [18] analysed this equation in the background of the linear theory of irreversible processes (stating that it is fully consistent with the mentioned theory). They also showed that the theoretical expression for the relaxation time depends on the second-order derivative of specific Helmholtz free energy a. This suggests that in the concerned modification of Eq. (9), implicitly, apart from ∇x, a second gradient ∇a is involved. Also, the same article shows how to attach the concerned relaxation equation to the system of the conservation equations (of mass, momentum and energy) to formulate the HRM. Subsequently, the authors described the difference between the HEM and HRM, focusing on a study of dispersion, characteristics, choking, and shock waves. However, no closure equation for the relaxation time θ was proposed; instead, several constant values for this parameter were assumed in the calculations.

2. REVIEW OF NONEQUILIBRIUM MASS TRANSFER RELAXATION MODELLING FOR TWO-PHASE FLOWS

This review is supposed to present selected available literature approaches that, according to the author's knowledge, cover all developed consistent modelling techniques.

To make this review more useful, it was decided to provide the numbers of equations and tables that describe the presented equations in the Original Article (OA). Those numbers are given in angle brackets.

The equations, that can be found in the OA, generally differ visually from the presented versions since they have been transformed to match the form given by Eq. (11) or Eq. (12), and the nomenclature of this work.

2.1. Simple thermodynamic relations

Only equations with one relaxation source term are classified into this category. These relations are most often encountered in BN-type models in the form of the phasic mass conservation equation (Eq. 1). Otherwise, they have the form of an advection equation:

$$\frac{\partial \beth_k}{\partial t} + u_k \nabla \beth_k = K_k, \tag{10}$$

where $\beth_k$, in the case of the two-phase flow, is the volume fraction α, the vapour mass fraction x or the mass fraction of the saturated phases y (the saturation index). In all these cases:

$$K_k = r(\psi_k - \psi_{k*}), \tag{11}$$

or

$$K_k = r(\psi_k - \psi_{eq}), \tag{12}$$

where ψ_k denotes a thermodynamic property of the k phase, ψ_{eq} denotes a value of that property at equilibrium (for $\psi_k = y$, $\psi_{eq} = y_{eq} = 1$), and r is a relaxation rate. The relaxation rate is inversely proportional to the relaxation time θ of the considered thermodynamic property ψ.

The above description is a generalisation (of the various approaches found in the literature – Tab.1 and Tab. 2) made to emphasise the fundamental concept: the difference of a certain property between phases (or its value at the equilibrium state) is the driving force of the mass exchange process. However, it has the following limitations: It can only be used for properties that, at equilibrium, have the same value in all phases. Eq. (11) has no physical meaning for properties describing mixture composition (such as the volume fraction and the mass fraction).

A drawback of the considered approaches is predicting zero interphase mass transfer when $\psi_k = \psi_{k*} \neq \psi_{eq}$ (or, in other words, when $\psi_k = \psi_{k*}$ and $r \neq \pm\infty$). Thus, the model formulation in which such a situation is possible (e.g., a model with constant r) cannot describe the equilibrium mass transfer.

Tab. 1 summarises the approaches described by Eqs. (10 – 12). Tab. 2 is a synopsis of the approaches described by Eqs. (1,11,12) for which $C = K$. The last column of each of these tables shows the reference index of the article from which the given equation comes (Original Article Index, OAI). This column also exhibits the fluid type for which the equation was developed. If the fluid is not specified, the equation is supposed to hold for

every fluid. Some approaches presented in Tab. 2 are based on the conservation equation of the phasic mass (Eq. 1) in which the left-hand side uses x instead of α. In this case, (x) is added just after the model type description.

Tab. 1. Review of the selected simple thermodynamic relations for nonequilibrium mass transfer modelling with a form given by Eq. (10). Denotations and remarks: T - in the description of the model type, indicates that the model is transient; a lack of T implies a steady-state flow model, S_I - the specific exchange surface; c- the acoustic speed; $\varphi = [p_{sat}(T) - p][p_c - p_{sat}(T)]^{-1}$; $p_{sat}(T)$ – the saturation pressure at a temperature of the metastable liquid T; p_c – the fluid's critical pressure; A, P – the flow channel cross-section area and its perimeter, respectively; $(2y - y^2)_{eq} = 1$; u_{L0} - the liquid superficial velocity; $\varepsilon = 75.28(A_{div} - A_{conv})(A_{ref} - A_{conv})^{-1}$; A_{div}, A_{conv} – the convergence and divergence rates of the converging-diverging nozzle, respectively; A_{ref} – referential nozzle divergence rate

Form of K / Model type	$\beth, \psi$	Relaxation rate r	OAI / Fluid
Eq. (11) < Eq. 3.1a > / 3D T BN	α, p	$\dfrac{S_I}{c_1\rho_1 + c_2\rho_2}$ < Eq. 3.1h >	[10]
Eq. (11) < Eq. 6a / Eq. 1a > / 3D T BN	α, p	Constant / Unspecified < Eq. 2a > / ---	[9/11]
Eq. (11) < Eq. 1a > / 3D T BN	$\alpha, \rho c^2$	$\dfrac{\nabla u}{\rho_1 c_1{}^2 \alpha_1{}^{-1} + \rho_2 c_2{}^2 \alpha_2{}^{-1}}$ < Eq. 1a >	[11]
Eq. (12). < Eq. 3 > / 1D T HRM	x, x	$\left[3.84 \cdot 10^{-7} \cdot \alpha_1{}^{-0.5} \varphi^{-1.8}\right]^{-1}$ < Eq. 11 >	[19] / H_2O
Eq. (12) < Eq. 4 > / 1D T HRM	x, x	$\left[2.14 \cdot 10^{-7} \cdot \alpha_1{}^{-0.5} \varphi^{-1.8}\right]^{-1}$ < Eq. 21 >	[24] / CO_2
Eq. (12) < Eq. 9 > / 3D HRM	x, x	$\left[\theta(p) \cdot \alpha_1{}^{a(p)} \varphi^{b(p)}\right]^{-1}$ < Eq. 12, Tab. 3 >	[25] / CO_2
Eq. (12) < Eq. 25 > / 1D DEM	y, y	$-0.02\dfrac{P}{A}\varphi^{0.25}$ < Eq. 25 >	[26] / H_2O
Eq. (12) < Eq. 8 > / 1D DEM	$y, 2y - y^2$	$-0.01\dfrac{P}{A}\left(\dfrac{|u|}{|u_{L0}|}\right)^{0.1}\varphi^{0.25}$ < Eq. 8 >	[27] / H_2O
Eq. (12) < Eq. 59, Eq. 60 > / 1D T DEM	y, y	$-(0.0084\dfrac{P}{A} + 0.6337)\varphi^{0.228}$ < Eq. 59, Eq. 60 >	[28] / H_2O
Eq. (12) < Eq. 24, Eq. 25 > / 1D DEM	y, y	$-(38 + 1.3 \cdot 10^{-39}e^{\varepsilon})\varphi^{-0.22}$ < Eq. 24, Eq. 25 >	[29] / CO_2
Eq. (12) < Eq. 2 > / 1D DEM	y, y	$-(0.1086\dfrac{P}{A} + 0.5958)\varphi^{0.228}$ < Eq. 2, Tab. 8 >	[30] / $C_2H_2F_4$

Tab. 2. Review of the selected simple thermodynamic relations for nonequilibrium mass transfer modelling with a form given by Eq. (1). Denotations and remarks: μ – the chemical potential; g – the specific Gibbs free energy; h - the specific enthalpy; ML – subscript denoting the metastable liquid; SL - subscript denoting the saturated vapor; SL – subscript denoting the saturated liquid; $(h_{ML}h_{SG})_{eq} = h_{SL}h_{SG}$; $\varkappa$ - isentropic exponent of the vapor

Form of C / Model type	ψ	Relaxation rate r	OAI / Fluid
Eq. (11) < Eq. 3.1.1 / Eq. 1b > / 1D/3D T BN	μ	Constant / Unspecified < Eq. 2c> / ---	[9/12]
Eq. (11) < Eq. 3, Eq. 47a / Eq. 3.1 > / 1D,3D T BN (x)	g	$\rho\, r_g(S_I, p, T)$ < Eq. 3, Eq. 47a / Eq. 3.1 >	[14/15]
Eq. (12) < Eq. 4a, Eq. 14 > / 1D T HRM (x)	$h_{ML}h_{SG}$	$\dfrac{\rho^2(x - x^2)(\varkappa - 1)}{\varkappa\, p\, (h_{SL} - h_{SG})\, \theta}$ < Eq. 4a, Eq. 14 >	[31] / H_2O
Eq. (12) < Eq. 6 > / 1D HRM (x)	x	$\dfrac{\rho}{2.15 \cdot 10^{-7} \cdot \alpha_1{}^{-0.54} \varphi^{-1.76}}$ < Eq. 10 >	[32] / CO_2

2.2. Empirical phenomenological relations

All equations described by the fourth and next rows of Tab. 1 and the equations from the two last rows of Tab. 2 can be classified as empirical phenomenological relations since their relaxation rates have been developed for specific fluids by adjusting the presented equations' constants so that the modelling would result with possibly best approximation of the experimental data.

2.3. More advanced thermodynamic approaches

This category gathers equations based on the idea presented in the previous section that have more than one relaxation source term. Additionally, approaches using more than one equation (from the set of all equations constituting the model) to describe non-equilibrium mass transport are also presented here. A notation that in each relaxation term, the difference in the property

sciendo

Wojciech Angielczyk
A Review of the Relaxation Models for Phase Transition Flows Centered on the Topological Aspects of the Nonequilibrium Mass Transfer Modelling

DOI 10.2478/ama-2024-0056

that is relaxed stands in the square brackets was used. Therefore, everything outside of these brackets, in a given relaxation term, is the relaxation rate of the considered property.

In [17], <Eq. 3.1.1>, an approach based on a 1D version of Eq. (1) is presented. In this equation the right-hand side is given in the following form <Eq. 3.1.8b>:

$$C = r[s_1 - s_2] + r\left[\frac{h_2}{T_2} - \frac{h_1}{T_1}\right] + r\left[\frac{L}{T_1} - \frac{L}{T_2}\right], \tag{13}$$

where s is the specific entropy, and $L = h_{SG} - h_{SL}$ is the specific latent heat. In this approach, there are three relaxation source terms that have a common relaxation rate r being a positive constant.

Another 1D equation is proposed in [33] in the following form <Eq. 43c>:

$$\frac{\partial \rho_2}{\partial t} + \frac{\partial(\rho_2 u)}{\partial x} = r[f(\rho_2) - f(\rho_1)] + r\mu(\rho_2)[\rho_1 - \rho_2], \tag{14}$$

where f is the Helmholtz free energy per unit volume. The relaxation rate in the first relaxation source term is $r = \gamma^{-1}(\rho - \rho_1)(\rho - \rho_2)$, while for the second relaxation source term it is $r\mu(\rho_2)$. Finally, γ is a relaxation parameter that determines the rate at which the chemical potentials and pressures of the two phases reach equilibrium.

The approach presented in [13], among the five model equations, contains three equations responsible for non-equilibrium mass transfer modelling. The first of them <Eq. 1a > has a form of the volume fraction advection equation (Eq. 10, $\beth = \alpha$). Thus, it is necessary to present only its right-hand side:

$$K = \nabla\boldsymbol{u}\omega[\rho_2 c_2^2 - \rho_1 c_1^2] + \rho v\left(\frac{c_2^2}{\alpha_2} + \frac{c_1^2}{\alpha_1}\right)\omega[g_2 - g_1] +$$
$$H\left(\frac{\Gamma_2}{\alpha_2} + \frac{\Gamma_1}{\alpha_1}\right)\omega[T_2 - T_1], \tag{15}$$

where $\omega = (\rho_1 c_1{}^2 \alpha_1{}^{-1} + \rho_2 c_2{}^2 \alpha_2{}^{-1})^{-1}$, v and H are relaxation coefficients, Γ is the Grüneisen coefficient. The second and third equations <Eq. 1b, Eq. 1c> are the phasic mass conservation equation (Eq. 1). Thus, it is necessary to present only their right-hand sides:

$$C = \pm \rho v[g_2 - g_1]. \tag{16}$$

Another approach using three equations (out of the six model equations) for non-equilibrium mass transfer modelling is presented in [34] and extended in [35]. The first equation is the volume fraction advection equation <Eq. 63a>, whose right-hand side reads:

$$K = \zeta_p[p_1 - p_2] + \frac{\zeta_T \hbar A_{\text{int}}}{V}[T_1 - T_2] - \frac{\zeta_G A_{\text{int}}}{V}[g_1 - g_2], \tag{17}$$

where $\hbar$ is the heat transfer coefficient, A_{int} is the interface area (the area of the interphase heat transfer), V is the volume of the considered mixture element, $\zeta_p, \zeta_T, \zeta_G$ are relaxation coefficients. The second and third equations <Eq. 63b, Eq. 63c> are the phasic mass conservation equation. Thus, it is necessary to present only their right-hand sides:

$$C = \pm \frac{A_{\text{int}}}{V}[g_1 - g_2]. \tag{18}$$

It is worth noting that Eqs. (15-18) contain relaxation terms based on the difference between phasic specific Gibbs free energy.

2.4. Relations based on the statistical phase change analysis

The fundamental works in this domain determine the net rate of molecular interfacial transport j that can be a basis for the calculation of the interphase mass transfer rate C. However, translating the equations for j, into those describing C, would significantly complicate them and impede recognition of the relaxed properties and the relaxation rates. Thus, it was decided to avoid that and to introduce only the nomenclature corrections.

In each equation (presented in this and the next section), the difference in the property that is relaxed is enclosed in curly brackets. Therefore, everything outside these brackets is the relaxation rate of the considered property.

The net rate of molecular Interfacial transport of gas into the liquid phase is derived in [36], <Eq. 33>, in the following form:

$$j = -\frac{2 A_s}{C_{eq}} \frac{p_1'}{\sqrt{2\pi m_1 k T_2}}\{C - C_{eq}\}, \tag{19}$$

where A_s is the fraction of the area available for absorbing gas molecules when the liquid is in equilibrium with the gas, p_1' is the partial pressure of the gas above the liquid, m_1 is the molar mass of the absorbed gas, k is the Boltzmann constant, C is the concentration of the gas in the liquid (number of the molecules per unit volume).

In [37] net rate of molecular interfacial transport <Eq. 57> is described as dependent on difference in phasic chemical potentials:

$$j = \frac{2K}{kT}\{\mu_1 - \mu_2\}, \tag{20}$$

where K is the equilibrium molecular exchange rate between phases, T is the two-phase mixture temperature.

Quite complex relation for net rate of molecular interfacial transport is presented in [38], <Eq. 42, Eq. 45>:

$$j = \frac{\eta\, p_{sat}(T_2)}{\sqrt{2\pi m k T_2}}\langle exp\left[\frac{\mu_2}{T_2} - \frac{\mu_1}{T_1} + h_1\left(\frac{1}{T_1} - \frac{1}{T_2}\right)\right] - exp\left[-\frac{\mu_2}{T_2} + \frac{\mu_1}{T_1} - h_1\left(\frac{1}{T_1} - \frac{1}{T_2}\right)\right]\rangle, \tag{21}$$

where $\eta = \exp\left(\frac{1}{\rho_2 k T_2}\left[p_{eq} - p_{sat}(T_2)\right]\right)$, m is the molar mass. However, it can be recognize as a relaxation equation only when transformed to the following form:

$$j = \frac{\eta\, p_{sat}(T_2)}{\sqrt{2\pi m k T_2}}\frac{1}{\dot{\rho}}\{\dot{\rho}^2 - 1\}, \tag{22}$$

where

$$\dot{\rho} = exp\left[\frac{\mu_2}{T_2} - \frac{\mu_1}{T_1} + h_1\left(\frac{1}{T_1} - \frac{1}{T_2}\right)\right].$$

In [39] an analysis of evaporation through the classical kinetic theory was conducted. The authors developed relation that reads:

$$j = \frac{\sigma}{1 - 0.5\sigma}(2\pi m k T_2)^{-0.5}\left\{\frac{p_{sat}(T_2)}{\sqrt{T_2}} - \frac{p_1}{\sqrt{T_1}}\right\}, \tag{23}$$

where σ is an empirical coefficient called the condensation coefficient.

2.5. Relations based on the nucleation theories and/or the kinetic theory of phase change

The equations presented here can also be included in the previous section because the kinetic theory of phase change and nucleation theories are closely related to the statistical phase change analysis. However, the rule here is that if the OA specifies the interphase mass transfer rate C instead of the net rate of molecular interfacial transport j, it qualifies for this section.

Of course, many equations for C derived within the theories considered here are not relaxation equations. Unfortunately, despite that some of them are used in the relaxation models. For example, in [40], an equation describing the evolution of the saturation index y, which does not have the features of a relaxation equation was introduced into the DEM. Consequently, the obtained model is not a relaxation model.

In [41], <Eq. 10>, an approach based on the steady-state version of Eq. (1) is presented (however, α is replaced by x). In this equation, the right-hand side is given in the following form, <Eq. 12>:

$$C = R_{boil} = \pm \left[\frac{\hat{\sigma}}{2-\hat{\sigma}}\right]\left(\frac{m\prime}{2\pi G_c T_{\text{sat}}}\right)^{1/2}\{p - p_{\text{sat}}\}, \tag{24}$$

where $\hat{\sigma}$ is the accommodation coefficient that represents the number of molecules passing during the phase change process (it is dependent on the flow conditions and the working fluid, thus, it needs to be adjusted according to the experimental data), G_c is the universal gas constant, m' is the molecular mass. In this approach, C is called the boiling source term R_{boil}.

The above approach is extended in [42] by enlarging the interphase mass transfer rate C by the so-called cavitation source R_{cav}:

$$C = R_{boil} + R_{cav} = R_{boil} \pm \frac{c_c\sqrt{K_T}}{\delta}\rho_1\rho_2(1-x)\left(\frac{2}{3}\frac{p_{\text{sat}}-p}{\rho_2}\right)^{0.5}, \tag{25}$$

where c_c is the cavitation constant, K_T is the turbulence kinetic energy, δ is the surface tension. It can be seen that since the last exponent on the right-hand side differs from one, R_{cav} is not strictly appropriate relaxation term.

3. TOPOLOGICAL ASPECTS OF RELAXATION MODELS

The topological analysis is crucial for understanding the elementary features of the considered models and applying the proper numerical procedures for determining the practically acceptable solutions. This analysis focuses on the steady-state version of the two-phase flow model and uses the theory of dynamical systems. It was conducted and described, in a rather detailed way, in [43] and (more recently) in [44]. Thus, solely elements of the theory that are critical for further investigation of the relaxation equations describing nonequilibrium interphase mass transfer are presented.

In case of doubts regarding the analysis, the reader is guided to the above-mentioned articles or the author's previous publications: A detailed description of the physical sense and mathematical formulation of relaxation models, as well as the Homogeneous Equilibrium Model is presented in [45], while procedures for determining the solutions are presented in [46]. Finally, practical applications of this two can be found in [47].

3.1. General form of the equation system and its transformations

Practically all known one-dimensional models of a steady-state flow can be presented in a form of the following nonlinear ordinary first order differential equation system [43]:

$$A(\sigma)\frac{d\sigma}{dz} = b(z,\sigma), \tag{26}$$

The size and elements of the matrix A and the vector σ depend on the model type. The vector σ consists of n quantities describing a thermodynamic state of the fluid, and if necessary, the velocity of the fluid. The elements of the matrix A depend only on σ's components, and b's elements additionally depend on the spatial coordinate z (specifying distance along the flow channel axis). The set of governing equations (26) supplied with the vector $\sigma_B = [\sigma_{1,B}, \sigma_{2,B}, \dots, \sigma_{n,B}]$ (describing the flow inlet conditions, the inlet is located at z_B) creates an initial-value problem. A solution to the problem is a trajectory $\sigma(z)$ in $n+1$ dimensional phase space Ω, which conventionally can be obtained by a numerical integration of the equation system (26).

The system of equations (26) can be solved with respect to the derivatives of σ's components by using Cramer's rule:

$$\frac{d\sigma_i}{dz} = \frac{N_i(z,\sigma)}{D(\sigma)}, \qquad i = 1,2,\dots,n, \tag{27}$$

where, D denotes the determinant of A, and N_i are determinants, each of which is created by replacing the i-th column of A with b. The most practically useful form of the equation system is obtained by application of the dummy parameter t [43]:

$$\frac{dz}{dt} = D, \frac{d\sigma_i}{dt} = N_i. \tag{28}$$

It is worth to notice that in the above autonomous form the independent variable is not z but the dummy parameter t.

3.2. Topological Structure of The Phase Space

Each possible state of a system is represented as a point in the phase space Ω. Thus, for example, if in the mathematical model n=3 then σ consists of 3 components, say: the pressure p, the enthalpy h, the velocity w. Consequently, the state of the fluid and flow in any cross-section of the nozzle is determined by three values of those parameters and value of the spatial coordinate z. Thus, the phase space of this example is 4-dimensional.

However, in general, the phase space is $n+1$ dimensional, thus for simplicity, the most interesting features of its structure are presented in the form of projections on a pressure p - spatial coordinate z plane depicted in Fig. 1. Accordingly, the black, green, and red curves present projections of $n+1$ dimensional trajectories on $p-z$ plane. Each solid line is a projection of a solution to the initial-value problem mentioned in the previous subsection. The inlet conditions related to those flows differ only in the velocities. Consequently, all trajectories related to the solid lines start form the same values of the inlet pressure p_B, the inlet density ρ_B, and the inlet specific enthalpy h_B but they are related to different mass flow rates.

It is necessary to distinguish three classes of points in the phase space Ω:
- Regular points at which $D \neq 0$. At each of these points the systems (26), (27) and (28) are equivalent. Any numerical

sciendo

Wojciech Angielczyk   DOI 10.2478/ama-2024-0056
A Review of the Relaxation Models for Phase Transition Flows Centered on the Topological Aspects of the Nonequilibrium Mass Transfer Modelling

forward-marching integration of system (26) that starts from the inlet conditions $B=[z_B, \sigma_{1,B}, \sigma_{2,B}, \ldots, \sigma_{n,B}]$ and passes towards the channel outlet, only through the regular points, gives a proper approximation to a physically acceptable analytic solution. The system (26) satisfies the existence and uniqueness requirements (only one trajectory passes through any regular point). A trajectory that consists of only regular points is fully subsonic or fully supersonic. Fig. 1 shows projections of three subsonic trajectories - the green curves that are called Possible Flow (PF) trajectories.

− Turning points at which $D = 0$ and and all $N_i \neq 0$ [18]. At those points the systems (26) and (27) are not equivalent. Numerical integration of (26), in the vicinity of the turning point, produces a systematically accumulating numerical error. As a result, the integration could become impossible even before reaching the turning point. This is because $|d\sigma_i/dz| \to \infty$ while $D \to 0$. However, the autonomous system (28) satisfies the existence and uniqueness requirements at those points. Hence, during its numerical integration, one can simply pass through a turning point and can obtain a proper approximation of a trajectory that, at the turning point, changes direction along the z-axis (the red curves in Fig. 1). The one-dimensional steady-state flow cannot change direction in the channel. Therefore, those trajectories are physically acceptable only if they pass through a point of inlet conditions B and the turning point is located at the end of the channel. Consequently, the solutions that pass through turning points localised inside of the channel are called Impossible Flow (IF) trajectories.
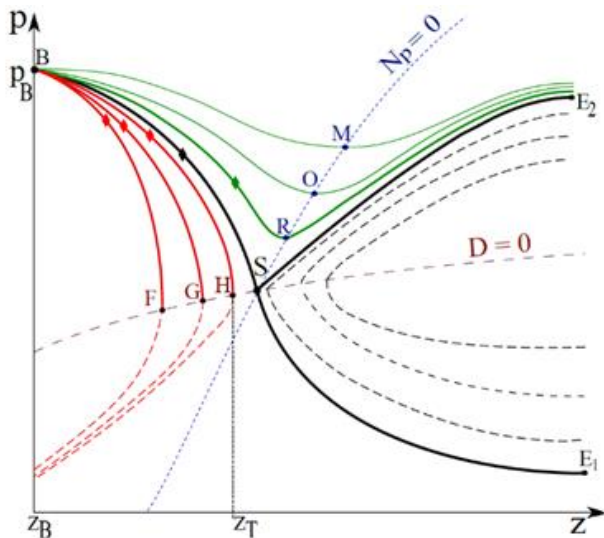
Fig. 1. Illustration of the structure of the considered phase space

In work [18] was shown that $D = 0$ occurring at the channel end is also a choking criterion (or critical flow condition) since $D = 0$ means that at this point the fluid's velocity reaches the local speed of sound, and as a result, the mass flow rate and the subsonic part of the trajectory are unchangeable even despite possible pressure drop occurring beyond the channel exit.

Fig. 1 shows projection of three IF trajectories (the red curves). They pass through the turning points F, G, H. Turning points determine the curve $D = 0$ (the projection of this curve is shown in Fig. 1 as the brown dashed line).

− Singular points at which $D = 0$ and all $N_i = 0$. Here are considered only nondegenerate singular saddle points like S at which $\mathrm{rank}(A) = n - 1$ and through which exactly two trajectories pass. Namely, *BSE$_1$* and *BSE$_2$* in Fig. 1 (but only *BSE$_1$* is "really" a transonic trajectory since on *BSE$_2$* the velocity of the fluid reaches the speed of sound merely at point *S* to decrease just after it [18]).

According to Eq. (28), $\mathrm{d}z = \mathrm{D}\mathrm{d}t$ and $\mathrm{d}\sigma_i = \mathrm{N}_i\mathrm{d}t$. Therefore, at those points, finite changes Δz and $\Delta\sigma_i$ calculated by the numerical methods are equal to 0 regardless of the integration step size Δt. It means that, the numerical algorithms cannot neither "start from" nor "pass through" this kind of points (they simply "get stuck" in these points - in the theory of differential equations, such points are called equilibrium points). Therefore, contrary to remaining trajectories, the transonic trajectory cannot be determined by a conventional numerical forward-marching integration (even of the system 28).

In [18] it was shown that when $D = 0$ and an arbitrary $N_i = 0$ then all remaining Ns also vanish.

3.3. The topological analysis of the nonequilibrium mass transfer modelling

The review of the equations describing nonequilibrium interphase mass transfer presented earlier revealed a high diversity of these equations. The equations differ in the method of derivation (the phenomenological approaches or the approaches rooted in theories that translate the microscopic behaviour of the fluid into its macroscopic properties), the type of quantity whose difference is relaxed, and the relaxation rates. However, it turns out that from the topological point of view, all these differences are not important. Crucial is the physical nature of the equation in which the relaxation is used. Namely, equations having the form of a conservation equation do not pose topological problems. However, relations in the form of an advection equation (Eq. 10) may be problematic.

To prove the above statement, let us start by noting that the conservation equations, regardless of the choice of σ's components, always contain at least two gradients of different physical quantities. On the contrary, the advection equation for some choices of σ has just one gradient.

The above statement is true for every conservation equation, but in this work, we analyse mass transfer approaches, so we need to prove it only for the mass conservation equation, Eq. (1). To do so, let us extend the divergence term from its left-hand side:

$$\nabla(\alpha_k\rho_k u_k) = \rho_k u_k \nabla\alpha_k + \alpha_k u_k \nabla\rho_k + \alpha_k\rho_k \nabla u_k. \tag{29}$$

We obtained two terms with different gradients ($\nabla\alpha_k$ and $\nabla\rho_k$) and the divergence term ($\alpha_k\rho_k\nabla u_k$).

Now, let us ponder on the advection equation, Eq. (10). It contains only one gradient $\nabla\beth_k$. The only possibility to get at least two gradients from it is to choose $\beth_k$ and σ in such a way that $\beth_k = \beth_k(\sigma_i, \sigma_j, \ldots)$. However, if this is not the case, Eq. 10 has only one gradient term that in general can be expressed as $a_c\nabla\sigma_j$, where a_c is a non-zero coefficient. Thus, the time-independent 1-D version of the advection equation (10) can be rearranged:

$$\frac{d\sigma_j}{dz} = \frac{K}{a_c}. \tag{30}$$

The next thing to do is to show that the models that contain equations with only one gradient are topologically degenerated. Let us rearrange the equation system of the flow model in the non-autonomous form (Eq. 27) into the following form:

$$\frac{d\sigma_j}{dz} D(\sigma) = N_j(z, \sigma). \qquad (31)$$

Substitution of Eq. (30) to Eq. (31) gives the equation that reads:

$$N_j(z, \sigma) = \frac{K}{a_c} D(\sigma). \qquad (32)$$

It means that when the main determinant $D = 0$ then $N_j = 0$. However, $N_j = 0$ does not vanish the remaining Ns (the author proved it for the HRM and the DEM by recalling the momentum equation in [48]).

The consequence of the demonstrated fact is that any model applying such a simple form of the closure equation ("single gradient" equation) does not possess turning points (Fig. 1, the points F, G and H). Instead, other kinds of points appear when $D = 0$. Suppose that $D \to 0$, then according to Eq. (30), $d\sigma_j/dz \to Ka_c^{-1}$, while according to Eq. (27), the remaining derivatives ($i \neq j$) approaches $\pm\infty$. However, there is no trajectory (a continuous multidimensional curve) that can satisfy such conditions (it is impossible for any curve to turn towards the opposite direction of the z-axis without being parallel to all other axes at the point of turning). In other words, the considered points refer to states that the analysed dynamical system [18] cannot archive.

In summary, the consequence of using "single gradient" closure equation is that the considered model is inconsistent with the previously described structure of the phase space (the model is degenerated). Therefore, there is a clear need for more complex closure equations. The obvious candidate to be tested is an equation containing two gradients.

To prove that the "two gradients" closure equation (thus, also any conservation equation with relaxation terms) does not introduce the topological flaw into the model, let us replace Eq. (30) with the following all-encompassing but simple relation:

$$a_{c,j} \frac{d\sigma_j}{dz} + a_{c,i} \frac{d\sigma_i}{dz} = K, \qquad (33)$$

were $a_{c,i}$ and $a_{c,j}$ are non-zero coefficients. Application of this equation to Eq. (31) results with:

$$\left(\frac{K}{a_{c,j}} - \frac{a_{c,i}}{a_{c,j}} \frac{d\sigma_i}{dz} \right) D(\sigma) = N_j(z, \sigma). \qquad (34)$$

From Eq. (27) we know that

$$d\sigma_i/dz = N_i(z, \sigma)/D(\sigma),$$

thus:

$$\frac{K}{a_{c,j}} D(\sigma) - \frac{a_{c,i}}{a_{c,j}} N_i = N_j(z, \sigma). \qquad (35)$$

We can see that when $D = 0$ then not necessarily $N_j = 0$. Hoverer, having $D = 0$ and any of Ns equal to zero, vanishes the remaining N.

This proves that the presence of at least two different gradient terms in the relaxation equation results in a model with the correct structure of the phase space (at least in the context of considered topological aspects).

4. CONCLUSIONS

This paper presents a brief succinct of the models for phase transition flow, followed by a quite detailed review of the two-phase flow relaxation models just to create an appropriate background for an in-depth review of the nonequilibrium mass transport relaxation equations used in the mentioned models. The reviewed approaches are divided into five classes: simple thermodynamic relations, empirical phenomenological relations, more advanced thermodynamic approaches, relations based on the statistical phase change analysis, relations based on the nucleation theories and/or the kinetic theory of phase change. However, it is just one of many possible systematisations. A broader categorization can be established as follows: phenomenological relations and approaches based on theories that bridge the microscopic depiction of the fluid to properties characterizing its macroscopic behaviour. The considered equations can also be classified as those that have a form of an advection equation (Eq. 10) and those having a form of phasic mass conservation (Eq. 1).

The conducted topological analysis revealed that relaxation equations taking the form of an advection equation (or more precisely equation containing only one term with a gradient of velocity-state vector component), when integrated into the flow model, can lead to a violation of the phase space structure (making the model topologically degenerated).

While the aforementioned sentence describes a potential rather than an inevitability, it becomes evident that all models linked to the equations outlined in Tab. 1 (three B-N type models and eight HRMs) and linked to Eq. (15) and Eq. (17) are inherently topologically degenerated. This does not mean that they are therefore worthless. In fact, the majority of these models have been validated through comparison with experimental results, affirming their practical utility. Then, what practical implications arise from the described degeneration? Let us start with the most certain things.

The practical ramification of the identified flaw is that the determinant described by Eq. (32) cannot be employed as the second condition (alongside $D = 0$) for determining the singular point. Moreover, this determinant is unsuitable for discerning whether the flow is Possible or Impossible. Therefore, if this determinant is used, the conventional solution algorithms of the initial-value problem (the PIF and the NCP algorithms) will not converge to the solution. In general, not being aware of the flaw may lead to the erroneous recognition of any of the turning points as a singular point. The significance of this issues diminishes when modelling unsteady flows or when a steady flow description is achieved by asymptotical convergence of the time-dependent solutions. Since, in such cases, the determination of singular points is not a pivotal step in the solution process. However, mentioned time-dependent solution methods require a significantly longer computation time.

The author also suspects that topologically degenerated models are more difficult to set up to accurately predict experimental flows. In other words, replacing the advection equations with the equations of the phasic mass conservation equation will simplify the mathematical form of correlations for the relaxation time (or the rate of relaxation) or the evolution of the saturation index.

Among the equations reviewed, many are based on the relaxation of the difference in the chemical potential and the specific Gibbs free energy. The author considers these approaches to be the most justified since these quantities are the same for both phases in equilibrium but more importantly, they appear in the first

sciendo

Wojciech Angielczyk

A Review of the Relaxation Models for Phase Transition Flows Centered on the Topological Aspects of the Nonequilibrium Mass Transfer Modelling

DOI 10.2478/ama-2024-0056

law of thermodynamics written for an open multi-component system.

According to the author, relying on the relaxation of other physical quantities (such as pressure, temperature, entropy, etc.) is not deemed a mistake, as long as only two different quantities are utilized. After all, the chemical potential and the specific Gibbs free energy can be conceptualized as functions defined by precisely two distinct intensive properties. Therefore, paradoxically, some equations classified as "more advanced thermodynamic approaches", in the author's opinion, are not physically justified.

Despite awareness of the topological flaw being able to prevent potential problems, The author recommends the use of appropriately complex closure equations to maintain the physical and mathematical consistency of the models. It was proved that closure equations with at least two gradients of different physical quantities should be employed. It is worth recalling that conservation equations exhibit this feature.

Please recall that the theoretical expression for the relaxation time depends on the second-order derivative of specific Helmholtz free energy. However, using such an approach is tantamount to formulating a non-hyperbolic model. Perhaps the physical correctness of modelling will dictate the development of new techniques for solving the considered flow cases.

As a final conclusion, let us note that in the time when the "oldest" of the models considered here were formulated, the conditions of their hyperbolicity were mainly studied. Subsequently, the models of the considered class were examined for their compliance with the second law of thermodynamics. Therefore, the author expresses curiosity about whether the field has now entered an era where scrutiny extends to verifying whether the formulated model is topologically non-degenerate. Nevertheless, the author is confident that in the forthcoming phase of the research, he will focus on delving into unravelling the connection between topological aspects and the model's adherence to the second law, as well as the model's hyperbolicity. However, this is a long-term investigation, while proposing closure equations for the DEM that are topologically correct is now a simple matter for the author and will probably be crowned with an appropriate publication soon.

REFERENCES

1. Saha P. Review of two-phase steam-water critical flow models with emphasis on thermal nonequilibrium.1977; NUREG/CR-0417. United States.
2. Richter HJ. Separated two-phase flow model: application to critical two-phase flow. International Journal of Multiphase Flow. 1983; 9(5): 511-530. ISSN 0301-9322.
https://doi.org/10.1016/0301-9322(83)90015-0
3. Ishii M, Hibiki T. Thermo-Fluid Dynamics of Two-Phase Flow. 1975. Second Edition. Springer.
4. Staedtke H. Gasdynamic Aspects of Two-Phase Flow: Hyperbolicity, Wave Propagation Phenomena and Related Numerical Methods. Wiley-VCH. 1st edition (October 6, 2006).
5. Baer MR, Nunziato, JW. A two-phase mixture theory for the deflagration-to-detonation transition (DDT) in reactive granular materials. Int. J. Multiph. Flow. 1986;12: 861-889.
6. Zhang C, Menshov I, Wang L, Shen Z. Diffuse interface relaxation model for two-phase compressible flows with diffusion processes. Journal of Computational Physics. 2022; 466: 111356. ISSN 0021-9991. https://doi.org/10.1016/j.jcp.2022.111356
7. Bdzil JB, Menikoff R, Kapila AK, Stewart DS. Two-phase modeling of deflagration-to-detonation transition in granular materials: A critical examination of modeling issues. Physics of fluids. 1999; 11: 2; 378-402. https://doi.org/10.1063/1.869887
8. Kapila AK , Menikoff R, Bdzil JB, Stewart DS. Two-phase modeling of deflagration-to-detonation transition in granular materials: Reduced equations. Physics of fluids. 2001;13(10):3002-3024. https://doi.org/10.1063/1.1398042
9. Pelanti M. Arbitrary-rate relaxation techniques for the numerical modeling of compressible two-phase flows with heat and mass transfer. International Journal of Multiphase Flow. 2022;153:104097. ISSN 0301-9322. https://doi.org/10.1016/j.ijmultiphaseflow.2022.104097
10. Saurel R, Petitpas F, Abgrall R. Modelling phase transition in metastable liquids: application to cavitating and flashing flows. Journal of Fluid Mechanics. 2008;60:313–350.
https://doi.org/10.1017/S0022112008002061
11. LeMartelot S, Nkonga B, Saurel R, Liquid and liquid-gas flows at all speeds. Journal of Computational Physics 255. 2013;53–82. https://doi.org/10.1016/j.jcp.2013.08.001
12. Lund H, Aursand P. Two-Phase Flow of CO2 with Phase Transfer. Energy Procedia. 2012;23:246-255. ISSN 1876-6102. https://doi.org/10.1016/j.egypro.2012.06.034
13. Le Martelot S, Saurel R, Nkonga B. Towards the direct numerical simulation of nucleate boiling flows. International Journal of Multiphase Flow. 2014;66:62-78. ISSN 0301-9322. https://doi.org/10.1016/j.ijmultiphaseflow.2014.06.010
14. Saurel R, Boivin P, Le Métayer O. A general formulation for cavitating, boiling and evaporating flows. Computers & Fluids. 2016; 128: 53-64, ISSN 0045-7930.
https://doi.org/10.1016/j.compfluid.2016.01.004
15. Chiapolino A, Boivin P, Saurel. A simple and fast phase transition relaxation solver for compressible multicomponent two-phase flows. Computers & Fluids. 2017; 150: 31-45. ISSN 0045-7930. https://doi.org/10.1016/j.compfluid.2017.03.022
16. Demou AD, Scapin N, Pelanti M, Brandt L. A pressure-based diffuse interface method for low-Mach multiphase flows with mass transfer. Journal of Computational Physics. 2022;448:110730. ISSN 0021-9991. https://doi.org/10.1016/j.jcp.2021.110730
17. Stewart HB, Wendroff B. Two-phase flow: models and methods. Journal of Computational Physics. 1984;56(3):363-409.
18. Bilicki Z, Kestin J. Physical aspects of the relaxation model in two-phase flow. Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences. 1990;428(1875):379-397.
19. Downar-Zapolski P, Bilicki Z, Bolle L, Franco J. The non-equilibrium relaxation model for one-dimensional flashing liquid flow. International Journal of Multiphase Flow. 1996;22(3): 473-483. ISSN 0301-9322. https://doi.org/10.1016/0301-9322(95)00078-X
20. Atkins P, de Paula J. Physical Chemistry. 2006; 8th ed. W.H. Freeman: 805-7. ISBN 0-7167-8759-8
21. Einstein A. Schallausbreitung in teilweise dissoziierten Gasen [Sound propagation in partly dissociated gases]: 380-385.
22. Mandelshtam, LI, Leontovich EM. A theory of sound absorption in liquids. Zh. Exp. Teor Fiz. 1937;7:434-449 (in Russian).
23. Bauer EG, Houdayer, GR, Sureau HM. A non-equilibrium axial flow model in application to loss-of-coolant accident analysis. The CYSTERE system code. OECD/NEA Specialist Meeting on Transient Two-phase Flow. 1976. Toronto Canada.
24. Angielczyk W, Bartosiewicz Y, Butrymowicz D, Seynhaeve J-M. 1-D modelling of supersonic carbon dioxide two-phase flow through ejector motive nozzle. International Refrigeration and Air Conditioning Conference. 2010. Purdue USA.
25. Haida M, Smolka J, Hafner A, Palacz M, Banasiak K, Nowak AJ. Modified homogeneous relaxation model for the R744 transcritical flow in a two-phase ejector, International Journal of Refrigeration. 2018;85:314-333. ISSN 0140-7007.
https://doi.org/10.1016/j.ijrefrig.2017.10.010
26. Feburie V, Giot M, Granger S, Seynhaeve J. A model for choked flow through cracks with inlet subcooling. International Journal of Multiphase Flow. 1993;19(4):541–562.
https//doi:10.1016/0301-9322(93)90087-b

27. Attou A, Seynhaeve JM. Steady-state critical two-phase flashing flow with possible multiple choking phenomenon. Part 1: physical modelling and numerical procedure. Journal of Loss Prevention in the Industries. 1999;12:335-345.
https://doi.org/10.1016/S0950-4230(98)00017-5

28. De Lorenzo M, Lafon P, Seynhaeve JM, Bartosiewicz Y. Benchmark of Delayed Equilibrium Model (DEM) and classic two-phase critical flow models against experimental data. International Journal of Multiphase Flow. 2017;92:112-130. ISSN 0301-9322.
https://doi.org/10.1016/j.ijmultiphaseflow.2017.03.004

29. Angielczyk W, Bartosiewicz Y, Butrymowicz D. Development of Delayed Equilibrium Model for CO2 convergent-divergent nozzle transonic flashing flow. International Journal of Multiphase Flow. 2020;131:103351. ISSN 0301-9322.
https://doi.org/10.1016/j.ijmultiphaseflow.2020.103351

30. Tammone C, Romei A, Persico G, Haglind F. Extension of the delayed equilibrium model to flashing flows of organic fluids in converging-diverging nozzles. International Journal of Multiphase Flow. 2024; 171:104661. ISSN 0301-9322.
https://doi.org/10.1016/j.ijmultiphaseflow.2023.104661

31. Ambroso A, Hérard J-M, Hurisse O. A method to couple HEM and HRM two-phase flow models. Computers & Fluids. 2009;38(4):738-756, ISSN 0045-7930.
https://doi.org/10.1016/j.compfluid.2008.04.016

32. Palacz M, Haida M, Smolka J, Nowak AJ, Banasiak K, Hafner A. HEM and HRM accuracy comparison for the simulation of CO2 expansion in two-phase ejectors for supermarket refrigeration systems. Applied Thermal Engineering. 2017;115:160-169. ISSN 1359-4311.
https://doi.org/10.1016/j.applthermaleng.2016.12.122

33. James F, Mathis H. A relaxation model for liquid-vapor phase change with metastability. 2015 arXiv preprint arXiv:1507.06333.
https://doi.org/10.48550/arXiv.1507.06333

34. De Lorenzo M, Lafon Ph, Pelanti M. A hyperbolic phase-transition model with non-instantaneous EoS-independent relaxation procedures. Journal of Computational Physics. 2019;379: 279-308. ISSN 0021-9991. https://doi.org/10.1016/j.jcp.2018.12.002

35. De Lorenzo M, Lafon Ph, Pelanti M, Pantano A, Di Matteo M, Bartosiewicz Y, Seynhaeve JM. A hyperbolic phase-transition model coupled to tabulated EoS for two-phase flows in fast depressurizations. Nuclear Engineering and Design. 2021;371:110954. ISSN 0029-5493. https://doi.org/10.1016/j.nucengdes.2020.110954

36. Ward CA. The rate of gas absorption at a liquid interface. The Journal of Chemical Physics. 1977; 67(1): 229-235.
https://doi.org/10.1063/1.434547

37. Ward CA, Findlay RD, Rizk M. Statistical rate theory of interfacial transport. I. Theoretical development. The Journal of Chemical Physics. 1982;76(11):5599-5605. https://doi.org/10.1063/1.442865

38. Ward CA, Fang G. Expression for predicting liquid evaporation flux: Statistical rate theory approach. Physical Review. 1999;59(1): 429.
https://doi.org/10.1103/PhysRevE.59.429

39. Schrage RW. A theoretical study of interphase mass transfer. 1953. Columbia University Press. https://doi.org/10.7312/schr90162

40. Banasiak K, Hafner A. 1D Computational model of a two-phase R744 ejector for expansion work recovery. International Journal of Thermal Sciences. 2011;50(11):2235-2247. ISSN 1290-0729.
https://doi.org/10.1016/j.ijthermalsci.2011.06.007

41. Bodys J, Smolka J, Palacz M, HaidaM, Banasiak K. Non-equilibrium approach for the simulation of CO2 expansion in two-phase ejector driven by subcritical motive pressure. International Journal of Refrigeration. 2020;114:32-46. ISSN 0140-7007.
https://doi.org/10.1016/j.ijrefrig.2020.02.015

42. Bodys J, Smolka J, Palacz M, Haida M, Banasiak K, Nowak AJ. Effect of turbulence models and cavitation intensity on the motive and suction nozzle mass flow rate prediction during a non-equilibrium expansion process in the CO2 ejector. Applied Thermal Engineering. 2022;201:117743, ISSN 1359-4311.
https://doi.org/10.1016/j.applthermaleng.2021.117743

43. Bilicki Z, Dafermos C, Kestin J, Majda G, Zeng DL. Trajectories and singular points in steady-state models of two-phase flows. International journal of multiphase flow. 1987; 13(4): 511-533.

44. De Sterck H. Critical point analysis of transonic flow profiles with heat conduction. SIAM Journal on Applied Dynamical Systems. 2007; 6(3): 645-662. https://doi.org/10.1137/060677458

45. Angielczyk W, Bartosiewicz Y, Butrymowicz , Seynhaeve, JM. 1-D modeling of supersonic carbon dioxide two-phase flow through ejector motive nozzle. International Refrigeration and Air Conditioning Conference. 2010.

46. Angielczyk W, Śmierciew K, Butrymowicz D. Application of a fast transonic trajectory determination approach in 1-D modelling of steady-state two-phase carbon dioxide flow. In E3S Web of Conferences. 2019; 128: 06005, EDP Sciences.
https://doi.org/10.1051/e3sconf/201912806005

47. Angielczyk W, Seynhaeve JM, Gagan J, Bartosiewicz Y, Butrymowicz D. Prediction of critical mass rate of flashing carbon dioxide flow in convergent-divergent nozzle. Chemical Engineering and Processing - Process Intensification. 2019; 143: 107599. ISSN 0255-2701. https://doi.org/10.1016/j.cep.2019.107599

48. Angielczyk W, Butrymowicz D. Revisiting the relaxation equations describing nonequilibrium mass transfer in the transonic homogeneous flashing flow models. Postępy w badaniach wymiany ciepła i masy: Monografia Konferencyjna XVI Sympozjum Wymiany Ciepła I Masy. 2022; 113-123. Białystok. Oficyna Wydawnicza Politechniki Białostockiej. ISBN 978-83-67185-30-1.
https://doi.org/10.24427/978-83-67185-30-1_13

Wojciech Angielczyk: (iD) https://orcid.org/0000-0001-6408-1354

COMPUTATIONAL ANALYSIS OF MHD NANOFLUID FLOW ACROSS A HEATED SQUARE CYLINDER WITH HEAT TRANSFER AND ENTROPY GENERATION

Madhu SHARMA*⊙, Bhupendra K. SHARMA**⊙, Chandan KUMAWAT***⊙, Arun K. JALAN****⊙, Neyara RADWAN*****⊙

*Department of Bioscience, CASH, Mody University of Science & Technology, Lakshmangarh, Rajasthan, India
**Department of Mathematics, Birla Institute of Technology and Science, Pilani, Rajasthan, India
***School of Computer Science and Artificial Intelligence, SR University, Warangal, Telangana, India
****Department of Mechanical Engineering, Birla Institute of Technology and Science, Pilani, Rajasthan, India
*****Industrial Engineering Department, College of Applied Sciences, AL MAAREFA UNIVERSITY, Riyadh, Saudi Arabia
*****Mechanical Department, Faculty of Engineering, Suez Canal University, El Salam District, Egypt

madhusharma5dec@gmail.com, bhupen_1402@yahoo.co.in, Chandankumawat000@gmail.com, arunjalan@pilani.bits-pilani.ac.in

Abstract: The mixed convection heat transfer of nanofluid flow in a heated square cylinder under the influence of a magnetic field is considered in this paper. ANSYS FLUENT computational fluid dynamics (CFD) software with a finite volume approach is used to solve unsteady two-dimensional Navier-Stokes and energy equations. The numerical solutions for velocity, thermal conductivity, temperature, Nusselt number and the effect of the parameters have been obtained; the intensity of the magnetic field, Richardson number, nanoparticle volume fraction, magnetic field parameter and nanoparticle diameter have also been investigated. The results indicate that as the dimensions of nanoparticles decrease, there is an observed augmentation in heat transfer rates from the square cylinder for a fixed volume concentration. This increment in heat transfer rate becomes approximately 2.5%–5% when nanoparticle size decreases from 100 nm to 30 nm for various particle volume fractions. Moreover, the magnitude of the Nusselt number enhances with the increase in magnetic field intensity and has the opposite impact on the Richardson number. The findings of the present study bear substantial implications for diverse applications, particularly in the realm of thermal management systems, where optimising heat transfer is crucial for enhancing the efficiency of electronic devices, cooling systems and other technological advancements.

Key words: CuO–Water nanofluid, MHD, heated square cylinder, heat transfer, entropy generation

1. INTRODUCTION

Liquid suspensions with nanometer-sized particles are known as nanofluids. Nanofluids are typically made up of nanoparticles ranging in size from 1 nm to 100 nm dispersed in a standard liquid such as water [1, 2], ethylene glycol [3], methanol [4] or blood [5]. The nanofluid can be used to improve the thermal properties of base fluids such as propylene glycol, ethylene glycol, water and oil, among others. They could be used in biomedical and engineering applications such as process industries, cooling, cancer therapy and a variety of others. X-rays, computers, vehicle engines, nuclear reactors, radiators and solar energy could all benefit from nanofluids. Wen and Ding [6] looked at the convective heat transfer properties of an Al2O3-water nanofluid in a copper tube and discovered that increasing the Reynolds number and volumetric ratio of particles improved the heat transfer coefficient. Shahi et al. [7] explored the laminar convective heat transfer of CuO–water nanofluid moving through a square cavity in a laminar flow regime. Bovand et al. [8] showed the effects of Al2O3 – water nanofluid on fluid flow and heat transfer around an equilateral triangle obstacle with varying orientations. Hayat et al. [9] investigated the flow of carbon water nanomaterials with coupled melting heat transfer and thermal radiation effects. The estimated findings revealed that the side-facing flow has the greatest influence of nanoparticles on heat transfer rate enhancement, while the ver-

tex-facing flow has the least. Hayat et al. [10] studied hydromagnetic tangent-hyperbolic fluid flow with varied characteristics and nanoparticles. Natural convection of a nanofluid in the presence of an electric field has recently been studied electrohydrodynamically (EHD). Sheikholeslami and Ellahi [11] investigated the hydrothermal treatment of Fe3O4-ethylene glycol nanofluid in a lid-driven chamber with a sinusoidal upper wall exposed to a non-uniform electric field. Sheikholeslami and Chamkha [12] investigated the heat transmission properties of electrohydrodynamic free convection of a Fe3O4-ethylene glycol nanofluid in a semi-annulus enclosure with a sinusoidal wall in a following investigation. The influence of a magnetic field on Fe3O4-plasma nanofluid flow in a vessel as a targeted medication delivery method was examined by Kandelousi and Ellahi [13]. It was discovered that the existence of a magnetic field had a significant impact on the flow field, and that increasing the Reynolds number and magnetic number reduced the skin friction coefficient. To enhance the energy transportation into fluid flow, Sarfarz et al. [14–15], introduced the ternary hybrid nanofluids over different surfaces such as inclined porous and spiraling disks. In these, they mixed the three different nanoparticles with the base fluid water and stated that, with introducing ternary hybrid nanoparticles, the energy storage is enhanced significantly. Flow behaviour under the effect of MHD through various geometries has been studied in a variety of cases, such as laminar mixed convection flow from a vertical surface with induced magnetic field studied by Chaudhary and Sharma

[16], analysis of heat and mass transfer for a non-Darcian porous medium [17], temperature-dependent viscosity and thermal conductivity analysis studied by Kumawat [18], and mixed convection flow through rotating channel under the presence of inclined magnetic field and joule effect discussed by Mishra and Sharma [19].

Flow past a heated or cooled square cross-section cylinder has drawn a lot of interest over the years. Such research has been motivated by its fundamental nature as well as its relevance in a wide range of engineering applications, including electronic component cooling, compact heat exchangers, combustion chambers in chemical processes, flow dividers in polymer processing applications and energy systems. Crystal growth, high-performance building insulations, multi-shield structures used in nuclear reactors, solar power collectors, food processing, float glass manufacturing, furnaces, drying technologies and other applications use mixed convection flow and heat transmission in a heated square cylinder. Due to their use, numerous writers (Sharma et al. [20], Turki et al. [21], Bouaziz et al. [22] and Hayat et al. [23]) have studied convective heat fluxes inside cavities such as triangular, trapezoidal, cylindrical, wavy, square and so on. Recently, Sharma has discussed the study of heat and mass transfer under the presence of different types of nanoparticles for the different base fluids on the various geometries with different circumstances such as Darcy-Forchheimer hybrid nanofluid flow over the rotating Riga disk using an artificial neural network approach [24], response surface optimisation discussion for the electromagnetohydrodynamic Cu-polyvinyl alcohol/water Jeffrey nanofluid [25], Bayesian regularisation networks approach for the micropolar ternary hybrid nanofluid flow of blood on the curved stretching sheet [26], Arrhenius activation energy determination for gyrotactic microorganism flows over the porous inclined stretching sheet [27].

The knowledge of thermodynamics for fluid circulation in confined regions is based on the fundamental idea of entropy generation, which measures the level of disorder within a system. The afore-mentioned principle exhibits a wide range of applications across several areas associated with thermodynamics. These fields encompass electronic systems, information systems, power collecting systems, and geothermal power plants. Dogonchi et al. [28] performed a mathematical analysis of the entropy generation on buoyancy-driven flow of $Fe_3O_4-H_2O$ nanofluid through two square cylinders under the porous enclosure condition. They used the Finite element approach to find the graphical profile of entropy and velocity in this analysis. Furthermore, they [29] extended their analysis to estimate the entropy generation profile for a porous cavity containing a base fluid mixed with nano-encapsulated phase change materials (NEPCMs) in the presence of an external heat source/sink. According to this study, the presence of the Rayleigh number has a positive impact on the entropy profile. Recently, the study of entropy analysis on water-mixed nanofluid under different fluid flow mediums with different nanoparticles performed by various researchers such as TiO2-water nanofluid through inclined U-shaped domain with non-Newtonian fluid properties [30], non-Newtonian NEPCM inside the inclined chamber under presence of heater [31], Al_2O_3-water nanofluid flows through a hexagonal-shaped geometry with a periodic magnetic field [32], and for both Newtonian and non-Newtonian models of bio-fluid through curved vessels under different types of stenosis conditions and nanofluids [33–35].

The depth analysis of the existing literature, it is evident that there is a conspicuous absence of experimental or numerical investigations pertaining to mixed convection heat transfer of nanofluids within a heated square cylinder subjected to the influence of a magnetic field. Therefore, this study presents the computational examination of CuO–water nanofluid impact on the heat transfer profile for a heated square cylinder with a magnetic field. The entropy generation also calculates different magnetic field intensities for various nanoparticle volume fractions. The introduction of nanoparticles into the fluid introduces alterations in its thermo-physical properties, thereby engendering intricate interactions among inertia, viscosity and buoyancy forces. This complexity renders the analysis of mixed convection in nanofluids a formidable task. In this work, the impact of several parameters such as nanoparticle volume fraction, nanoparticle size, magnetic field intensity, Richardson number on the velocity, thermal conductivity, Nusselt number and entropy profile have been discussed with the help of both graphs and contours.

Beyond its intrinsic scientific value, this research contributes practical insights with implications for a spectrum of contemporary applications. By providing information on heat exchange enhancement, the study addresses a pressing need for improved thermal management systems. The findings are poised to influence the design and optimisation of various technological applications, spanning electronic devices, cooling systems and other engineering domains where efficient heat transfer is paramount. Thus, the research not only expands the current scientific understanding but also directly informs and advances technology with real-world applications.

1.1. Formulation of the problem

Consider a single-phase approach for nanofluids [36, 37], where the base fluid and nanoparticles are properly blended and may thus be treated as a homogenous mixture. In addition, due to the ultra-fine and low volume fraction of the solid particles, the fluid phase and solid particles are considered to be in thermal equilibrium and travel at the same local velocity. The flow in a vertical plane channel with a built-in heated square cylinder is the system of interest in this scenario. Fig. 1 shows the computational domain and coordinate system for the setup considered in this investigation. In the channel axis, the square cylinder is symmetrically placed. With a uniform free-stream velocity of u_0 and a constant temperature of T_C at the inlet, fluid moving from bottom to top approaches the square cylinder of side h. The temperature of the square cylinder is kept constant. The square cylinder's bottom face is at a distance of X_u = 10 from the channel inlet. X_d = 15 is the distance between the body and the channel outlet. Except for the body force element in the momentum equation, all thermo-physical parameters (e.g. heat capacity and thermal conductivity) are considered to be temperature-independent in this study (Boussinesq approximation). The time-dependent, two-dimensional Navier Stokes and energy equations of incompressible nanofluid are the conservation equations that describe laminar flow and heat transport. The governing equations for nanofluid flow and heat transfer can be stated as follows using the Boussinesq approximation [38, 39]:

$$\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} = 0 \qquad (1)$$

$$\rho_{nf}\left(\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} + v\frac{\partial u}{\partial y}\right) = -\frac{\partial p}{\partial x} + \mu_{nf}\left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2}\right) +$$
$$\mu_0 M\frac{\partial \bar{H}}{\partial x} - \sigma_{nf}B_y^2 u + \sigma_{nf}B_x B_y v, \qquad (2)$$

sciendo

Madhu Sharma, Bhupendra K. Sharma, Chandan Kumawa, Arun K. Jalan, Neyara Radwan
Computational Analysis of MHD Nanofluid Flow Across a Heated Square Cylinder with Heat Transfer and Entropy Generation

DOI 10.2478/ama-2024-0057

$$\rho_{nf}\left(\frac{\partial v}{\partial t} + u\frac{\partial v}{\partial x} + v\frac{\partial v}{\partial y}\right) = -\frac{\partial p}{\partial y} + \mu_{nf}\left(\frac{\partial^2 v}{\partial x^2} + \frac{\partial^2 v}{\partial y^2}\right) +$$
$$\mu_0 M\frac{\partial \bar{H}}{\partial y} - \sigma_{nf}B_x^2 v + \sigma_{nf}B_x B_y u + \ [\emptyset\rho_{np}\beta_s +$$
$$(1-\emptyset)\rho_f]g(T - T_c), \tag{3}$$

$$(\rho C_p)_{nf}\left(\frac{\partial T}{\partial t} + u\frac{\partial T}{\partial x} + v\frac{\partial T}{\partial y}\right) = k_{nf}\left(\frac{\partial^2 T}{\partial x^2} + \frac{\partial^2 T}{\partial y^2}\right) +$$
$$\sigma_{nf}\left(uB_y - vB_x\right)^2 - \mu_0 T\frac{\partial M}{\partial T}\left(u\frac{\partial \bar{H}}{\partial x} + v\frac{\partial \bar{H}}{\partial y}\right) +$$
$$\mu_{nf}\left\{2\left(\frac{\partial u}{\partial x}\right)^2 + 2\left(\frac{\partial v}{\partial y}\right)^2 + \left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y}\right)^2\right\}, \tag{4}$$

where ρ, μ, β, α and ɸare the density, the dynamic viscosity, the coefficient of thermal expansion, the thermal diffusivity and the nanoparticles volume fraction, respectively, taking into account s for solid, subscripts f for fluid and nf for nanofluid. In the above equations, the space coordinates, time, velocities and pressure are normalised with the width of the square cylinder h, the characteristic time $\frac{h}{u_0}$, the maximum velocity of the channel inlet u_0 and the characteristic pressure $\rho_f\mu_0^2$, respectively. The dimensionless quantities are defined as follows:

$$\theta = \frac{T - T_c}{T_h - T_c}, \quad R_e = \frac{\rho_f u_o h}{\mu_f},$$

$$R_i = \frac{Gr}{R_e^2}, \quad P_r = \frac{v_f}{\alpha_f},$$

$$Gr = \frac{g\beta_f h(T_h - T_c)}{v_f^2}$$

$$\tau = \frac{t}{\frac{h}{u_0}}, \quad X = \frac{x}{h}, \quad Y = \frac{y}{h}, \quad U = \frac{u}{u_0}, \quad V = \frac{v}{u_0},$$

$$P = \frac{p}{\rho_{nf}u_0^2}$$

$$H_x^* = \frac{H_x}{H_0}, \quad H_y^* = \frac{H_y}{H_0}, \quad Mn_f = \frac{\mu_0\,\chi\,H_0^2}{\rho\,u_0^2},$$

$$N = \frac{\sigma\,\mu_0^2\,H_0^2\,h}{\rho\,u_0^2}, \quad Ec = \frac{u_0^2}{C_p\,(T_h - T)}$$

where T_h and T_c are hot and cold temperatures, respectively, v_f is the kinematic viscosity of the base fluid.

Boundary conditions: All solid walls are considered to have no-slip boundary conditions for velocities. The adiabatic properties of both channel walls are described. At the channel inlet, the normal component of velocity is considered to be zero, and the axial velocity is assumed to have a fully formed parabolic profile, given by

$$U = \frac{Y}{4}(4 - Y),$$

The convective boundary condition (CBC), at the exit of the channel is given by:

$$\frac{\partial\emptyset}{\partial t} + u_{av}\frac{\partial\emptyset}{\partial X} = 0,$$

where the variable φ is the dependent variable (U, V, θ). In comparison to the Neumann boundary condition, as described by Sohankar et al. [40] and Abbassi et al. [41], the CBC reduces the number of iterations per time step and requires a smaller upstream computing domain. The square cylinder is considered to be isothermally heated at T_h and exchanging heat with the cold

fluid flowing around it, which is at T_c at the channel inlet and is at a uniform temperature.
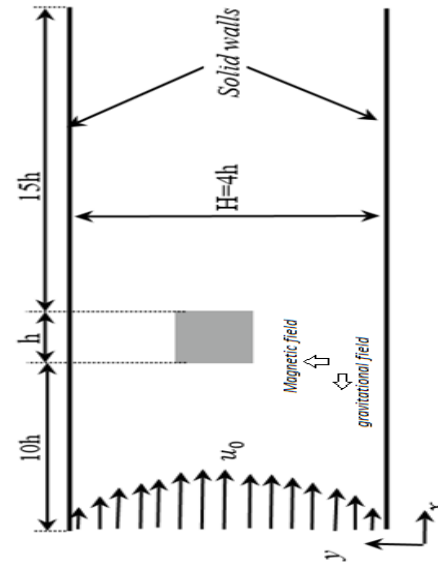
Fig. 1. Physical model of the problem

1.2. Modelling of nanofluids

Based on published relationships, the thermophysical parameters of CuO–water no fluid such as viscosity, density, specific heat and thermal conductivity are estimated.

Viscosity: Masoumi et al. [42] provided a theoretical model based on Brownian motion for predicting the effective viscosity of nanofluids. Their model was demonstrated to effectively predict the effective viscosity of various nanofluids, which is expressed as:

$$\mu_{nf} = \mu_f + \frac{\rho_{np}V_B d_{np}^2}{72C\delta},$$

where

$$V_B = \frac{1}{d_{np}}\sqrt{\frac{18KT}{\pi\rho_{np}d_{np}}}, \quad \delta = \sqrt[3]{\frac{\pi}{6\varphi}}d_{np}$$

$$C = \mu_f^{-1}\left[\left(c_1 d_{np} \times 10^9 + c_2\right)\varphi + \left(c_3 d_{np} \times 10^9 + c_4\right)\right],$$

in which

c_1=−0.000001133, c_2=−0.000002771, c_3= 0.00000009, c_4= 0.000000393.

Density: The effective density ρ_{nf} of the nanofluid is expressed as:

$$\rho_{nf} = (1 - \emptyset)\rho_f + \emptyset\rho_{np}.$$

Specific heat: The following equation, proposed by Xuan and Roetzel [43], is used to compute the specific heat of nanofluids, assuming thermal equilibrium between the base fluid and the nanoparticles:

$$(C_p)_{nf} = \frac{(1-\emptyset)(\rho C_p)_f + \emptyset(\rho C_p)_{np}}{(1-\emptyset)\rho_f + \emptyset\rho_{np}}$$

The thermal conductivity model established by Koo and Kleinstreuer [36] uses two-term functions to account for the impacts of

particle size d_{np}, particle volumetric concentration φ, temperature T, Brownian motion of nanoparticles and the characteristics of the base fluid. Vajjha and Das [44] improved their model using a larger set of data, as follows:

$$K_{nf} = \frac{(K_{np}+2K_f)-2\Phi(K_f-K_{np})}{(K_{np}+2K_f)+\Phi(K_f-K_{np})} K_f + 5.10^4 \beta \varphi \rho c_{p,f} \sqrt{\frac{kt}{\rho_{nf} d_{np}}} f(T,\varphi)$$

where

$$f(T,\varphi) = (2.8217.10^{-2}\varphi + 3.917.10^{-3})\frac{T}{T_0} + (-3.0669.10^{-2}\varphi - 3.91123.10^{-3})$$

in which, T_0 is set at 273 K and the expression of β for CuO nanoparticles is given as follows:

$$\beta = 9.881(100\,\phi)^{-0.9446},$$

which is valid for 298 $K \leq T \leq 363\ K$ and 1%$\leq \varphi \leq$6%.

The values of the nanoparticles and the base fluid properties are required for the thermophysical properties of nanofluid summarised by the preceding equations. In the current work, the CuO nanoparticle characteristics are kept constant in the operating range of 300 K to 350 K, as shown in (Tab. 1):

Tab. 1. Physical Properties of CuO nanoparticles (Dogonchi et al. [45], Abbas et al. [46], Sivaraj et Al. [47])

$\rho(kg/m^3)$	C_p(J/kg.K)	K(W/mK)	$\beta \times 10^5$(1/K)
6,350	535.6	76.5	1.61

The thermophysical properties of water are considered as a function of the temperature with the following equations (ASHRAE Handbook [48]):

$$\rho_f = -0.0036T^2 + 1.9159T + 748.19,$$

$$c_{p,f} = 0.0001T^3 + 0.1155T^2 - 0.41296T + 0.90178,$$
$$k_f = -8 \times 10^{-6}T^2 + 0.0062T - 0.5388,\ \mu_f = -0.00002414 \times 10^{(\frac{247.8}{T-140})}.$$

1.3. Non-dimensional governing equations

The dimensional governing equations will be converted into non-dimensional governing equations with the help of non-dimensional parameters.

$$\frac{\partial U}{\partial X} + \frac{\partial U}{\partial Y} = 0 \tag{5}$$

$$\left(\frac{\partial U}{\partial \tau} + U\frac{\partial U}{\partial X} + V\frac{\partial U}{\partial Y}\right) - \frac{\partial P}{\partial X} + \frac{\mu_{nf}}{\mu_f}\frac{\rho_f}{\rho_{nf}}\frac{1}{Re}\left(\frac{\partial^2 U}{\partial X^2} + \frac{\partial^2 U}{\partial Y^2}\right) + \frac{\rho_{nf}}{\rho_f}Mn_f H^* \frac{\partial H^*}{\partial X} - \frac{\sigma_{nf}}{\sigma_f}\frac{\rho_f}{\rho_{nf}}N\left(UB_y^{*2} - VB_x^* B_y^*\right) \tag{6}$$

$$\left(\frac{\partial V}{\partial \tau} + U\frac{\partial V}{\partial X} + V\frac{\partial V}{\partial Y}\right) = -\frac{\partial P}{\partial Y} + \frac{\mu_{nf}}{\mu_f}\frac{\rho_f}{\rho_{nf}}\frac{1}{Re}\left(\frac{\partial^2 V}{\partial X^2} + \frac{\partial^2 V}{\partial Y^2}\right) + \frac{\rho_{nf}}{\rho_f}Mn_f H^* \frac{\partial H^*}{\partial Y} - \frac{\sigma_{nf}}{\sigma_f}\frac{\rho_f}{\rho_{nf}}N\left(VB_x^{*2} - UB_x^* B_y^*\right) + Ri\frac{\rho_f}{\rho_{nf}}\left(1 - \phi + \phi\frac{\rho_{np}}{\rho_f}\frac{\beta_{np}}{\beta_f}\right)\theta \tag{7}$$

$$\left(\frac{\partial \theta}{\partial \tau} + U\frac{\partial \theta}{\partial X} + V\frac{\partial \theta}{\partial Y}\right) = \frac{(\rho C_p)_{nf}}{(\rho C_p)_f}\left[\frac{k_{nf}}{k_f}\frac{1}{Re\,Pr}\left(\frac{\partial^2 \theta}{\partial X^2} + \frac{\partial^2 \theta}{\partial Y^2}\right) + \frac{\sigma_{nf}}{\sigma_f}N\,Ec\,(UB_y^* - VB_x^*)^2 - Mn_f\,Ec\,\theta\,H^* \frac{\partial H^*}{\partial \theta}\left(U\frac{\partial H^*}{\partial X} + V\frac{\partial H^*}{\partial Y}\right) + \frac{\mu_{nf}}{\mu_f}\frac{Ec}{Re}\left\{2\left(\frac{\partial U}{\partial X}\right)^2 + 2\left(\frac{\partial V}{\partial Y}\right)^2 + \left(\frac{\partial U}{\partial Y} + \frac{\partial V}{\partial X}\right)^2\right\}\right] \tag{8}$$

1.4. Magnetisation equation

A large external magnetic field has been employed to recirculate flow in dilated regions and increase flow flux into the artery walls. A current plate beneath the bulges generates this magnetisation, which follows the Maxwell's law. The dimensional form of Maxwell's law is given as follows:

$$\nabla \times \bar{H} = J = \sigma(\vec{V} \times \mathrm{B} \tag{9}$$

$$\nabla.\,\mathrm{B} = \nabla.\,(\mathrm{H} + \mathrm{M}) = 0 \tag{10}$$

also

$$B = \mu_0(M + H),$$

where, μ_0 defines the magnetic permeability of the system, J denotes the electric current density, $\bar{H}$, B represents the magnetic field strength and applied magnetic induction, respectively.

The magnetic field strength and magnetic induction in x & y directions is represented as:

$$\bar{H} = (\bar{H}_x, \bar{H}_y)\ \text{and}\ B = (B_x, B_y).$$

The magnetisation process happening in the system is represented by M and it shows the direct relationship with the magnetic field strength.

The mathematical representation of magnetisation is written as follows:

$$M = \chi\ H$$

where χ describes the magnetic susceptibility.

With doing non-dimensional process using non-dimensional parameters, the above equation can be reduced as:

$$\nabla \times H^* = J = \sigma(\vec{V^*} \times B^*) \tag{11}$$

$$\nabla.\,B^* = \nabla.\,(H^* + M^*) \tag{12}$$

The magnetic field intensity of a current plate is given by:

$$H_y^* = -\frac{H_0}{2}\left[Ln\left(\frac{(x^*-x_2^*)^2+(y^*-y_0^*)^2}{(x^*-x_1^*)^2+(y^*-y_0^*)^2}\right)\right]$$

$$H_x^* = H_0\left[tan^{-1}\left(\frac{(x^*-x_2^*)}{(y^*-y_0^*)}\right) - tan^{-1}\left(\frac{(x^*-x_1^*)}{(y^*-y_0^*)}\right)\right]$$

$$H^* = \sqrt{H_y^{*2} + H_x^{*2}}$$

H_0 is magnetic field strength which depends on applied magnetic induction B, and x_1^*, x_2^*, y_0^* are the position of the horizontal plate. As previously stated, the magnetisation force is created by multiplying the magnetic field and its gradient at various points in space. As a result, the magnetic force near the plate's borders is quite high. This force can alter ferrofluid flow and cause it to deviate from its normal path. The X-component of this force tries to reverse fluid motion, while the Y-component draws particles up-

Madhu Sharma, Bhupendra K. Sharma, Chandan Kumawa, Arun K. Jalan, Neyara Radwan
Computational Analysis of MHD Nanofluid Flow Across a Heated Square Cylinder with Heat Transfer and Entropy Generation

DOI 10.2478/ama-2024-0057

wards and back into the circulating flow inside the bulges.

The magnetisation property (M*), which is an innate attribute of the ferrofluid, determines the effect of the magnetic field on the flow. This property can be described using a variety of equations, however, in this article, we'll utilise the linear formula [49] that links magnetisation to magnetic field strength and temperature:

$$M^* = \chi_m H^* \tag{13}$$

χ_m is the magnetic susceptibility and varies with temperature:

$$\chi_m = \frac{\chi_0}{1+\beta(T^*-T_0)} \tag{14}$$

χ_0, β and T_0 are constant parameters that are obtained by experimental research.

The applied magnetic induction components in non-dimensional form are written as:

$$B_x^* = (H_x^* + M^*) = (H_x^* + \chi_m H^*) = \left(H_x^* + \chi_m \sqrt{H_y^{*2} + H_x^{*2}}\right) \tag{15}$$

$$B_y^* = (H_y^* + M^*) = (H_y^* + \chi_m H^*) = \left(H_y^* + \chi_m \sqrt{H_y^{*2} + H_x^{*2}}\right) \tag{16}$$

1.5. Numerical method of solution

The non-linear governing equations of laminar mixed convection heat transfer in a heated square cylinder with a constant heat flux are solved using ANSYS FLUENT computational fluid dynamics (CFD) coupled with a finite volume approach in a heated square cylinder with constant heat flux. To execute all of the simulations, the ANSYS FLUENT commercial solver was utilised, along with User-Defined Functions that were written to include the effects of magnetic field and represent the MHD-FHD governing equations. A typical scalar transport equation is converted into an algebraic equation that can be solved numerically using the control volume-based technique. It is made up of the following steps:

Using a computational grid, divide the domain into discrete control volumes.

− Integrating the governing equations on individual control volumes to create algebraic equations for discrete dependent variables like velocities, pressure and temperature.

− Linearisation of discretised equations and solution of the resulting linear equation system to produce updated dependent variable values. Details about the solver algorithms used by ANSYS FLUENT can be found in [50].

1.6. Entropy generation

The knowledge of thermodynamics for fluid circulation in confined regions is based on the fundamental idea of entropy generation, which measures the level of disorder within a system.

By taking into account those considerations, a common expression for entropy generation may be formulated as follows [28, 33]:

$$E^{gen} = \frac{k_{nf}(\nabla T)}{T_c^2} + \frac{\mu_{nf}}{T_c}\left\{2\left(\frac{\partial u}{\partial x}\right)^2 + 2\left(\frac{\partial v}{\partial y}\right)^2 + \left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y}\right)^2\right\} + \frac{J^2}{\sigma_{nf} T_c} \tag{17}$$

Therefore, the entropy generation expression for fluid flow through a heated square cylinder is

$$E^{gen} = \frac{k_{nf}\left[\left(\frac{\partial T}{\partial x}\right)^2 + \left(\frac{\partial T}{\partial y}\right)^2\right]}{T_c^2} + \frac{\mu_{nf}}{T_c}\left\{2\left(\frac{\partial u}{\partial x}\right)^2 + 2\left(\frac{\partial v}{\partial y}\right)^2 + \left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y}\right)^2\right\} + \frac{\sigma_{nf}(uB_y-vB_x)^2}{T_c} \tag{18}$$

By using Tab. 1, the above dimensional form of entropy generation can be reduced into non-dimensional form, and it will be written as:

$$\bar{E}^{gen} = \frac{E^{gen}}{E}$$

$$\bar{E}^{gen} = \frac{k_{nf}}{k_f}\left[\left(\frac{\partial\theta}{\partial X}\right)^2 + \left(\frac{\partial\theta}{\partial Y}\right)^2\right] + \frac{\mu_{nf}}{\mu_f}\frac{Br}{AA}\left[2\left(\frac{\partial U}{\partial X}\right)^2 + 2\left(\frac{\partial V}{\partial Y}\right)^2 + \left(\frac{\partial U}{\partial X} + \frac{\partial V}{\partial Y}\right)^2\right] + \frac{\sigma_{nf}}{\sigma_f}\frac{N\,Br\,Re}{AA}(UB_y^* - VB_x^*)^2 \tag{19}$$

where

$$B_r = P_r.E_c, \quad AA = \frac{T_h-T_c}{T_c}.$$

1.7. Grid-Independence Analysis

To ensure the stability of numerical simulations, the truncation error becomes negligible when the calculations yield consistent results irrespective of grid density or sparsity. Consequently, the truncation error and the overall validity of numerical outputs hinge on the grid's independence. The reliability of numerical simulations can be significantly impacted by the stability of results with varying grid resolutions. While, in theory, a dense grid could mitigate this concern in grid-independent scenarios, it may lead to unnecessary utilisation of computational resources. Therefore, identifying the optimal grid size is crucial.

Tab. 2. Grid-independence test

Grid	$\frac{k_{nf}}{k_f}$	Nusselt number
100 ×100	1.154	5.144
200 ×200	1.125	5.118
400 ×400	1.1145	5.105
800 ×800	1.1128	5.085
1,000 ×1,000	1.1121	5.0846

In the present investigation, determining the optimal grid system involves conducting a 'grid independence test'. This test serves as a mechanism for selecting a mesh density that is both computationally accurate and economically feasible. The optimal grid size, specified as (800 × 800), is outlined in Tab. 2; further refinement of the mesh size does not enhance accuracy.

2. RESULTS AND DISCUSSION

A complete numerical research employing CuO–water nanofluid and varying magnetic field intensities to analyse laminar mixed convection heat transfer around a heated square cylinder. For various values of magnetic field, Richardson number, volume fraction, size and type (CuO–water) of nanoparticles, numerical simulation results are reported in terms of streamlines, isotherms, velocity magnitude, temperature and average Nusselt number. When compared to some previously published experimental results of CuO–water nanofluids, the current data is found to be similar to Lee et al. [54], implying that they might be used as heat transfer fluids. Tab. 3, provides the default values for different physical parameters and their source of references.

Tab. 3. Range of physical parameters with sources

Parameter	Range	Source
Re	$1 - 1,000$	Uddin et al. [51]
Mn_f	$0 - 2.02 \times 10^{12}$	Abdi et al. [52]
N	$0 - 0.1$	Tzirtzilakis and Xenos [53]
P_r	4.623	Abdi et al. [52]
R_i	$0 - 10$	Shahi et al. [7]

Figs. 2a and b demonstrate the validation process for the mathematical model formulation and computational method et al. [22], subject to certain assumptions. Specifically, our model disregards the influence of the magnetic field, validation of the Nusselt number is presented graphically, while Fig. 2b illustrates the validation of velocity through nanoparticle absence and the effect of viscous dissipation, while the Newtonian fluid properties have been considered instead of considering non-Newtonian fluid (Power-law fluid) in Bouaziz et al.'s[22] work. The validation process follows accuracy employed in the current research. In Fig. 2a, the contour representation. This validation is conducted by comparing the current research with the findings of Bouaziz the control volume finite element method used in [22], and the simulation of the present work is executed in ANSYS FLUENT Software employing the finite volume methodology. Fig. 2a demonstrates a substantial agreement in Nusselt number between the current study and the observations in [22]. Similarly, Fig. 2b depicts the validation of velocity contours, aligning well with the outcomes of the prior work [22].
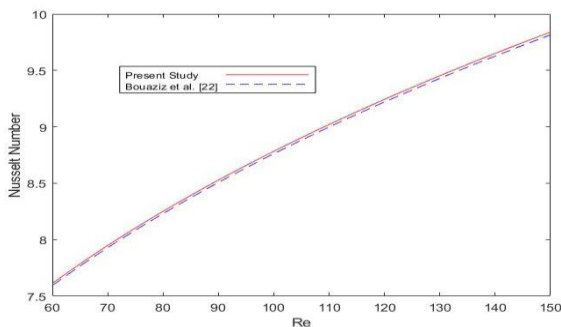
Fig. 2 (a) Validation of Present study for Nusselt number with Bouaziz et al. [22]
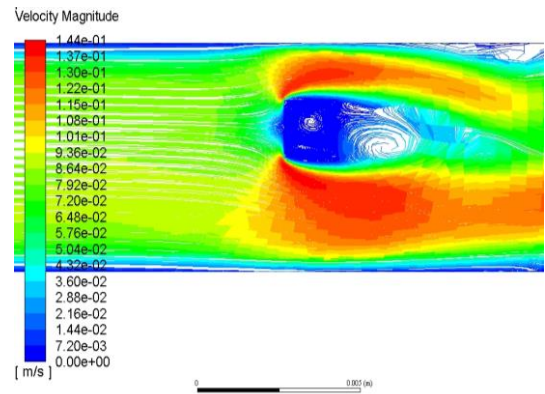
Fig. 2 (b) Contour of velocity magnitude
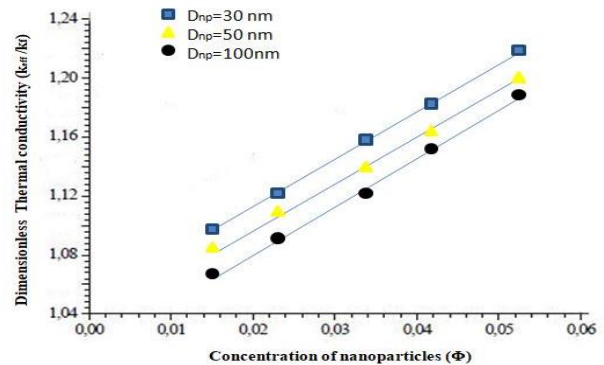
Fig. 3. Thermal conductivity vs concentration of nanoparticles for different values of nanoparticles diameter

Fig. 3 shows the effective thermal conductivity, defined as keff/kf, as a function of particle volume percent and nanoparticle diameter at 300 K. This figure shows that nanoparticles CuO is more thermally conductive than the base fluid water, therefore, when mixed into water, it increases the nanofluid thermal conductivity. Furthermore, as the volume fraction of CuO nanoparticles in the flow increases, the thermal conductivity ratio improves, and this improvement is approximately 12.5% for 0.05 volume fraction when compared to 0.01 volume fraction. This is consistent with other studies, which show that the thermal conductivity of nanofluids rises linearly with particle loading (Philip et al. 55, Shima et al. 56). It is also observed that, for a fixed particle volume fraction, when the smaller diametric size nanoparticles are added into water, it produced more thermal conductivity compared with mixing large diametric nanoparticles. This happens because, small diametric particles quickly dissolve into water and have a higher efficiency of forming strong bonds with water molecules. Therefore, from the figure, it can be easily seen that the thermal conductivity of a particle with a diameter of 30 nm is higher than that of a particle with a diameter of 50 nm or 100 nm.

In Fig. 4, the impact of varied nanoparticle sizes on thermal conductivity (keff/kf) is depicted across different temperatures. It is essential to highlight that as the temperature (T) increases, there is a corresponding increase in thermal conductivity. This phenomenon can be attributed to the incorporation of nanoparticles into the base fluid (water), resulting in an elevation of the fluid's thermal conductivity. Furthermore, the presence of ultra-fine particles induces irregular and chaotic movement within the fluid, thereby enhancing its energy exchange rates [57]. Comparing our findings with previously published experimental results on CuO–water nanofluids, particularly those by Das et al. [58] and Liu et al. [59],

a consistent pattern emerges. The observed similarity suggests robustness and reliability in the behaviour of nanofluids containing different sizes of CuO nanoparticles. Specifically, Fig. 4 elucidates that nanofluids based on water and featuring 30 nm CuO nanoparticles exhibit a significantly higher sensitivity to temperature changes compared to those incorporating particles of sizes 50 nm and 100 nm. This heightened temperature sensitivity underscores the intricate interplay between nanoparticle dimensions and their influence on the thermal properties of the nanofluid system.
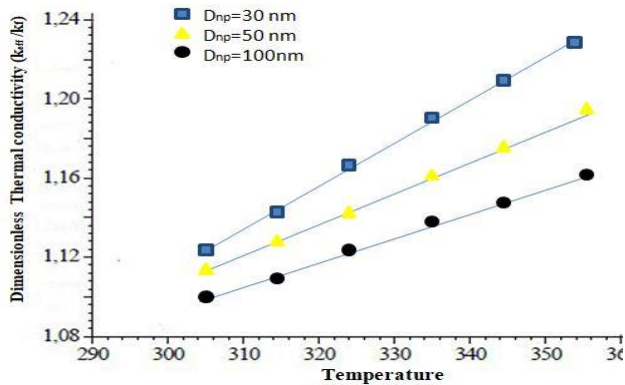
Fig. 4. Thermal conductivity for different values of nanoparticles diameter
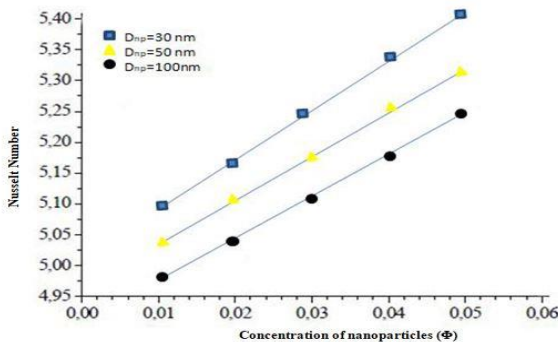
Fig. 5. Nusselt number vs concentration of nanoparticles for different values of nanoparticle diameter

Fig. 5 shows the Nusselt number in relation to nanoparticle concentration for various values of nanoparticle diameter. The heat transfer rate increases monotonically as *dnp* grows for each value of *dnp*, and this enhancement is found to be more substantial for *dnp* = 30 nm than 50 nm and 100 nm. Furthermore, the higher thermal conductivity of the nanofluid is linked to the sensitivity of thermal boundary layer thickness with mass fraction. As a result, higher thermal conductivity values are coupled with higher thermal diffusivity values. As shown, a high degree of thermal diffusivity produces a decrease in temperature gradients and, as a result, an increase in boundary thickness. The Nusselt number is reduced when the thermal boundary layer thickness increases, however, the Nusselt number is a multiplication of the temperature gradient and the heat transfer coefficient. Because the reduction in temperature gradient caused by the presence of nanoparticles is significantly smaller than the thermal conductivity ratio, increasing the bulk fraction improves Nusselt. According to Martnez-Cuencaco et al. [60], the heat transfer improvement achieved in nanofluids is primarily due to a Pr number change (viscosity

change). Buschmann [61] also determined that using a combination of Reynolds and Prandtl numbers to describe laminar nanofluid pipe flow with inserted twisted tape is sufficient because two-phase flow effects such as Brownian and thermophoretic diffusion are minor.
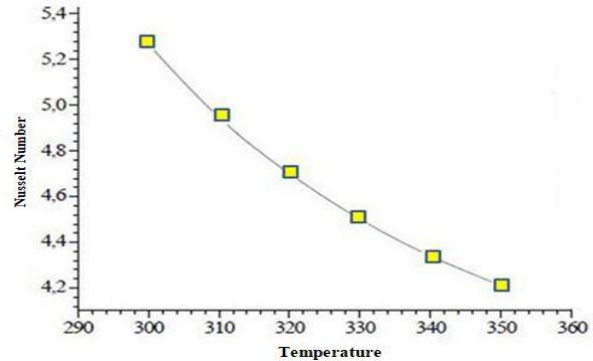
Fig. 6. Nusselt number vs temperature at φ = 5%

In Fig. 6, the variation in global time-averaged Nusselt number under the presence of 5% CuO nanoparticle, is graphically represented in relation to the temperature variable T, specifically over the heat transmission surface of the square cylinder. An interesting phenomenon is visible, in which an increase in temperature T coincides with an observable decrease in heat transfer rate. The diminution in heat transfer with escalating temperature is attributable to a complex interplay of physical factors. One key contributor is the alteration in fluid properties induced by changing temperatures. The transition from T = 300 K to T = 350 K leads to variations in the fluid's viscosity and density, influencing the dynamics of the boundary layer and subsequently affecting the efficiency of convective heat transfer. Additionally, fluctuations in the thermal conductivity of the fluid contribute to its altered capacity for heat conduction. Therefore, a quantified 22% decrease in heat transfer when transitioning from T = 300 K to T = 350 K underscores the system's sensitivity to thermal variations.

Fig. 7 illustrates the fluctuation of the Nusselt number concerning diverse concentrations of nanoparticles under distinct magnetic field strengths. It is evident that the Nusselt number exhibits an upward trend with escalating particle concentration and magnetic field strength. Notably, when the magnetic field intensity is <200 T, the rise in the Nusselt number appears to be more restrained, registering a moderate increase of 6%. However, when the magnetic field strength surpasses 200 T, there is a pronounced and steep ascent in the Nusselt number, reaching an impressive 35%. The observed correlation between increasing particle concentration, magnetic field strength, and the Nusselt number can be elucidated through the underlying physical processes. At lower magnetic field intensities, the response of nanoparticles to the magnetic field is comparatively subdued, resulting in a more controlled enhancement of the Nusselt number. As the magnetic field strength >200 T, the intensified magnetic forces exerted on the nanoparticles lead to a more substantial impact on fluid flow and heat transfer. This heightened influence manifests in a sharp escalation of the Nusselt number, indicating a more pronounced augmentation in convective heat transfer.

Fig. 8 depicts the relationship between the global time-averaged Nusselt Number and the Richardson number for various magnetic field intensities. It is evident that the Nusselt number

falls as the Richardson number rises, whereas the magnetic field has the opposite effect. When the magnetic field strength is increased from 400 T to 500 T, there is a significant difference. The Lorentz force increases as the magnetic field intensity increases, causing a reduction in fluid flow convection in the cylinder and, as a result, an increase in heat transfer across the cylinder as more heat transfer occurs in the system.
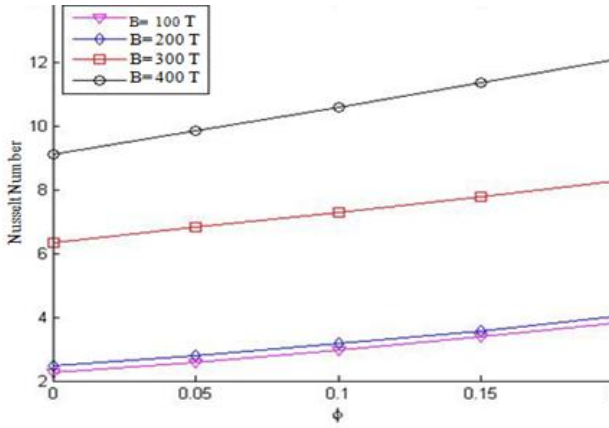
Fig. 7. Nusselt number vs concentration of nanoparticles for different values of magnetic field
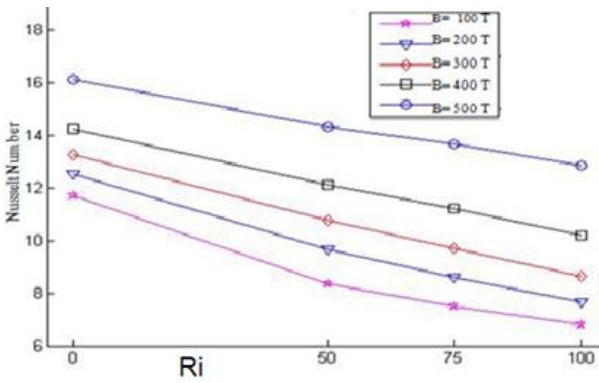
Fig. 8. Nusselt number vs Richardson number for different values of magnetic field
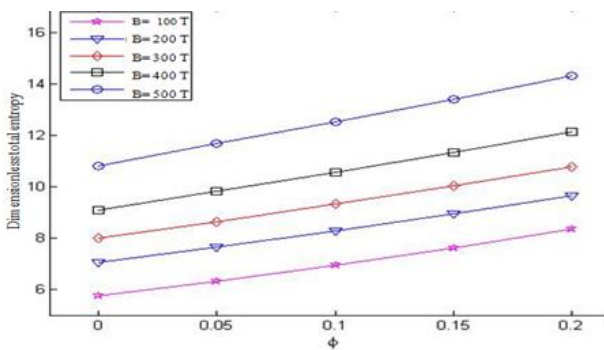
Fig. 9. Entropy for varying intensity of magnetic field

Fig. 9 shows how the dimensionless total entropy changes when the nanoparticle concentration changes at different magnetic strengths. Entropy is basically, the unused energy which further cannot be used for doing mechanical work. The nanoparticles are higher thermal conductive than the water, so when particles are mixed with the water, it enhances the base fluid's thermal conduc-

tivity, therefore the temperature of nanofluid enhances significantly according to the volume percentage of nanoparticles. This increased temperature, produces more energy into the flow, but due to sudden surge in energy, most of the energy is not taken for useful works. As a result, the entropy of the flow rises with the mixing of nanoparticles which have more thermal conductivity than the base fluid. Also, it is observed that flow temperature enhances significantly with the increase in magnetic field intensity, due to this, the flow generates more entropy. Therefore, it can be seen through the figure that with the increase in magnetic field intensity, the entropy production into the flow rises significantly.

Contours of static temperature for different sizes of nanofluids are shown in Figs. 10a–c. It is noticed that the contours become less condensed in the region of the top of the cylinder when the size of the nanoparticle decreases. However, when the nanoparticle size increases, the temperature contours are slightly shifted away from the cylinder's surface, resulting in a low-temperature gradient visible from the contours, especially near the rear face, and the thermal boundary layer around the solid walls of the square cylinder is found to be thick. It is obvious from these findings that the nanoparticle dimension dnp = 30 nm has a major impact on the cylinder's stability. Furthermore, the heat transfer rate improves significantly when the size of nanoparticles decreases at the same concentration. Also, the colour range from blue to red represents the temperature variation in the cylinder from minimum to high. It is noted that the length of temperature contours increases when the size of nanoparticle decreases. This is in agreement with the previous study on nanofluids under the influence of magnetic field.
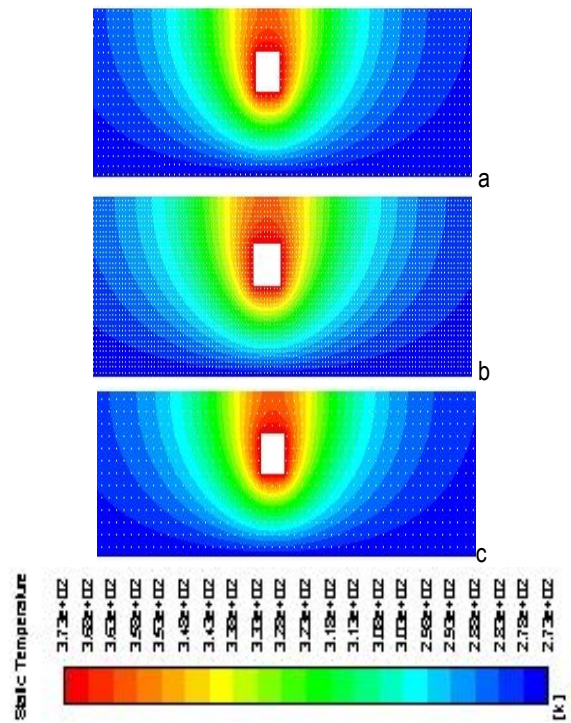
Fig. 10. Contour of total temperature for B = 100 T and dnp = (a) 30 nm (b) 50 nm (c) 80 nm

Figs. 11a–c show the contour of velocity magnitude in the cylinder of CuO–water nanofluid with magnetic field B = 100 T and different sizes of nanoparticles. It can be noted that, for dnp = 50 nm and dnp = 80 nm, the periodic flow is defined by the alternate shedding of vortices into the stream from the bottom face of the

square cylinder. Vortex shedding on both sides of the square cylinder is suppressed at dnp = 30 nm. Vortex shedding is a broad term for a variety of physical phenomena. Vortices are shed alternately from the top and bottom of the cylinder when this happens. The physical changes in the local flow are always followed by the shedding of vortices. It causes pressure and flow velocity to fluctuate in the cylinder's immediate vicinity.
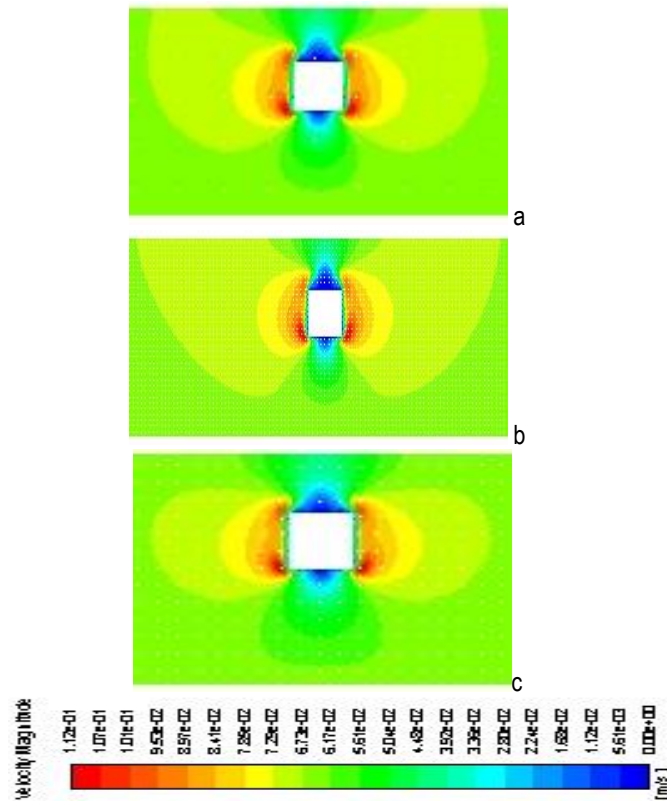
Fig. 11. Contour of velocity magnitude with magnetic field B = 100 T and dnp = (a) 30 nm (b) 50 nm (c) 80 nm
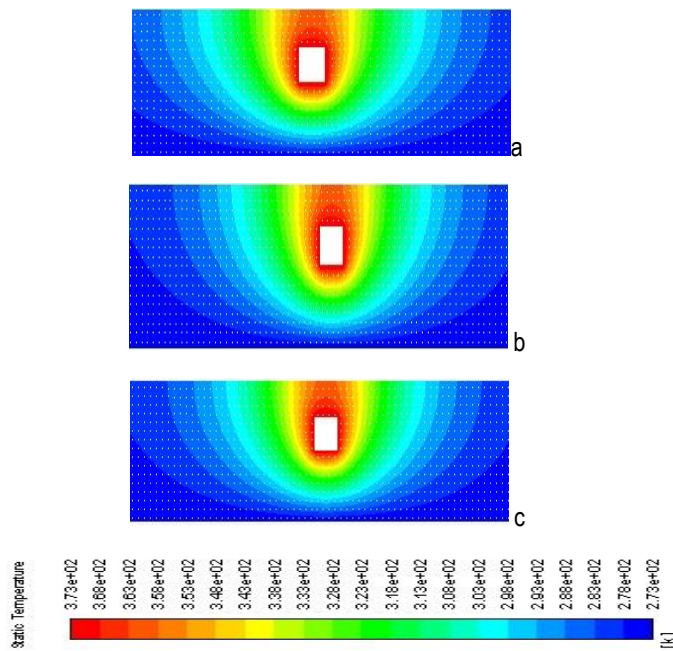
Fig. 12. Contour of static temperature for dnp = 30 nm and magnetic field B (a) 100 T (b) 200 T (c) 300 T
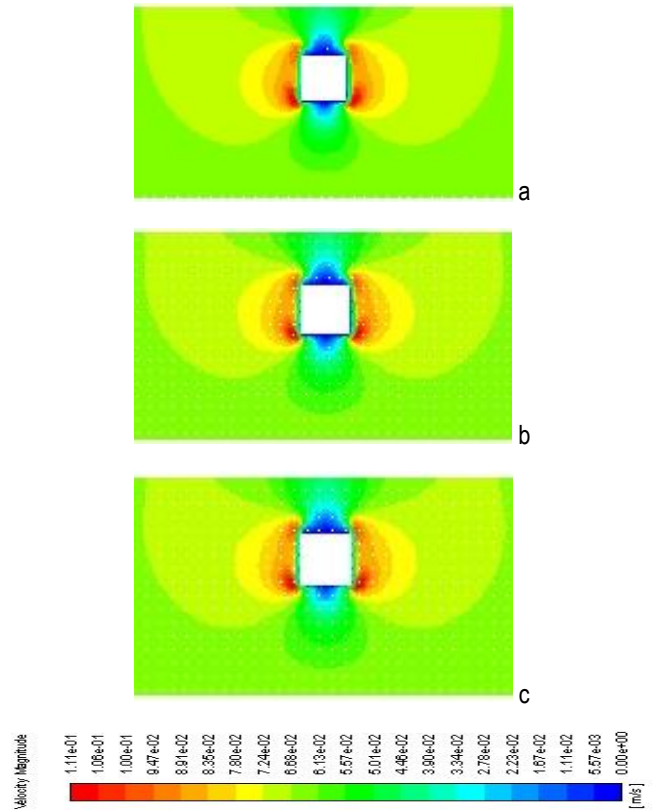
Fig. 13. Contour of velocity magnitude for magnetic field B = (a) 100 T (b) 200 T (c) 300 T

Contours of static temperature for different intensities of magnetic field are shown in Figs. 12a–c. It is observed that the temperature contours are much denser near the front surface of the cylinder, which explains the higher temperature gradient. By increasing the magnetic field strength, the secondary eddy gradually became bigger and the primary eddy became smaller. The isotherms show that the effect of the magnetic field is suppressing the convective heat transfer mechanism. This is due to uniformly distributed isotherms in the cavity, especially at the bottom. It is also clear that the heat transfer curve reaches a peak at the corners of the top and bottom surfaces of the cylinder which is due to the high temperature gradients at these points.

Contours of velocity magnitude for dnp = 30 nm in the cylinder with different intensities of magnetic fields are shown in Fig. 13a–c. When the intensity of the magnetic field drops, the length of the velocity contours increases. This means that the magnetic field causes the fluid flow to slow down, resulting in higher heat transfer in the system. As a result of the increased magnetic field, the heat transfer across the cylinder increases, and the velocity of the CuO–water nanofluid drops. The Lorentz force increases as the magnetic field strength increases, resulting in less fluid flow convection in the cylinder, as evidenced by the decrease in vortexes as the magnetic field intensity rises.

Fig. 14 depicts the streamlines of the fluid crossing the square cylinder in the channel at various intensities of the applied magnetic field (a–c). With a rise in the magnetic field, the heat convection across the cylinder increases but the velocity falls marginally. The periodic flow defined by the alternate shedding of vortices from the bottom face of the square cylinder into the stream may be seen for B = 200 T and B = 300 T. Vortex shedding is suppressed at the top of the square cylinder when B = 100 T.
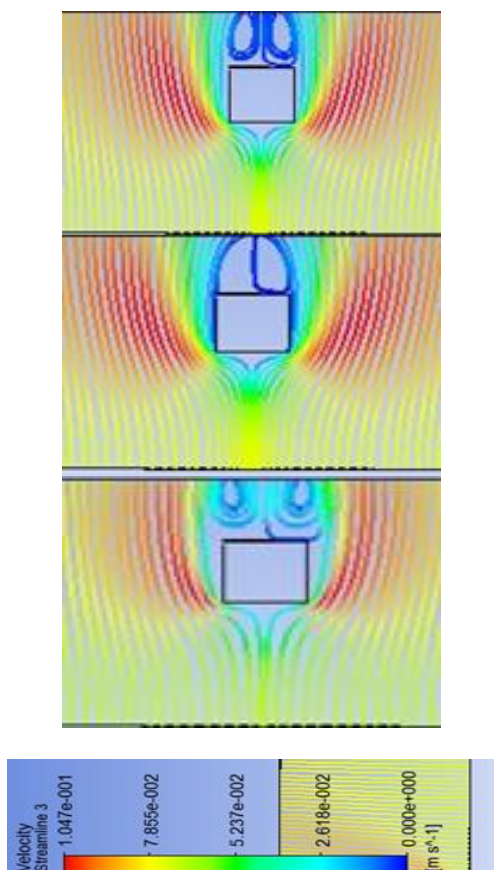
Fig. 14. Streamlines Velocity in the cylinder with dnp = 30 nm and the influence of magnetic field B (a) 100 T (b) 200 T (c) 300 T

3. CONCLUSIONS

A numerical analysis of unsteady, laminar flow and heat transfer of CuO–water nanofluid past a heated square cylinder inside a vertical channel has been discussed in the presence of a magnetic field. The entropy generation also calculates different magnetic field intensities for various nanoparticle volume fractions. In this work, the impact of several parameters such as nanoparticle volume fraction, nanoparticle size, magnetic field intensity, Richardson number on the velocity, thermal conductivity, Nusselt number and entropy profile have been discussed with the help of both graphs and contours. This computational study is performed using the ANSYS FLUENT CFD software, which uses the finite volume approach for solving governing equations. The findings are poised to influence the design and optimisation of various technological applications, spanning electronic devices, cooling systems and other engineering domains where efficient heat transfer is paramount. Thus, the research not only expands the current scientific understanding but also directly informs and advances technology with real-world applications. The following is the conclusion reached as a result of this research:

− As the Richardson number rises, it induces a reduction in the Nusselt number, while conversely, the intensity of the magnetic field exerts a contrasting influence, leading to an augmentation in its effect.
− Diminishing the nanoparticle size results in an elevation of the Nusselt number at a specified concentration, whereas amplifying the nanoparticle concentration leads to an increased Nusselt number at a given particle size.

− The incorporation of nanoparticles improves thermal conductivity and alters the structure of the flow field, consequently leading to heightened heat transfer.
− The stability of the cylinder is notably influenced by the nanoparticle diameter, specifically when the diameter is observed to be *dnp* = 30 nm.
− The entropy of a flow is directly related to nanoparticle volume fraction and magnetic field intensity.

List of Symbols:
C_p= specific heat of the fluid
H = channel width
d = diameter
h = side length of a square cylinder
t = time
k = thermal conductivity
L = length of the channel
P_r = Prandtl number
R_i = Richardson number
T = dimensional temperature
t = time nondimensionalised by u_0/h
u_o= mean cHannel inlet velocity
V = velocity vector numdimensionalised by U_o
X_u= distance from body to inlet
X_d= distance from body to outlet
μ_f= base fluid viscocity
μ_{nf} = nanofluid viscocity
K_s= thermal conductivity of metallic nanoparticle
K_f= thermal conductivity of base fluid
K_{nf}= thermal conductivity of nanofluid
ρ_s= metallic nanoparticle density
ρ_f= base fluid density
ρ_{nf}= nanofluid density
$(\rho C_p)_f$= heat capacitance of base fluid
$(\rho C_p)_{nf}$= heat capacitance of nanofluid
$(\rho\gamma)_{nf}$= thermal expansion coefficient of nanofluid
ϕ = volume fraction
G_r= Grashof number
R_e= Reynolds number
B_r= Brinkman number
E_c= Eckert number
E^{gen} = entropy generation
B_{np}= coefficient of thermal expansion
B_x= x-component of the magnetic flux density
B_y= y-component of the magnetic flux density

REFERENCES

1. Seyyedi SM, Hashemi-Tilehnoee M, del Barrio EP, Dogonchi AS, Sharifpur M. Analysis of magneto-natural-convection flow in a semi-annulus enclosure filled with a micropolar-nanofluid; a computational framework using CVFEM and FVM. Journal of Magnetism and Magnetic Materials. 2023; 568:170407. https://doi.org/10.1016/j.jmmm.2023.170407
2. Abbas N, Nadeem S, Issakhov A. Transportation of modified nanofluid flow with time dependent viscosity over a Riga plate: exponentially stretching. Ain Shams Engineering Journal. 2021;12(4):3967-73. https://doi.org/10.1016/j.asej.2021.01.034
3. Lee S, Choi SS, Li SA, Eastman JA. Measuring thermal conductivity of fluids containing oxide nanoparticles. 1999;121:280–289. https://doi.org/10.1115/1.2825978
4. Mostafizur RM, Saidur R, Aziz AA, Bhuiyan MH. Thermophysical properties of methanol based Al2O3 nanofluids. International Journal of Heat and Mass Transfer. 2015;85:414-9. https://doi.org/10.1016/j.ijheatmasstransfer.2015.01.075

Madhu Sharma, Bhupendra K. Sharma, Chandan Kumawa, Arun K. Jalan, Neyara Radwan
Computational Analysis of MHD Nanofluid Flow Across a Heated Square Cylinder with Heat Transfer and Entropy Generation

DOI 10.2478/ama-2024-0057

5. Sharma BK, Kumawat C, Bhatti MM. Optimizing energy generation in power-law nanofluid flow through curved arteries with gold nanoparticles. Numerical Heat Transfer, Part A: Applications. 2023;1-33. https://doi.org/10.1080/10407782.2023.2232123

6. Wen D, Ding Y. Experimental investigation into convective heat transfer of nanofluids at the entrance region under laminar flow conditions. International journal of heat and mass transfer. 2004;47(24):5181-8. https://doi.org/10.1016/ j.ijheatmasstransfer.2004.07.012

7. Shahi M, Mahmoudi AH, Talebi F. Numerical study of mixed convective cooling in a square cavity ventilated and partially heated from the below utilizing nanofluid. International Communications in Heat and Mass Transfer. 2010;37(2):201-13. https://doi.org/10.1016/j.icheatmasstransfer.2009.10.002

8. Bovand M, Rashidi S, Esfahani JA. Enhancement of heat transfer by nanofluids and orientations of the equilateral triangular obstacle. Energy conversion and management. 2015;97:212-23. https://doi.org/10.1016/j.enconman.2015.03.042

9. Hayat T, Khan MI, Waqas M, Alsaedi A, Farooq M. Numerical simulation for melting heat transfer and radiation effects in stagnation point flow of carbon–water nanofluid. Computer methods in applied mechanics and engineering. 2017;315:1011-24. https://doi.org/10.1016/j.cma.2016.11.033

10. Hayat T, Waqas M, Alsaedi A, Bashir G, Alzahrani F. Magnetohydrodynamic (MHD) stretched flow of tangent hyperbolic nanoliquid with variable thickness. Journal of molecular liquids. 2017 Mar 1;229:178-84. Available from : https://doi.org/10.1016/j.molliq.2016.12.058

11. Sheikholeslami M, Ellahi R. Electrohydrodynamic nanofluid hydrothermal treatment in an enclosure with sinusoidal upper wall. Applied Sciences. 2015;5(3):294-306. https://doi.org/10.3390/app5030294

12. Sheikholelami M, Chamkha AJ. Electrohydrodynamic free convection heat transfer of a nanofluid in a semi-annulus enclosure with a sinusoidal wall. Numerical Heat Transfer, Part A: Applications. 2016;69(7):781-93. https://doi.org/10.1080/10407782.2015.1090819

13. Kandelousi MS, Ellahi R. Simulation of ferrofluid flow for magnetic drug targeting using the lattice Boltzmann method. Zeitschrift für Naturforschung A. 2015;70(2):115-24. https://doi.org/10.1515/zna-2014-0258

14. Sarfraz M, Khan M, Al-Zubaidi A, Saleem S. Tribology-informed analysis of convective energy transfer in ternary hybrid nanofluids on inclined porous surfaces. Tribology International. 2023;188:108860. https://doi.org/10.1016/j.triboint.2023.108860

15. Sarfraz M, Khan M, Al-Zubaidi A, Saleem S. Enhancing energy transport in Homann stagnation-point flow over a spiraling disk with ternary hybrid nanofluids. Case Studies in Thermal Engineering. 2023;49:103134. https://doi.org/10.1016/j.csite.2023.103134

16. Chaudhary RC, Sharma BK. Combined heat and mass transfer by laminar mixed convection flow from a vertical surface with induced magnetic field. Journal of Applied Physics. 2006;99(3):034901. https://doi.org/10.1063/1.2161817

17. Sharma BK, Mishra A, Gupta S. Heat and mass transfer in magneto-biofluid flow through a non-Darcian porous medium with Joule effect. Journal of Engineering Physics and Thermophysics. 2013;86:766-74.

17. https://link.springer.com/article/10.1007/s10891-013-0893-0

18. Raj Kumawat S, Vyas H, Mohan R, Sasidharan R, Yadav B, Gupta N. 90 versus 60 min of early skin-to-skin contact on exclusive breast-feeding rate in healthy infants'≥ 35 weeks: A randomised controlled trial. Acta Paediatrica. 2024;113(2):199-205. https://doi.org/10.1111/apa.17021

19. Mishra A, Sharma BK. MHD mixed convection flow in a rotating channel in the presence of an inclined magnetic field with the Hall effect. Journal of Engineering Physics and Thermophysics. 2017; 90:1488-99. https://doi.org/10.1007/s10891-017-1710-y

20. Sharma S, Maiti DK, Alam MM, Sharma BK. Nanofluid flow and heat transfer from heated square cylinder in the presence of upstream rectangular cylinder under Couette-Poiseuille flow. Wind Struct. 2019;29(1):65-75. https://doi.org/10.12989/was.2019.29.1.065

21. Turki S, Abbassi H, Nasrallah SB. Effect of the blockage ratio on the flow in a channel with a built-in square cylinder. Computational Mechanics. 2003;33:22-9. https://doi.org/10.1007/s00466-003-0496-2

22. Bouaziz M, Kessentini S, Turki S. Numerical prediction of flow and heat transfer of power-law fluids in a plane channel with a built-in heated square cylinder. International Journal of Heat and Mass Transfer. 2010;53(23-24):5420-9. https://doi.org/10.1016/j.ijheatmasstransfer.2010.07.014

23. Hayat T, Anwar MS, Farooq M, Alsaedi A. Mixed convection flow of viscoelastic fluid by a stretching cylinder with heat transfer. Plos one. 2015;10(3):e0118815. https://doi.org/10.1371/journal.pone.0118815

24. Sharma BK, Sharma P, Mishra NK, Fernandez-Gamiz U. Darcy-Forchheimer hybrid nanofluid flow over the rotating Riga disk in the presence of chemical reaction: artificial neural network approach. Alexandria Engineering Journal. 2023;76:101-30. https://doi.org/10.1016/j.aej.2023.06.014

25. Kumar A, Sharma BK, Gandhi R, Mishra NK, Bhatti MM. Response surface optimization for the electromagnetohydrodynamic Cu-polyvinyl alcohol/water Jeffrey nanofluid flow with an exponential heat source. Journal of Magnetism and Magnetic Materials. 2023;576:170751. https://doi.org/10.1016/j.jmmm.2023.170751

26. Sharma BK, Sharma P, Mishra NK, Noeiaghdam S, Fernandez-Gamiz U. Bayesian regularization networks for micropolar ternary hybrid nanofluid flow of blood with homogeneous and heterogeneous reactions: Entropy generation optimization. Alexandria Engineering Journal. 2023;77:127-48. https://doi.org/10.1016/j.aej.2023.06.080

27. Sharma BK, Khanduri U, Mishra NK, Chamkha AJ. Analysis of Arrhenius activation energy on magnetohydrodynamic gyrotactic microorganism flow through porous medium over an inclined stretching sheet with thermophoresis and Brownian motion. Proceedings of the Institution of Mechanical Engineers, Part E: Journal of Process Mechanical Engineering. 2023;237(5):1900-14. https://doi.org/10.1177/09544089221128768

28. Dogonchi AS, Mishra SR, Chamkha AJ, Ghodrat M, Elmasry Y, Alhumade H. Thermal and entropy analyses on buoyancy-driven flow of nanofluid inside a porous enclosure with two square cylinders: Finite element method. Case Studies in Thermal Engineering. 2021;27:101298. https://doi.org/10.1016/j.csite.2021.101298

29. Afshar SR, Mishra SR, Dogonchi AS, Karimi N, Chamkha AJ, Abulkhair H. Dissection of entropy production for the free convection of NEPCMs-filled porous wavy enclosure subject to volumetric heat source/sink. Journal of the Taiwan Institute of Chemical Engineers. 2021;128:98-113. https://doi.org/10.1016/j.jtice.2021.09.006

30. Shao W, Nayak MK, El-Sapa S, Chamkha AJ, Shah NA, Galal AM. Entropy optimization of non-Newtonian nanofluid natural convection in an inclined U-shaped domain with a hot tree-like baffle inside and considering exothermic reaction. Journal of the Taiwan Institute of Chemical Engineers. 2023;148:104990. https://doi.org/10.1016/j.jtice.2023.104990

31. Dogonchi AS, Bondareva NS, Sheremet MA, El-Sapa S, Chamkha AJ, Shah NA. Entropy generation and heat transfer performance analysis of a non-Newtonian NEPCM in an inclined chamber with complicated heater inside. Journal of Energy Storage. 2023;72:108745. https://doi.org/10.1016/j.est.2023.108745

32. Nayak MK, Dogonchi AS, Rahbari A. Free convection of Al2O3-water nanofluid inside a hexagonal-shaped enclosure with cold diamond-shaped obstacles and periodic magnetic field. Case Studies in Thermal Engineering. 2023;50:103429. https://doi.org/10.1016/j.csite.2023.103429

33. Sharma BK, Kumawat C, Makinde OD. Hemodynamical analysis of MHD two phase blood flow through a curved permeable artery having variable viscosity with heat and mass transfer. Biomechanics and Modeling in Mechanobiology. 2022;21(3):797-825. https://doi.org/10.1007/s10237-022-01561-w

34. Sharma BK, Kumawat C, Khanduri U, Mekheimer KS. Numerical investigation of the entropy generation analysis for radiative mhd power-law fluid flow of blood through a curved artery with hall effect. Waves in Random and Complex Media. 2023:1-38. https://doi.org/10.1080/17455030.2023.2226228

35. Kumawat C, Sharma BK, Al-Mdallal QM, Rahimi-Gorji M. Entropy generation for MHD two phase blood flow through a curved permeable artery having variable viscosity with heat and mass transfer. International Communications in Heat and Mass Transfer. 2022;133: 105954. https://doi.org/10.1016/j.icheatmasstransfer.2022.105954

36. Koo J, Kleinstreuer C. Laminar nanofluid flow in microheat-sinks. International journal of heat and mass transfer. 2005;48(13):2652-61. https://doi.org/10.1016/j.ijheatmasstransfer.2005.01.029

37. Santra AK, Sen S, Chakraborty N. Study of heat transfer due to laminar flow of copper–water nanofluid through two isothermally heated parallel plates. International journal of thermal sciences. 2009;48(2):391-400.https://doi.org/10.1016/j.ijthermalsci.2008.10.004

38. Yasmeen T, Hayat T, Khan MI, Imtiaz M, Alsaedi A. Ferrofluid flow by a stretched surface in the presence of magnetic dipole and homogeneous-heterogeneous reactions. Journal of Molecular liquids. 2016;223:1000-5. https://doi.org/10.1016/j.molliq.2016.09.028

39. Nawaz M, Nazir U, Saleem S, Alharbi SO. An enhancement of thermal performance of ethylene glycol by nano and hybrid nanoparticles. Physica A: Statistical Mechanics and its Applications. 2020;551:124527. https://doi.org/10.1016/j.physa.2020.124527

40. Sohankar A, Norberg C, Davidson L. Low-Reynolds-number flow around a square cylinder at incidence: study of blockage, onset of vortex shedding and outlet boundary condition. International journal for numerical methods in fluids. 1998;26(1):39-56. https://doi.org/10.1002/(SICI)1097-0363

41. Abbassi H, Turki S, Nasrallah SB. Channel flow past bluff-body: outlet boundary condition, vortex shedding and effects of buoyancy. Computational Mechanics.2002;28(1):10-6. https://doi.org/10.1007/s004660100261

42. Masoumi N, Sohrabi N, Behzadmehr A. A new model for calculating the effective viscosity of nanofluids. Journal of Physics D: Applied Physics. 2009;42(5):055501. DOI 10.1088/0022-3727/42/5/055501

43. Xuan Y, Roetzel W. Conceptions for heat transfer correlation of nanofluids. International Journal of heat and Mass transfer. 2000;43(19):3701-7. https://doi.org/10.1016/S0017-9310(99)00369-5

44. Vajjha RS, Das DK. Experimental determination of thermal conductivity of three nanofluids and development of new correlations. International journal of heat and mass transfer. 2009;52(21-22):4675-82. https://doi.org/10.1016/j.ijheatmasstransfer.2009.06.027

45. Dogonchi AS, Waqas M, Afshar SR, Seyyedi SM, Hashemi-Tilehnoee M, Chamkha AJ, Ganji DD. Investigation of magnetohydrodynamic fluid squeezed between two parallel disks by considering Joule heating, thermal radiation, and adding different nanoparticles. International Journal of Numerical Methods for Heat & Fluid Flow. 2020;30(2):659-80. https://doi.org/10.1108/HFF-05-2019-0390

46. Abbas N, Nadeem S, Issakhov A. Transportation of modified nanofluid flow with time dependent viscosity over a Riga plate: exponentially stretching. Ain Shams Engineering Journal. 2021;12(4):3967-73. https://doi.org/10.1016/j.asej.2021.01.034

47. Sivaraj R, Animasaun IL, Olabiyi AS, Saleem S, Sandeep N. Gyrotactic microorganisms and thermoelectric effects on the dynamics of 29 nm CuO-water nanofluid over an upper horizontal surface of paraboloid of revolution. Multidiscipline Modeling in Materials and Structures. 2018 Oct 8;14(4):695-721. https://doi.org/10.1108/MMMS-10-2017-0116

48. Owen MS. ASHRAE Handbook: Fundamentals, American Society of Heating. Refrigeration and Air-Conditioning Engineers. 2009.

49. Scarpa F, Smith FC. Passive and MR fluid-coated auxetic PU foam–mechanical, acoustic, and electromagnetic properties. Journal of intelligent material systems and structures. 2004;15(12):973-9. https://doi.org/10.1177/1045389X04046610

50. ANSYS C. Reference Guide. Release 12.1. ANSYS. Inc. 2009.

51. Uddin MJ, Rasel SK, Rahman MM, Vajravelu K. Natural convective heat transfer in a nanofluid-filled square vessel having a wavy upper surface in the presence of a magnetic field. Thermal Science and Engineering Progress. 2020;19:100660. https://doi.org/10.1016/j.tsep.2020.100660

52. Abdi H, Motlagh SY, Soltanipour H. Study of magnetic nanofluid flow in a square cavity under the magnetic field of a wire carrying the electric current in turbulence regime. Results in Physics. 2020;18:103224. https://doi.org/10.1016/j.rinp.2020.103224

53. Tzirtzilakis EE, Xenos MA. Biomagnetic fluid flow in a driven cavity. Meccanica. 2013;48:187-200. https://doi.org/10.1007/s11012-012-9593-7

54. Lee S, Choi SS, Li SA, Eastman JA. Measuring thermal conductivity of fluids containing oxide nanoparticles.1999;121(2): 280-289. https://doi.org/10.1115/1.2825978

55. Philip J, Shima PD, Raj B. Evidence for enhanced thermal conduction through percolating structures in nanofluids. Nanotechnology. 2008;19(30):305706. DOI 10.1088/0957-4484/19/30/305706

56. Shima PD, Philip J, Raj B. Influence of aggregation on thermal conductivity in stable and unstable nanofluids. Applied Physics Letters. 2010;97(15). https://doi.org/10.1063/1.3497280

57. Xuan Y, Li Q. Heat transfer enhancement of nanofluids. International Journal of heat and fluid flow. 2000;21(1):58-64. https://doi.org/10.1016/S0142-727X(99)00067-3

58. Das SK, Putra N, Thiesen P, Roetzel W. Temperature dependence of thermal conductivity enhancement for nanofluids. J. Heat Transfer. 2003;125(4):567-74. https://doi.org/10.1115/1.1571080

59. Liu MS, Lin MC, Huang IT, Wang CC. Enhancement of thermal conductivity with CuO for nanofluids. Chemical Engineering & Technology: Industrial Chemistry-Plant Equipment-Process Engineering-Biotechnology. 2006;29(1):72-7. https://doi.org/10.1002/ceat.200500184

60. Martínez-Cuenca R, Mondragón R, Hernández L, Segarra C, Jarque JC, Hibiki T, Juliá JE. Forced-convective heat-transfer coefficient and pressure drop of water-based nanofluids in a horizontal pipe. Applied Thermal Engineering. 2016;98:841-9. https://doi.org/10.1016/j.applthermaleng.2015.11.050

61. Buschmann MH. Nanofluid heat transfer in laminar pipe flow with twisted tape. Heat Transfer Engineering. 2017;38(2):162-76. https://doi.org/10.1080/01457632.2016.1177381

Madhu Sharma: https://orcid.org/0000-0002-9154-6708

Bhupendra K. Sharma: https://orcid.org/0000-0002-2051-9681

Chandan Kumawat: https://orcid.org/0000-0002-3366-2807

Arun K. Jalan: https://orcid.org/0000-0001-5382-1521

Neyara Radwan: https://orcid.org/0000-0001-8756-4834

ANALYTICAL ANALYSIS OF THERMAL CONDUCTIVITY BY LAMINAR FLOW UNDER THE INFLUENCE OF MHD USING VOGEL'S MODEL

Rashid NAWAZ*/** , **Muhammad FAROOQ***
Siddiq Ur REHMAN* , **Gamal M. ISMAIL***** , **Hijaz AHMAD***/****/*****/********

*Department of Mathematics, Abdul Wali Khan University, Mardan, Khyber Pakhtunkhwa, Pakistan
**UniSa STEM, University of South Australia, Australia
***Department of Mathematics, Faculty of Science, Islamic University of Madinah, Madinah, 42351, Saudi Arabia
****Near East University, Operational Research Center in Healthcare, TRNC Mersin 10, Nicosia, 99138, Turkey
*****Department of Computer Science and Mathematics, Lebanese American University, Beirut, Lebanon
******Department of Mathematics and Informatics, Azerbaijan University, Jeyhun Hajibeyli street, 71, AZ1007, Baku, Azerbaijan

rashidnawaz@awkum.edu.pk, muhammadfarooq@awkum.edu.pk,
siddiqurrehman@awkum.edu.pk, gamalm2010@yahoo.com, hijaz555@gmail.com

Abstract: The main aim of this work is to study the influence of thermal conductivity of uniform couple stress fluid of inclined Poiseuille flow in the presence of magneto hydrodynamic (MHD) between two parallel plates. A well-known Vogel's viscosity model is used. The momentum and energy equations are solved analytically by utilizing Homotopy Perturbation method (HPM) and Optimal Homotopy Asymptotic Method (OHAM). The results include the velocity profile, average velocity, volume flux, Shear Stress, Skin friction and the temperature distribution between the plates. Particular attention is given to the effect of MHD Γ on the velocity field and temperature distribution. As can be seen, that there exists direct relation between MHD parameter Γ and velocity profile and bears inverse relation with temperature distribution. In addition to that, influence of non-dimensional parameters like G, A, a, B and γ on the velocity field and temperature distribution are also discussed graphically. The physical characteristics of the problem have been well discussed in graphs for several parameters of interest. The results reveal that both techniques are reliable and are in great agreement with each other.

Keywords: OHAM, HPM, Couple Stress Fluid, Magnetohydrodynamics, Vogel's Model

1. INTRODUCTION

In recent years, the flow of non-Newtonian fluids has attracted a great deal of interest due to the expanding industrial, medical, and technological applications. Numerous researchers have investigated certain flow issues with non-Newtonian fluids, such as couple stress fluid. Stokes's [1] theory of couple stress fluids has several medical, industrial, and scientific applications, including as the modelling of synthetic fluids, polymer thickened oils, liquid crystals, animal blood, and synovial fluid. The book by Stokes [2] covers some of the earliest developments in coupled stress fluid theory and basic flows. A few scholars have recently investigated some couple stress fluid flows for various flow geometries [3]-[8].

There are a number of suggested constitutive equations for non-Newtonian fluids, each of which attempts to provide an explanation for their behavior. Couple stresses, a non-symmetric stress tensor, and body couples are only a few of the unique characteristics of Stokes' couple stress fluid model. Models for fluids with mechanically important microstructures are proposed in Stokes' couple stress fluid theory. Microstructure can affect a liquid if the distinctive geometric dimension of the problem is on the same scale as the microstructure's size [9]. The introduction of a size-dependent effect is one of the major aspects of couple stresses. Classical continuum mechanics disregards the influence

of material particle size inside the continuum. It is compatible with the symmetry of the force-stress tensor to disregard the rotational interaction between fluid particles. Clearly, this is not the case, and a size-dependent couple-stress theory is necessary in situations such as fluid flow with suspended particles when it must be used. The couple-stress tension caused by the microrotation of these freely floating particles led to the formation of couple-stress fluid. As shown by the couplet stress fluid, fluids may be used to describe a variety of lubricants, suspension fluids, blood, etc. These fluids are employed in a range of industrial processes, including liquid crystal solidification, polymer fluid extrusion, colloidal solutions, and the cooling of metallic plates in a bath.

A number of technological processes rely on heat transmission. Non-Newtonian mixtures require heat transfer for processing and management [10]. Mathematicians, engineers, and scientists find it hard to understand how nonlinear fluid flows work because the nonlinearity can show up in different ways. For example, studying reactive variable viscosity flows in a slit with wall injection or suction is a good example. In our case, the fact that viscosity changes with temperature is one of the reasons why the coupled ordinary differential equations are not linear. Different researchers look into flows with viscosity that changes with temperature [11]-[14]. Moreover, several scientists examined the effect of MHD on squeezing fluid flow through porous medium in [15]-[20]. Furthermore, the effect of MHD on micropolar fluid were analyzed by researchers in [21]-[24].

Many numerical, semi-numerical, analytical, and homotopy based strategies have been presented by mathematicians to solve nonlinear differential equations and related systems. The Optimal Homotopy Asymptotic Method (OHAM), which is used to solve differential equations, is one of the efficient homotopy based techniques. This technique was presented by Marinca et al. to solve differential equations. This perturbation approach works without small or big parameters, unlike other perturbation methods, and it does not require discretization, which requires time. Unlike iterative approaches, the procedure does not require a preliminary guess. The method's convergence is also changed by using a more variable function called the auxiliary function. One of OHAM's shortcomings is that the unknown auxiliary parameters which control the convergence are determined using the least squares approach, which requires extra time when dealing with highly nonlinear problem. The collocation approach is an alternative technique; however, accuracy would be compromised in the latter situation. Numerous issues have been resolved using this approach. When dealing with heat flux conditions, researchers converted highly nonlinear partial differential equations (PDEs) [25] [26] into structured ordinary differential equations (ODEs) with appropriate restrictions (OHAM). Recently, scientists and engineers have increased their usage of the Homotopy perturbation technique (HPM) [27] in nonlinear problems due to the fact that this method converts the tough problem under inquiry into a clear, easily-solvable problem. He invented and modified the homotopy perturbation method. after initially proposing [28] it in 1998. In the majority of instances, the approach results in a fairly quick convergence of the solution series. After only one iteration, the outcome is often quite accurate. By combining linear and nonlinear problems, He's homotopy perturbation approach was created to solve beginning and boundary value problems. Most perturbation methods assume the presence of a tiny parameter, yet the vast majority of nonlinear problems lack any type of small parameter. Scientists have just begun to use homotopy perturbation theory, which, when properly combined with perturbation theory, has [29] [30] the potential to be a very effective mathematical tool. El-Shahed [31] recently used Volterra's integro differential equations to apply He's homotopy perturbation approach.

Keeping in mind the above literature survey, there is no analysis about the steady inclined Poiseuille flow under the influence of magneto hydrodynamic (MHD) between two parallel plates. Prior to computing the analytical results, the regulating flow expression is converted to an ordinary system. The two reliable analytical techniques OHAM and HPM are used to solve the governing system of couple equations. Velocity profile and temperature distribution are calculated numerically and graphically presented. The effect of various parameters i.e. G, A,Γ, γ and a are also discussed.

2. CHAPTER TITLE

Consider a fluid that is viscous, incompressible, and electrically conducts as it moves between two infinitely parallel inclined plates at the positions of y = −H (lower plate) and y = H (upper plate) under the influence of a persistent pressure gradient in the direction of motion and a steady transverse magnetic field Γ. The fluid still flows even while the plates are still. Temperatures of Θ_0 and Θ_1 are maintained for the bottom and top plates, respectively.
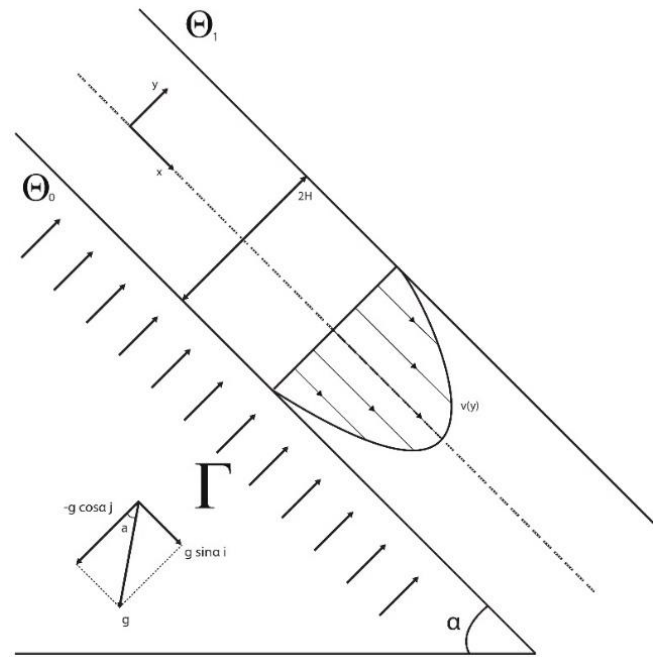
Fig 1. Geometry of the problem

Fig. (1) depicts the coordinate system that was chosen. The angle of the plate with regard to the horizontal direction is α. The viscosity η of a fluid is stated to be dependent on Θ. Velocity and temperature field are of the form:

$$Y = Y(v,0,0), v = v(y), \Theta = \Theta(y), \tag{1}$$

The continuity equation [32]

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho Y) = 0, \tag{2}$$

is identically satisfied after applying assumptions (1). The three components of the momentum equation [32]

$$\rho \frac{DY}{Dt} = \nabla \cdot T - \eta \nabla^4 Y + J \times B + \rho g, \tag{3}$$

become

$$0 = -\frac{\partial p}{\partial x} + \frac{\partial}{\partial x}(S_{xx}) + \frac{\partial}{\partial y}(S_{xy}) + \frac{\partial}{\partial z}(S_{xz}) - \eta \frac{d^4 v}{dy^4} - \sigma B_0 v + \rho g sin(\alpha), \tag{4}$$

$$0 = -\frac{\partial p}{\partial y} + \frac{\partial}{\partial x}(S_{yx}) + \frac{\partial}{\partial y}(S_{yy}) + \frac{\partial}{\partial z}(S_{yz}) - \rho g cos(\alpha), \tag{5}$$

$$0 = -\frac{\partial p}{\partial z} + \frac{\partial}{\partial x}(S_{zx}) + \frac{\partial}{\partial y}(S_{zy}) + \frac{\partial}{\partial z}(S_{zz}), \tag{6}$$

where B_0 denotes the applied magnetic field and σ denotes the fluid's electric conductivity, we know that

$$S_{xx} = S_{yy} = S_{zz} = S_{xz} = S_{zx} = S_{yz} = S_{zy} = 0, S_{xy} = \mu \frac{dv}{dy} = S_{yx}, \tag{7}$$

the following equations are derived by inserting these values into (4-6)

$$0 = -\frac{\partial p}{\partial x} + \frac{d}{dy}\left(\mu \frac{dv}{dy}\right) - \eta (d^4 v)/(dy^4) - \sigma \beta_0 v + \rho g sin(\alpha), \tag{8}$$

$$0 = -\frac{\partial p}{\partial y} - \rho g cos(\alpha), \tag{9}$$

Rashid Nawaz, Muhammad Farooq, Siddiq Ur Rehman, Gamal M. Ismail, Hijaz Ahmad
Analytical Analysis of Thermal Conductivity by Laminar Flow under the Influence of MHD using Vogel's Model

DOI 10.2478/ama-2024-0058

$$0 = -\frac{\partial p}{\partial z}, \tag{10}$$

The velocity profile is obtained from (4). This equation may also be expressed as

$$\eta \frac{d^4 v}{dy^4} - \mu \frac{d^2 v}{dy^2} - \frac{d\mu}{dy}\frac{dv}{dy} + \frac{\partial p}{\partial x} + \sigma\beta_0 v - \rho g \sin(\alpha) = 0. \tag{11}$$

After applying all presumptions, the energy equation [32]

$$\rho c_p \frac{D\Theta}{Dt} = \kappa\nabla^2 \Theta + tr(T \cdot L), \tag{12}$$

simplifies to the Eq. (13)

$$\frac{d^2 \Theta}{dy^2} + \frac{\mu}{\kappa}\left(\frac{dv}{dy}\right)^2 + \frac{\mu}{\kappa}\left(\frac{d^2 v}{dy^2}\right)^2 = 0, \tag{13}$$

The corresponding boundary conditions are

$$v(y = \pm H) = 0, \tag{14}$$

$$\frac{d^2}{dy^2} v(y = \pm H) = 0, \tag{15}$$

$$\Theta(y = -H) = \Theta_0, \Theta(y = H) = \Theta_1. \tag{16}$$

The standard no-slip boundary conditions are given in Eq. (14). According to Eq. (15), couple stresses at the plates are zero. The following non-dimensional parameters are introduced [32]:

$$v^{\cdot} = \frac{v}{V}, y^{\cdot} = \frac{y}{H}, x^{\cdot} = \frac{x}{H}, \Theta^{\cdot} = \frac{\Theta - \Theta_0}{\Theta_1 - \Theta_0}, \mu^{\cdot} = \frac{\mu}{\mu_0}, p^{\cdot} = \frac{p}{\mu_0 \frac{V}{H}},$$

$$\gamma = \frac{\mu_0 V^2}{\kappa(\Theta_1 - \Theta_0)}, B^2 = \frac{\mu_0(H^2)}{\eta}, G = -\frac{B^2 H^5}{\mu_0 V}\frac{\partial p}{\partial x}, A = \frac{\rho g H^4}{\eta V}.$$

Hence Γ in terms of parameters are as follows $\Gamma = \frac{\sigma B_0^2 H^4}{\eta}$. Where γ is the Brinkman number, μ Is the reference viscosity, and V is the reference velocity. Using these dimensionless parameters, equation (11) and (13) along with the boundary conditions becomes (dropping dots)

$$\frac{d^4 v}{dy^4} - B^2 \mu \frac{d^2 v}{dy^2} - B^2 M \frac{d\mu}{dy}\frac{dv}{dy} + \Gamma v + A\sin(\alpha) - G = 0, v(y = \pm 1) = \frac{d^2}{dy^2} v(y = \pm 1) = 0, \tag{17}$$

$$\frac{d^2 \Theta}{dy^2} + \gamma\mu\left(\frac{dv}{dy}\right)^2 + \frac{\gamma}{B^2}\left(\frac{d^2 v}{dy^2}\right)^2 = 0, \Theta(y = -1) = 0, \Theta(y = 1) = 1. \tag{18}$$

The following is an expression for the dimensionless form of the Vogel's viscosity model [33]:

$$\mu = \mu_* \exp\left(\frac{A_0}{B_0 + \Theta} - \Theta_w\right). \tag{19}$$

Using Taylor series expansion on Eq. (19) we get

$$\mu = a^2\left(1 - \frac{A_0 \Theta}{B_0^2}\right). \tag{20}$$

Assume that $M = A_0/(B_0^2)$, then (20) becomes

$$\mu = a^2(1 - M\Theta), \frac{d\mu}{dy} = -a^2 M \frac{d\Theta}{dy}. \tag{21}$$

The coupled system shown below is generated by substituting Eq. (20) in the governing Eqs. (17) and (18):

$$\frac{d^4 v}{dy^4} - B^2 a^2 (1 - M\Theta)\frac{d^2 v}{dy^2} + B^2 Ma^2 \frac{d\Theta}{dy}\frac{dv}{dy} + \Gamma v + A\sin(\alpha) - G = 0, v(y = \pm 1) = \frac{d^2}{dy^2} v(y = \pm 1) = 0, \tag{22}$$

$$\frac{d^2 \Theta}{dy^2} + \gamma a^2 (1 - M\Theta)\left(\frac{dv}{dy}\right)^2 + \frac{\gamma}{B^2}\left(\frac{d^2 v}{dy^2}\right)^2 = 0, \Theta(y = -1) = 0, \Theta(y = 1) = 1. \tag{23}$$

3. BASIC NOTION OF METHODS

3.1. Fundamental concept of OHAM

Let us examine a boundary value problem, as well as the boundary conditions

$$\chi[X(y)] + g(y) + \xi[X(y)] = 0, \tag{24}$$

$$\beta\left(X, \frac{\partial X}{\partial y}\right) = 0. \tag{25}$$

In this notation, y is the dependent variable, $X(y)$ is the unknown function, $g(y)$ is the known function, χ, ξ and β are the linear, nonlinear, and boundary operators, respectively.

The OHAM states that the homotopy $\hbar(G(y, \rho), \rho): R \times [0,1] \to R$ satisfies

$$(1 - \rho)[\chi(G(y, \rho)) + g(y)] = \hbar(\rho)[\chi(G(y, \rho)) + g(y) + \xi(G(y, \rho))], \tag{26}$$

$$\beta\left(G(y, \rho), \frac{\partial G(y, \rho)}{\partial y}\right) = 0,$$

for $y \in R$ and $\rho \in [0,1]$, is generated using the non-zero auxiliary function $\hbar(\rho)$ for $\rho \neq 0$ with $\hbar(0) = 0$ and a function $G(y, \rho)$ whose value is unknown. $G(y, 0) = X_0(y)$ and $G(y, 1) = X(y)$ remain true when $\rho = 0$ and 1. Hence, as ρ varies from 0 to 1, the solution $G(y, \rho)$ changes from $X_0(y)$ to $X(y)$.
for $\rho = 0$,

$$\chi[X_0(y)] + g(y) = 0, \beta\left(X_0(y), \frac{dX_0(y)}{dy}\right) = 0. \tag{27}$$

We pick the auxiliary function $\hbar(\rho)$ in such a manner that

$$\hbar(\rho) = \sum_{i=0}^{n} \rho^i C_i, \tag{28}$$

where C_i are the constants governing convergence that must be calculated. Expanding Taylor's series $G(y, \rho)$ about ρ to obtain an approximation

$$G(y, \rho, C_i) = X_0(y) + \sum_{j=1}^{n} X_j(y, C_1, C_2, \dots, C_j)\rho^j. \tag{29}$$

Equating the coefficients of identical powers of ρ by putting (29) into (26) provides the following findings.
for $\rho = 1$,

$$\chi[X_1(y)] + g(y) = C_1 \xi_0[X(y)], \beta\left(X_1(y), \frac{dX_1(y)}{dy}\right) = 0. \tag{30}$$

for $\rho = 2$,

$$\chi[X_2(y)] - \chi[X_1(y)] = C_2 \xi_0[X_0(y)] + C_1[\chi(X_1(y)) + \xi_1(X_0(y), X_1(y))], \beta\left(X_2(y), \frac{(dX_2(y))}{dy}\right) = 0. \tag{31}$$

generally

$$\chi[X_n(y)] - \chi[X_{n-1}(y)] =$$
$$C_n\,\xi_0\,[X_0\,(y)] +$$
$$\left(\sum_{j=1}^{n-1}\left[\chi[X_{n-j}(y)] + \xi_{n-j}(X_0(y), X_1(y), \cdots, X_{n-1}(y))\right]\right),$$
(32)

with boundary conditions

$$\beta\left(X_k\,(y), \frac{dX_k\,(y)}{dy}\right) = 0, k = 2,3,4\ldots$$
(33)

Where $\xi_m\left(X_0\,(y), X_1\,(y), \cdots, X_{m-1}\,(y)\right)$ is the coefficient of ρ^m in the expansion of $\xi\big(G\,(y, \rho)\big)$ about ρ as

$$\xi\big(G(y, \rho, C_i\,)\big) =$$
$$\xi_0\,[X_0\,(y)] +$$
$$\sum_{m=1}^{\infty}\xi_m(X_0(y), X_1(y), \cdots, X_{m-1}(y))\,\rho^m$$
(34)

The convergence of the series (29) is determined by C_i. For convergence at $\rho = 1$, the rth-order approximation G is

$$G(y, C_1, C_2, \cdots, C_r\,) = X_0\,(y) + \sum_{j=1}^{r}X_j(y, C_1, C_2, \cdots, C_j).$$
(35)

Placing (35) in (24) the expression for the residual is

$$R(z, C_1, C_2, \cdots, C_r\,) = \chi[G(y, C_1, C_2, \cdots, C_r\,)] + g(y) + \xi[G(y, C_1, C_2, \cdots, C_r\,)].$$
(36)

If $R = 0, G$ will be the exact solution, however this is typically not the case in nonlinear situations. There are numerous approaches for determining the optimal values of the constants C_i. Following is the application of the least-squares approach.

$$J = \int_a^b R^2(z, C_1, C_2, \cdots, C_r\,)dz.$$
(37)

Minimizing this function, we have

$$\frac{\partial J}{\partial c_i}\,(y, C_1, C_2, \cdots, C_r\,) = 0, i = 1,2,3,\cdots, r.$$
(38)

Where a and b are chosen from the problem's domain in order to locate C_i. For these values of C_i, the approximation solution is well-defined.

3.2. Fundamental concept of HPM

Take the following nonlinear differential equation to illustrate the operation of HPM:

$$\Xi[X(y)] + g(r) = 0, r \in \Omega,$$
(39)

with the boundary conditions

$$\beta\left(X(y), \frac{\partial X(y)}{\partial y}\right) = 0, r \in \Lambda.$$
(40)

Ξ is a basic differential operator, β is a boundary operator, $g(r)$ is a well-known analytic function, and Λ is the domain boundary for Ω.

The operator Ξ is decomposed into two pieces, χ and ξ, where χ is linear and ξ is nonlinear. Hence, Eq. (39) may be expressed as follows:

$$\chi[X(y)] + \xi[X(y)] - g(r) = 0,$$
(41)

He [9] built a homotopy $\chi: \Omega \times [0,1] \to R$ that meets the condition,

$$H(X, \rho) = (1 - \rho)[\chi(X) - \chi(X\_0\,)] + \rho[\Xi(X) - g(r)]$$
(42)

or

$$H(X, \rho) = \chi(X) - \chi(X_0\,) + \rho[\chi(X_0\,)] + \rho[\xi(X) - g(r)].$$
(43)

Where $r \in \Omega, \rho \in [0,1]$ which is referred to as the homotopy parameter, and X_0 is the initial approximation of the function (39). Hence, it concludes that

$$H(X, 0) = \chi(X) - \chi(X_0) = 0, H(X, 1) = \Xi(X) - g(r) = 0,$$
(44)

and the procedure of moving ρ from 0 to 1 is identical to that of $H(X, \rho)$ from $\chi(X) - \chi(X_0\,)$ to $\Xi(X) - g(r)$. This is known as deformation in topology, $\chi(X) - \chi(X_0\,)$ and $\Xi(X) - g(r)$ are called homotopic. Using the perturbation approach [10], and assuming that $0 \le \rho \le 1$ is a small parameter, we may suppose that the solution of (42) or (43) can be written as a series in ρ, as shown below

$$X = X_0 + \rho X_1 + \rho^2 X_2 + \rho^3 X_3 + \cdots$$
(45)

when ρ → 1, (42) or (43) corresponds to (41) and becomes the approximate solution of (41), i.e.

$$X(y) = \lim_{\rho \to 1} X = X_0 + X_1 + X_2 + X_3 + \ldots$$
(46)

The convergence rate of the series (46) is dependent on $\Xi(X)$ [11] in the majority of situations.

4. SOLUTIONS OF THE PROBLEM

4.1. OHAM solution

Zero component of velocity and temperature distribution along with the boundary conditions are given by

$$\frac{d^4}{dy^4}\,v_0\,(y) + A\sin(\alpha) - G = 0,$$
(47)

$$v_0\,(y = \pm 1) = 0, \frac{d^2}{dy^2}\,v_0\,(y = \pm 1) = 0,$$

$$\frac{d^2}{dy^2}\,\Theta_0\,(y) = 0,$$
(48)

$$\Theta_0\,(y = -1) = 0, \Theta_0\,(y = 1) = 1.$$

Their solutions are

$$v_0(y) = \frac{1}{24}\,(5G - 6Gy^2 + Gy^4 - 5A\sin(\alpha) + 6Ay^2\sin(\alpha) - Ay^4\sin(\alpha)),$$
(49)

$$\Theta_0(y) = \frac{1+y}{2}.$$
(50)

First component of velocity and temperature distribution along with the boundary conditions are given by

$$-a^2\,B^2\,c_1\,M\frac{d}{dy}\,v_0\,(y)\frac{d}{dy}\,\Theta_0\,(y) -$$
$$a^2\,B^2\,c_1\,M\Theta_0\,(y)\frac{d^2}{dy^2}\,v_0\,(y) + a^2\,B^2\,c_1\,\frac{d^2}{dy^2}\,v_0\,(y) -$$
$$A\sin(\alpha) - Ac_1\sin(\alpha) + c_1\,G - c_1\,\Gamma v_0\,(y) -$$
$$c_1\,\frac{d^4}{dy^4}\,v_0\,(y) + G - \frac{d^4}{dy^4}\,v_0\,(y) + \frac{d^4}{dy^4}\,v_1\,(y) = 0,$$
(51)

$$v_1\,(y = \pm 1) = 0, \frac{d^2}{dy^2}\,v_1\,(y = \pm 1) = 0.$$

Rashid Nawaz, Muhammad Farooq, Siddiq Ur Rehman, Gamal M. Ismail, Hijaz Ahmad
Analytical Analysis of Thermal Conductivity by Laminar Flow under the Influence of MHD using Vogel's Model

DOI 10.2478/ama-2024-0058

$$-a^2 \gamma c_4 \left(\frac{d}{dy} v_0(y)\right)^2 + a^2 M \gamma c_4 \Theta_0(y) \left(\frac{d}{dy} v_0(y)\right)^2 -$$

$$\frac{1}{B^2} \gamma c_4 \left(\frac{d^2}{dy^2} v_0(y)\right)^2 - \frac{d^2}{dy^2} \Theta_0(y) - c_4 \frac{d^2}{dy^2} \Theta_0(y) +$$

$$\frac{d^2}{dy^2} \Theta_1(y) = 0, \tag{52}$$

$$\Theta_1(y=-1) = 0, \Theta_1(y=1) = 0.$$

Their solutions are

$$v_1(y) =$$
$$\frac{1}{40320} (c_1 (y^2-1)(G -$$
$$A\sin(\alpha))(4a^2 B^2 (M(y(y(y(y(4y+7)-38)-98)+$$
$$74)+427)-14(y^4-14y^2+61)) + \Gamma(y^6-27y^4+$$
$$323y^2-1385))), \tag{53}$$

$$\Theta_1(y) = -\frac{1}{181440B^2} \left(\gamma c_4 (y^2-1)\left(a^2 B^2 (35My^7 +\right.\right.$$
$$45(M-2)y^6 - 325My^5 - 459(M-2)y^4 + 809My^3 +$$
$$1431(M-2)y^2 + 809My + 1431(M-2)) -$$
$$\left.\left. 1512(y^4-4y^2+11)\right)(G-A\sin(\alpha))^2 \right), \tag{54}$$

Second component of velocity and temperature distribution along with the boundary conditions are given by

$$Gc_2 - A\sin(\alpha)c_2 - \Gamma c_2 v_0(y) - \Gamma c_1 v_1(y) -$$
$$a^2 B^2 M c_2 \frac{d}{dy} v_0(y) \frac{d}{dy} \Theta_0(y) -$$
$$a^2 B^2 M c_1 \frac{d}{dy} v_1(y) \frac{d}{dy} \Theta_0(y) -$$
$$a^2 B^2 M c_1 \frac{d}{dy} v_0(y) \frac{d}{dy} \Theta_1(y) + a^2 B^2 c_2 \frac{d^2}{dy^2} v_0(y) -$$
$$a^2 B^2 c_2 \Theta_0(y) \frac{d^2}{dy^2} v_0(y) -$$
$$a^2 B^2 M c_1 \Theta_1(y) \frac{d^2}{dy^2} v_0(y) + a^2 B^2 c_1 \frac{d^2}{dy^2} v_1(y) -$$
$$a^2 B^2 M c_1 \Theta_0(y) \frac{d^2}{dy^2} v_1(y) - c_2 \frac{d^4}{dy^4} v_0(y) -$$
$$\frac{d^4}{dy^4} v_1(y) - c_1 \frac{d^4}{dy^4} v_1(y) + \frac{d^4}{dy^4} v_2(y) = 0, \tag{55}$$

$$v_2(y=\pm 1) = 0, \frac{d^2}{dy^2} v_2(y=\pm 1) = 0,$$

$$-a^2 \gamma c_5 \left(\frac{d}{dy} v_0(y)\right)^2 + a^2 M \gamma c_5 \Theta_0(y) \left(\frac{d}{dy} v_0(y)\right)^2 +$$

$$a^2 M \gamma c_4 \Theta_1(y) \left(\frac{d}{dy} v_0(y)\right)^2 -$$

$$2a^2 c_4 \gamma \frac{d}{dy} v_0(y) \frac{d}{dy} v_1(y) +$$

$$2a^2 M \gamma c_4 \Theta_0(y) \frac{d}{dy} v_0(y) \frac{d}{dy} v_1(y) -$$

$$\frac{1}{B^2} \gamma c_5 \left(\frac{d^2}{dy^2} v_0(y)\right)^2 -$$

$$\frac{1}{B^2} 2\gamma c_4 \frac{d^2}{dy^2} v_0(y) \frac{d^2}{dy^2} v_1(y) - c_5 \frac{d^2}{dy^2} \Theta_0(y) -$$

$$\frac{d^2}{dy^2} \Theta_1(y) - c_4 \frac{d^2}{dy^2} \Theta_1(y) + \frac{d^2}{dy^2} \Theta_1(y) = 0, \tag{56}$$

$$\Theta_2(y=-1) = 0, \Theta_2(y=1) = 0.$$

Their solutions are

$$v_2(y) =$$
$$-\frac{1}{261534873600} (y^2-1)(G -$$
$$A\sin(\alpha))(c_1 (2a^2 \gamma c_4 M (a^2 B^2 (1540My^{13} +$$
$$2475(M-2)y^{12} - 31010My^{11} - 55674(M-2)y^{10} +$$
$$237622My^9 + 511893(M-2)y^8 - 573188My^7 -$$
$$1515132(M-2)y^6 - 1035936My^5 - 2947563(M-$$
$$2)y^4 + 3822918My^3 + 18538902(M-2)y^2 -$$
$$3536826My - 54064629(M-2)) - 6552(21y^{10} -$$
$$287y^8 + 2188y^6 - 10748y^4 + 27367y^2 -$$
$$64109))(G-A\sin(\alpha))^2 -$$
$$6486480(4a^2 B^2 (M(y(y(y(y(4y+7)-38)-98)+$$
$$74)+427) - 14(y^4-14y^2+61)) + \Gamma(y^6-27y^4+$$
$$323y^2-1385))) -$$
$$6486480c_2 (4a^2 B^2 (M(y(y(y(y(4y+7)-38)-98)+$$
$$74)+427) - 14(y^4-14y^2+61)) + \Gamma(y^6-27y^4+$$
$$323y^2-1385)) -$$
$$546c_1^2 (132a^4 B^4 (M^2 (y(y(y(y(y(y(y(7y+25)-$$
$$83)-515)-209)+3643)+7111)-6737) -$$
$$32714) + 2M(y(6737-y(y(5y(y(y(5y+9)-103)-$$
$$243)+3643)+14535)) + 62325 + 90(y^6-27y^4+$$
$$323y^2-1385)) + 12a^2 B^2 (3960(M(y(y(y(y(4y+$$
$$7)-38)-98)+74)+427) - 14(y^4-14y^2+61)) +$$
$$\Gamma(6My^9 + 11(M-2)y^8 - 214My^7 - 484(M-2)y^6 +$$
$$3614My^5 + 11066(M-2)y^4 - 27010My^3 -$$
$$129844(M-2)y^2 + 50804My + 555731(M-2))) +$$
$$\Gamma(11880(y^6-27y^4+323y^2-1385) + \Gamma(y^{10} -$$
$$65y^8 + 2410y^6 - 53954y^4 + 631621y^2 -$$
$$2702765)))), \tag{57}$$

$$\Theta_2(y) = -\frac{1}{160059342643200B^2} (y^2-1)\gamma(G -$$
$$A\sin(\alpha))^2 \left(-\left((A^2+2G^2)M \left(a^2 B^2 (1576575My^{15} +\right.\right.\right.$$
$$2297295(M-2)y^{14} - 31679505My^{13} -$$
$$49808385(M-2)y^{12} + 251771355My^{11} +$$
$$443901195(M-2)y^{10} - 866969541My^9 -$$
$$1717393941(M-2)y^8 + 732117949My^7 +$$
$$1691135469(M-2)y^6 + 2148126829My^5 +$$
$$5197726557(M-2)y^4 - 2312301143My^3 -$$
$$7951990023(M-2)y^2 - 2312301143My -$$
$$7951990023(M-2)) - 3392928(30y^{12} - 425y^{10} +$$
$$2851y^8 - 11384y^6 + 25198y^4 - 19847y^2 -$$
$$\left.19847)\right)\gamma a^2 \bigg) + AM \left(a^2 B^2 (1576575My^{15} +\right.$$
$$2297295(M-2)y^{14} - 31679505My^{13} -$$
$$49808385(M-2)y^{12} + 251771355My^{11} +$$
$$443901195(M-2)y^{10} - 866969541My^9 -$$
$$1717393941(M-2)y^8 + 732117949My^7 +$$
$$1691135469(M-2)y^6 + 2148126829My^5 +$$
$$5197726557(M-2)y^4 - 2312301143My^3 -$$
$$7951990023(M-2)y^2 - 2312301143My -$$
$$7951990023(M-2)) - 3392928(30y^{12} - 425y^{10} +$$
$$2851y^8 - 11384y^6 + 25198y^4 - 19847y^2 -$$
$$\left.19847)\right)\gamma(A\cos(2\alpha) + 4G\sin(\alpha))a^2 +$$
$$882161280 \left(a^2 B^2 (35My^7 + 45(M-2)y^6 -\right.$$
$$325My^5 - 459(M-2)y^4 + 809My^3 + 1431(M-$$
$$2)y^2 + 809My + 1431(M-2)) - 1512(y^4-4y^2+$$

11))$\Big)c_4^2 + 205632\Big(4290\Big(a^2 B^2\Big(35My^7 +$
$45(M-2)y^6 - 325My^5 - 459(M-2)y^4 + 809My^3 +$
$1431(M-2)y^2 + 809My + 1431(M-2)\Big) -$
$1512(y^4 - 4y^2 + 11)\Big) + \Big(39a^4\Big(70M^2 y^{10} +$
$210(M-2)My^9 - 14(M(61M + 44) - 44)y^8 -$
$3640(M-2)My^7 + (M(1621M + 12254) -$
$12254)y^6 + 22100(M-2)My^5 + (M(15129M -$
$89386) + 89386)y^4 - 47728(M-2)My^3 +$
$\big((257114 - 65391M)M - 257114\big)y^2 - 35518(M -$
$2)My + M(257114 - 65391M) - 257114\Big)B^4 +$
$a^2\Big(\big(165My^{11} + 195(M-2)y^{10} - 5451My^9 -$
$6669(M-2)y^8 + 79634My^7 + 102726(M-2)y^6 -$
$503806My^5 - 714090(M-2)y^4 + 1144841My^3 +$
$2033655(M-2)y^2 + 1144841My + 2033655(M -$
$2)\big)\Gamma - 1287\big(140My^7 + 225(M-2)y^6 - 1300My^5 -$
$2715(M-2)y^4 + 3236My^3 + 8835(M-2)y^2 -$
$3484My - 22665(M-2)\big)\Big)B^2 - 1716(7y^8 - 173y^6 +$
$1717y^4 - 5423y^2 + 13792)\Gamma\Big)c_1\Big)c_4 +$
$882161280\Big(a^2 B^2\Big(35My^7 + 45(M-2)y^6 -$
$325My^5 - 459(M-2)y^4 + 809My^3 + 1431(M -$
$2)y^2 + 809My + 1431(M-2)\Big) - 1512(y^4 - 4y^2 +$
11)$\Big)c_5\Big),$ (58)

The second order OHAM solution of velocity profile and temperature distribution are as under.

$$v_{OHAM}(y) = v_0(y) + v_1(y) + v_2(y) \quad (59)$$

and

$$\Theta_{OHAM}(y) = \Theta_0(y) + \Theta_1(y) + \Theta_2(y). \quad (60)$$

Using collocation method, we have found the values of c_1, c_2, c_4, c_5 for velocity profile as

c_1	c_2	c_4	c_5
-0.9721794673016152	-3.9622629501480556×10^{-4}	1.3773239590060564	-0.6750848800383696

After using the values of c_i and $\gamma = 0.1, B = 0.1, a = 0.2, A = 0.4, \alpha = 165, M = 0.5, G = 1, \Gamma = 0.3$ in Eq. (59) we get

$v_{OHAM}(y) =$
$-1.36845 \times 10^{-17}(-1 + y^2)(8.71528 \times 10^{15} +$
$y(6.12648 \times 10^9 + y(-1.72871 \times 10^{15} +$
$y(-3.14025 \times 10^9 + y(-8.41587 \times 10^{12} +$
$y(3.2572 \times 10^8 + y(3.07209 \times 10^{11} + y(1.33903 \times$
$10^6 + y(5.70556 \times 10^8 + y(-37273.9 +$
$y(-1.22634 \times 10^7 + y(-20.1364 + y(1.33903 \times$
$10^6 + y(5.70556 \times 10^8 + y(-37273.9 +$
$y(-1.22634 \times 10^7 + y(-20.1364 + y(-62.6786 +$
$1y))))))))))))))).$ (61)

Using the same method we have found the values of c_1, c_2, c_4, c_5 for temperature distribution as

c_1	c_2	c_4	c_5
0.8022157425325956	-2.6532547581107915	0.45512453948245185	-2.1169027772515703

Using the values of c_i and $\gamma = 0.1, B = 1, A = 0.5, \alpha = 165, M = 0.5, G = 1, \Gamma = 0.1$ in Eq. (60) we get
$\Theta_{OHAM}(y) = 0.664312 + y(0.5 + y(-0.223863 +$
$y(-1.50995 \times 10^{-7} + y(0.0742752 + y(3.25779 \times$
$10^{-7} + y(-0.0146459 + y(-1.03106 \times 10^{-7} +$
$y(-0.0000814195 + y(9.87488 \times 10^{-9} +$
$y(3.14246 \times 10^{-6} + y(2.95234 \times 10^{-11} +$
$y(3.16566 \times 10^{-10} + y(-8.63231 \times 10^{-13} +$
$y(-1.82434 \times 10^{-11} + y(-1.19257 \times 10^{-16} +$
$(-3.21288 \times 10^{-16} + 5.65363 \times$
$10^{-18}y)y)))))))))))))))).$ (62)

4.2. HPM solution

Zero component of velocity and temperature distribution along with the boundary conditions are given by

$$\frac{d^4}{dy^4} v_0(y) + A \sin(\alpha) - G = 0, \quad (63)$$

$$v_0(y = \pm 1) = 0, \frac{d^2}{dy^2} v_0(y = \pm 1) = 0,$$

$$\frac{d^2}{dy^2} \Theta_0(y) = 0, \quad (64)$$

$$\Theta_0(y = -1) = 0, \Theta_0(y = 1) = 1.$$

Their solutions are

$$v_0(y) = \frac{1}{24}(5G - 6Gy^2 + Gy^4 - 5A\sin(\alpha) + 6Ay^2\sin(\alpha) - Ay^4\sin(\alpha)), \quad (65)$$

$$\Theta_0(y) = \frac{1+y}{2}. \quad (66)$$

First component of velocity and temperature distribution along with the boundary conditions are given by

$$a^2 B^2 M \frac{dv_0(y)}{dy}\frac{d\Theta_0(y)}{dy} + a^2 B^2 M\Theta_0(y)\frac{d^2v_0(y)}{dy^2} - a^2 B^2 \frac{d^2v_0(y)}{dy^2} + \Gamma v_0(y) + \frac{d^4v_1(y)}{dy^4} = 0, \quad (67)$$

$$v_1(y = \pm 1) = 0, \frac{d^2}{dy^2}v_1(y = \pm 1) = 0,$$

$$\frac{\gamma\left(\frac{d^2v_0(y)}{dy^2}\right)^2}{B^2} + \frac{\partial^2\Theta_1(y)}{dy^2} - a^2\gamma M\Theta_0(y)\left(\frac{dv_0(y)}{dy}\right)^2 + a^2\gamma\left(\frac{dv_0(y)}{dy}\right)^2 = 0, \quad (68)$$

$$\Theta_1(y = -1) = 0, \Theta_1(y = 1) = 0._b$$

Their solutions are

$v_1(y) =$
$\frac{1}{40320}(3416a^2 AB^2 \sin(\alpha) - 1708a^2 AB^2 M \sin(\alpha) +$
$16a^2 AB^2 My^7 \sin(\alpha) + 28a^2 AB^2 My^6 \sin(\alpha) -$
$168a^2 AB^2 My^5 \sin(\alpha) - 420a^2 AB^2 My^4 \sin(\alpha) +$

Rashid Nawaz, Muhammad Farooq, Siddiq Ur Rehman, Gamal M. Ismail, Hijaz Ahmad
Analytical Analysis of Thermal Conductivity by Laminar Flow under the Influence of MHD using Vogel's Model

DOI 10.2478/ama-2024-0058

$448a^2AB^2My^3\sin(\alpha) + 2100a^2AB^2My^2\sin(\alpha) - 296a^2AB^2My\sin(\alpha) - 56a^2AB^2y^6\sin(\alpha) + 840a^2AB^2y^4\sin(\alpha) - 4200a^2AB^2y^2\sin(\alpha) - 16a^2B^2GMy^7 - 28a^2B^2GMy^6 + 168a^2B^2GMy^5 + 420a^2B^2GMy^4 - 448a^2B^2GMy^3 - 2100a^2B^2GMy^2 + 296a^2B^2GMy + 1708a^2B^2GM + 56a^2B^2Gy^6 - 840a^2B^2Gy^4 + 4200a^2B^2Gy^2 - 3416a^2B^2G + 1385A\Gamma\sin(\alpha) + A\Gamma y^8\sin(\alpha) - 28A\Gamma y^6\sin(\alpha) + 350A\Gamma y^4\sin(\alpha) - 1708A\Gamma y^2\sin(\alpha) - 1385\Gamma G + \Gamma(-G)y^8 + 28\Gamma Gy^6 - 350\Gamma Gy^4 + 1708\Gamma Gy^2),$ (69)

$\Theta_1(y) =$
$\frac{1}{181440B^2}(2862a^2A^2B^2\gamma\sin^2(\alpha) - 1431a^2A^2B^2\gamma M\sin^2(\alpha) + 35a^2A^2B^2\gamma My^9\sin^2(\alpha) + 45a^2A^2B^2\gamma My^8\sin^2(\alpha) - 360a^2A^2B^2\gamma My^7\sin^2(\alpha) - 504a^2A^2B^2\gamma My^6\sin^2(\alpha) + 1134a^2A^2B^2\gamma My^5\sin^2(\alpha) + 1890a^2A^2B^2\gamma My^4\sin^2(\alpha) - 809a^2A^2B^2\gamma My\sin^2(\alpha) - 90a^2A^2B^2\gamma y^8\sin^2(\alpha) + 1008a^2A^2B^2\gamma y^6\sin^2(\alpha) - 3780a^2A^2B^2\gamma y^4\sin^2(\alpha) - 5724a^2AB^2\gamma G\sin(\alpha) + 2862a^2AB^2\gamma GM\sin(\alpha) - 70a^2AB^2\gamma GMy^9\sin(\alpha) - 90a^2AB^2\gamma GMy^8\sin(\alpha) + 720a^2AB^2\gamma GMy^7\sin(\alpha) + 1008a^2AB^2\gamma GMy^6\sin(\alpha) - 2268a^2AB^2\gamma GMy^5\sin(\alpha) - 3780a^2AB^2\gamma GMy^4\sin(\alpha) + 1618a^2AB^2\gamma GMy\sin(\alpha) + 180a^2AB^2\gamma Gy^8\sin(\alpha) - 2016a^2AB^2\gamma Gy^6\sin(\alpha) + 7560a^2AB^2\gamma Gy^4\sin(\alpha) + 2862a^2B^2\gamma G^2 - 1431a^2B^2\gamma G^2M + 35a^2B^2\gamma G^2My^9 + 45a^2B^2\gamma G^2My^8 - 360a^2B^2\gamma G^2My^7 - 504a^2B^2\gamma G^2My^6 + 1134a^2B^2\gamma G^2My^5 + 1890a^2B^2\gamma G^2My^4 - 809a^2B^2\gamma G^2My - 90a^2B^2\gamma G^2y^8 + 1008a^2B^2\gamma G^2y^6 - 3780a^2B^2\gamma G^2y^4 + 16632A^2\gamma\sin^2(\alpha) - 1512A^2\gamma y^6\sin^2(\alpha) + 7560A^2\gamma y^4\sin^2(\alpha) - 22680A^2\gamma y^2\sin^2(\alpha) - 33264A\gamma G\sin(\alpha) + 3024A\gamma Gy^6\sin(\alpha) - 15120A\gamma Gy^4\sin(\alpha) + 45360A\gamma Gy^2\sin(\alpha) + 16632\gamma G^2 - 1512\gamma G^2y^6 + 7560\gamma G^2y^4 - 22680\gamma G^2y^2),$ (70)

Second component of velocity and temperature distribution along with the boundary conditions are given by

$a^2B^2M\Theta_0(y)\frac{d^2v_1(y)}{dy^2}a^2B^2M\Theta_1(y)\frac{d^2v_0(y)}{dy^2} - a^2B^2\frac{d^2v_1(y)}{dy^2} + \frac{d^4v_2(y)}{dy^4}a^2B^2M\frac{dv_1(y)}{dy}\frac{d\Theta_0(y)}{dy} + a^2B^2M\frac{dv_0(y)}{dy}\frac{d\Theta_1(y)}{dy} + \Gamma v_1(y) = 0,$ (71)

$v_2(y = \pm 1) = 0, \frac{d^2}{dy^2}v_2(y = \pm 1) = 0,$

$\frac{2\gamma\frac{d^2v_0(y)}{dy^2}\frac{d^2v_1(y)}{dy^2}}{B^2} + \frac{d^2\Theta_2(y)}{dy^2} + 2a^2\gamma M\Theta_0(y)\frac{dv_0(y)}{dy}\frac{dv_1(y)}{dy} - a^2\gamma M\Theta_1(y)\left(\frac{dv_0(y)}{dy}\right)^2 + 2a^2\gamma\frac{dv_0(y)}{dy}\frac{dv_1(y)}{dy} = 0,$ (72)

$\Theta_2(y = -1) = 0, \Theta_2(y = 1) = 0.$

Their solutions are

$v_2(y) =$
$\frac{1}{261534873600}\Big((y^2 - 1)(G - A\sin(\alpha))\Big(a^4B^2\Big(72072B^2\Big(M^2(y(y(y(y(y(y(y(7y +$

$25) - 83) - 515) - 209) + 3643) + 7111) - 6737) - 32714) + 2M(y(6737 - y(y(5y(y(y(5y + 9) - 103) - 243) + 3643) + 14535)) + 62325) + 90(y^6 - 27y^4 + 323y^2 - 1385)\Big) - \gamma M(A^2 + 2G^2)(1540My^{13} + 2475(M - 2)y^{12} - 31010My^{11} - 55674(M - 2)y^{10} + 237622My^9 + 511893(M - 2)y^8 - 573188My^7 - 1515132(M - 2)y^6 - 1035936My^5 - 2947563(M - 2)y^4 + 3822918My^3 + 18538902(M - 2)y^2 - 3536826My - 54064629(M - 2))\Big) +$

$6552a^2\Big(\gamma M(21y^{10} - 287y^8 + 2188y^6 - 10748y^4 + 27367y^2 - 64109)(A^2 + 2G^2) + B^2\Gamma(6My^9 + 11(M - 2)y^8 - 214My^7 - 484(M - 2)y^6 + 3614My^5 + 11066(M - 2)y^4 - 27010My^3 - 129844(M - 2)y^2 + 50804My + 555731(M - 2))\Big) +$

$a^2A\gamma M\Big(a^2B^2(1540My^{13} + 2475(M - 2)y^{12} - 31010My^{11} - 55674(M - 2)y^{10} + 237622My^9 + 511893(M - 2)y^8 - 573188My^7 - 1515132(M - 2)y^6 - 1035936My^5 - 2947563(M - 2)y^4 + 3822918My^3 + 18538902(M - 2)y^2 - 3536826My - 54064629(M - 2)) - 6552(21y^{10} - 287y^8 + 2188y^6 - 10748y^4 + 27367y^2 - 64109)\Big)(A\cos(2\alpha) + 4G\sin(\alpha)) + 546\Gamma^2(y^{10} - 65y^8 + 2410y^6 - 53954y^4 + 631621y^2 - 2702765)\Big)\Big),$ (73)

$\Theta_2(y) =$
$\frac{1}{160059342643200B^2}\Big(\gamma(y^2 - 1)(G - A\sin(\alpha))^2\Big(a^4B^2\Big(\gamma M(A^2 + 2G^2)(1576575My^{15} + 2297295(M - 2)y^{14} - 31679505My^{13} - 49808385(M - 2)y^{12} + 251771355My^{11} + 443901195(M - 2)y^{10} - 866969541My^9 - 1717393941(M - 2)y^8 + 732117949My^7 + 1691135469(M - 2)y^6 + 2148126829My^5 + 5197726557(M - 2)y^4 - 2312301143My^3 - 7951990023(M - 2)y^2 - 2312301143My - 7951990023(M - 2)) - 8019648B^2(70M^2y^{10} + 210(M - 2)My^9 - 14(M(61M + 44) - 44)y^8 - 3640(M - 2)My^7 + (M(1621M + 12254) - 12254)y^6 + 22100(M - 2)My^5 + (M(15129M - 89386) + 89386)y^4 - 47728(M - 2)My^3 + ((257114 - 65391M)M - 257114)y^2 - 35518(M - 2)My + M(257114 - 65391M) - 257114)\Big) - 102816a^2\Big(33\gamma M(30y^{12} - 425y^{10} + 2851y^8 - 11384y^6 + 25198y^4 - 19847y^2 - 19847)(A^2 + 2G^2) + 2B^2\Big(\Gamma(165My^{11} + 195(M - 2)y^{10} - 5451My^9 - 6669(M - 2)y^8 + 79634My^7 + 102726(M - 2)y^6 - 503806My^5 - 714090(M - 2)y^4 + 1144841My^3 + 2033655(M - 2)y^2 + 1144841My + 2033655(M - 2)) - 1287(140My^7 + 225(M - 2)y^6 - 1300My^5 - 2715(M - 2)y^4 + 3236My^3 + 8835(M - 2)y^2 -$

$3484My - 22665(M-2)))) -$

$a^2 A\gamma M \big(a^2 B^2 (1576575My^{15} + 2297295(M-2)y^{14} - 31679505My^{13} - 49808385(M-2)y^{12} + 251771355My^{11} + 443901195(M-2)y^{10} - 866969541My^9 - 1717393941(M-2)y^8 + 732117949My^7 + 1691135469(M-2)y^6 + 2148126829My^5 + 5197726557(M-2)y^4 - 2312301143My^3 - 7951990023(M-2)y^2 - 2312301143My - 7951990023(M-2)) - 3392928(30y^{12} - 425y^{10} + 2851y^8 - 11384y^6 + 25198y^4 - 19847y^2 - 19847) \big) (A\cos(2\alpha) + 4G\sin(\alpha)) + 352864512\Gamma(7y^8 - 173y^6 + 1717y^4 - 5423y^2 + 13792) \big) \Big).$ \hfill (74)

The second order HPM solution of velocity profile and temperature distribution are as under.

$$v_{HPM}(y) = v_0(y) + v_1(y) + v_2(y). \tag{75}$$

and

$$\Theta_{HPM}(y) = \Theta_0(y) + \Theta_1(y) + \Theta_2(y). \tag{76}$$

Using of $\gamma = 0.1, B = 0.1, A = 0.04, \alpha = 135, M = 0.005, G = 1, \Gamma = 0.01, a = 0.2$ in Eq. (75) and Eq. (76) one obtains the following solutions.

$v_{HPM}(y) = 0.207222 + y(1.46023 \times 10^{-8} + y(-0.248654 + y(-2.21003 \times 10^{-8} + y(0.0414248 + y(8.28676 \times 10^{-9} + y(7.44514 \times 10^{-6} + y(-7.88697 \times 10^{-10} + y(-2.44068 \times 10^{-7} + y(-1.20551 \times 10^{-13} + y(-3.21233 \times 10^{-10} + y(2.91432 \times 10^{-15} + y(2.10224 \times 10^{-11} + y(9.8514 \times 10^{-18} + (2.98879 \times 10^{-16} - 4.66088 \times 10^{-19}y)y))))))))))))),$ \hfill (77)

and

$\Theta_{HPM}(y) = 1.40694 + y(0.5 + y(-1.23656 + y(-2.20644 \times 10^{-7} + y(0.41193 + y(2.72583 \times 10^{-7} + y(-0.0822714 + y(-8.65162 \times 10^{-8} + y(-0.0000441843 + y(8.403 \times 10^{-9} + y(1.50419 \times 10^{-6} + y(1.59405 \times 10^{-12} + y(3.82264 \times 10^{-9} + y(-4.06999 \times 10^{-14} + y(-2.50695 \times 10^{-10} + y(-1.63881 \times 10^{-16} + (-4.51697 \times 10^{-15} + 7.76913 \times 10^{-18}y)y)))))))))))))))).$ \hfill (78)

4.3. Volumetric flux, average velocity, shear stress, skin friction

Dimensionless representation of the volume flux can be expressed as

$$Q = \int_{-1}^{1} v(y)\, dy. \tag{79}$$

By plugging in Eq.(61) and (77) into Eq.(79), we get the following equations

$Q_{OHAM} = -\frac{1}{261534873600}((G - A\sin(\alpha))(a^4 B^2 (B^2 ((1.08116 \times 10^{10} - 2.844 \times 10^9 M)M - 1.08116 \times 10^{10}) + \gamma G^2 M(1.80917 \times 10^8 M - 3.61833 \times 10^8)) +$

$a^2 A\gamma M \sin(\alpha)(A\sin(\alpha)(a^2 B^2 (1.80917 \times 10^8 M - 3.61833 \times 10^8) - 1.3911 \times 10^9) + G(a^2 B^2 (7.23667 \times 10^8 - 3.61833 \times 10^8 M) + 2.78219 \times 10^9)) + a^2 (B^2 (-8.76344 \times 10^9 \Gamma + (4.38172 \times 10^9 \Gamma - 1.41092 \times 10^{10})M + 2.82185 \times 10^{10}) - 1.3911 \times 10^9 \gamma G^2 M) + \Gamma(1.14349 \times 10^{10} - 1.77584 \times 10^9 \Gamma) - 6.97426 \times 10^{10})),$ \hfill (80)

$Q_{HPM} = -\frac{1}{85135050}((G - A\sin(\alpha))(a^4 B^2 (21991\gamma(M-2)M(A^2 + 2G^2) - 1092B^2 (M(897M - 3410) + 3410)) + 42a^2 (13B^2 (2764\Gamma - 8415)(M-2) - 4026\gamma M(A^2 + 2G^2)) - a^2 A\gamma M(21991a^2 B^2 (M-2) - 169092)(A\cos(2\alpha) + 4G\sin(\alpha)) - 56(\Gamma(10922\Gamma - 66495) + 405405))).$ \hfill (81)

The couple stress fluid's average velocity is represented by the symbol $\bar{v}$ and its definition is as follows:

$$\bar{v} = \frac{Q}{d}. \tag{82}$$

Dimensionless form of (82) correspondence with the flow rate given in (80) and (81).

On the surface of the upper plate, the dimensionless shear stress S_p may be calculated using the following formula:

$$S_p = -\mu \frac{dv}{dy}\Big|_{y=1}. \tag{83}$$

In this case, there is a minus sign because the top plate is pointing in the negative y direction of the coordinate system. Putting v_{OHAM} and v_{HPM} in (83) we get the following equations:

$S_{OHAM} = \frac{1}{130767436800}(\mu(G - A\sin(\alpha))(c_1 (2a^2 \gamma c_4 M(a^2 B^2 (-39529728(M-2) - 1114880M) + 298561536)(G - A\sin(\alpha))^2 - 12972960(4a^2 B^2 (376M - 672) - 1088\Gamma)) - 6486480(c_2 (4a^2 B^2 (336(M-2) + 40M) - 1088\Gamma) - 6720) - 546c_1^2 (132a^4 B^4 (-29472M^2 + 105088M - 97920) + 12a^2 B^2 (\Gamma(436480(M-2) + 27200M) + 3960(376M - 672)) + \Gamma(-2122752\Gamma - 12925440)))),$ \hfill (84)

$S_{HPM} = -\frac{1}{510810300}(\mu(G - A\sin(\alpha))(a^4 B^2 (2\gamma M(79384M - 154413)(A^2 + 2G^2) - 9009B^2 (M(921M - 3284) + 3060)) - 1638a^2 (712\gamma M(A^2 + 2G^2) + 5B^2 (44(62\Gamma - 189) + 9(517 - 161\Gamma)M)) - 2a^2 A\gamma M(a^2 B^2 (79384M - 154413) - 583128)(A\cos(2\alpha) + 4G\sin(\alpha)) - 3276(\Gamma(1382\Gamma - 8415) + 51975))).$ \hfill (85)

The opposite resistive force which is created between the surface of the body and particles of the fluid is called skin friction coefficient. The formula given below represents the skin friction coefficient at both plates.

$$\Theta'(y) = \frac{1.328}{\sqrt{\frac{\rho v H}{\mu}}}. $$

The volume fluxes that were calculated using OHAM and HPM are denoted here by Q_{OHAM} and Q_{HPM} respectively. Moreover, Eqs. (84) and (85) represents the shear stresses obtained by putting v_{OHAM} & v_{HPM} in eq. (83).

5. RESULTS AND DISCUSSION

Throughout the course of this investigation, we monitored the flow of couple stress fluids as heat was transferred between two inclined plates that were parallel to one another. Analytical models of the velocity field and temperature distribution were built with the help of the OHAM and HPM. In this scenario, the impacts of several non-dimensional parameters on the velocity field, temperature distribution, volumetric flow rate, and shear stress are visually depicted. Tables 1 - 5 display OHAM and HPM solutions for fluid velocity, temperature distributions, residuals, absolute differences and skin friction for both approaches. It has been demonstrated that OHAM and HPM yield comparable outcomes.

Tab. 1. Comparing the results of OHAM & HPM for velocity profile, when $\gamma = 0.2, A = 0.4, B = 0.1, a = 0.2, \alpha = 165, M = 0.01, G\,1, \Gamma = 0.3$.

y	v_{OHAM}	Residual v_{OHAM}	v_{HPM}	Residual v_{HPM}	Average Residual
-1.	0.	-2.72735×10^{-9}	1.82554×10^{-17}	-1.0742×10^{-7}	5.50735×10^{-8}
-0.9	0.0190055	-5.93355×10^{-9}	0.0190078	1.43614×10^{-5}	7.17775×10^{-6}
-0.8	0.037468	-1.78626×10^{-8}	0.0374725	2.84835×10^{-5}	1.42328×10^{-5}
-0.7	0.054899	-4.10482×10^{-8}	0.0549056	4.19107×10^{-5}	2.09348×10^{-5}
-0.6	0.070869	-7.42714×10^{-8}	0.0708775	5.43114×10^{-5}	2.71186×10^{-5}
-0.5	0.085007	-1.1398×10^{-7}	0.0850172	6.53782×10^{-5}	3.26321×10^{-5}
-0.4	0.0969998	-1.55482×10^{-7}	0.0970115	7.48361×10^{-5}	3.73403×10^{-5}
-0.3	0.106592	-1.93888×10^{-7}	0.106605	8.24494×10^{-5}	4.11278×10^{-5}
-0.2	0.113585	-2.24802×10^{-7}	0.113599	8.80278×10^{-5}	4.39015×10^{-5}
-0.1	0.117837	-2.44795×10^{-7}	0.117851	9.14316×10^{-5}	4.55934×10^{-5}
0.	0.119264	-2.51702×10^{-7}	0.119278	9.25755×10^{-5}	4.61619×10^{-5}
0.1	0.117837	-2.44786×10^{-7}	0.117851	9.14309×10^{-5}	4.55931×10^{-5}
0.2	0.113585	-2.24786×10^{-7}	0.113599	8.80265×10^{-5}	4.39008×10^{-5}
0.3	0.106592	-1.93866×10^{-7}	0.106605	8.24476×10^{-5}	4.11268×10^{-5}
0.4	0.0969998	-1.55457×10^{-7}	0.0970115	7.48339×10^{-5}	3.73392×10^{-5}
0.5	0.085007	-1.13957×10^{-7}	0.0850172	6.53756×10^{-5}	3.26308×10^{-5}
0.6	0.070869	-7.4254×10^{-8}	0.0708776	5.43086×10^{-5}	2.71172×10^{-5}
0.7	0.054899	-4.10381×10^{-8}	0.0549056	4.19079×10^{-5}	2.09334×10^{-5}
0.8	0.037468	-1.78602×10^{-8}	0.0374725	2.84808×10^{-5}	1.42315×10^{-5}
0.9	0.0190055	-5.93575×10^{-9}	0.0190078	1.4359×10^{-5}	7.17655×10^{-6}
1.	0.	-2.72655×10^{-9}	1.75113×10^{-17}	-1.09349×10^{-7}	5.6038×10^{-8}

Tab. 2. Comparing the results of OHAM & HPM for temperature distribution, when $\gamma = 0.03, A = 0.004, B = 0.1, \alpha = 135, M = 0.00005, G = 0.1, \Gamma = 0.01, a = 0.2$.

y	Θ_{OHAM}	Residual Θ_{OHAM}	Θ_{HPM}	Residual Θ_{HPM}	Average Residual
-1.	-1.06962×10^{-19}	4.58588×10^{-10}	-9.06168×10^{-19}	1.2674×10^{-11}	2.35631×10^{-10}
-0.9	0.0503963	9.85987×10^{-8}	0.0503956	2.13349×10^{-9}	5.03661×10^{-8}
-0.8	0.10079	3.47926×10^{-7}	0.100788	8.04974×10^{-9}	1.77988×10^{-7}
-0.7	0.151173	6.89008×10^{-7}	0.151171	1.69256×10^{-8}	3.52967×10^{-7}
-0.6	0.201537	1.07253×10^{-6}	0.201534	2.77568×10^{-8}	5.50143×10^{-7}
-0.5	0.25187	1.45826×10^{-6}	0.251867	3.94455×10^{-8}	7.48854×10^{-7}
-0.4	0.302162	1.81409×10^{-6}	0.302158	5.08813×10^{-8}	9.32485×10^{-7}
-0.3	0.352401	2.11508×10^{-6}	0.352397	6.1022×10^{-8}	1.08805×10^{-6}

-0.2	0.402579	2.34273×10^{-6}	0.402574	6.89681×10^{-8}	1.20585×10^{-6}
-0.1	0.452689	2.48428×10^{-6}	0.452684	7.40271×10^{-8}	1.27915×10^{-6}
0.	0.502726	2.53228×10^{-6}	0.502721	7.5763×10^{-8}	1.30402×10^{-6}
0.1	0.552689	2.48428×10^{-6}	0.552684	7.40271×10^{-8}	1.27915×10^{-6}
0.2	0.602579	2.34273×10^{-6}	0.602574	6.8968×10^{-8}	1.20585×10^{-6}
0.3	0.652401	2.11508×10^{-6}	0.652397	6.1022×10^{-8}	1.08805×10^{-6}
0.4	0.702162	1.81409×10^{-6}	0.702158	5.08813×10^{-8}	9.32486×10^{-7}
0.5	0.75187	1.45826×10^{-6}	0.751867	3.94454×10^{-8}	7.48855×10^{-7}
0.6	0.801537	1.07253×10^{-6}	0.801534	2.77567×10^{-8}	5.50143×10^{-7}
0.7	0.851173	6.89009×10^{-7}	0.851171	1.69256×10^{-8}	3.52967×10^{-7}
0.8	0.90079	3.47927×10^{-7}	0.900788	8.04971×10^{-9}	1.77988×10^{-7}
0.9	0.950396	9.85988×10^{-8}	0.950396	2.13348×10^{-9}	5.03661×10^{-8}
1.	1.	4.58565×10^{-10}	1.	1.26734×10^{-11}	2.35619×10^{-10}

Tab. 3. Comparison of OHAM & HPM for velocity profile, when $\gamma = 0.2, A = 0.04, B = 0.1, a = 0.2, \alpha = 165, M = 0.05, G = 1, \Gamma = 0.3$.

y	v_{OHAM}	v_{HPM}	Absolute Difference
-1.	0.	-7.32191×10^{-19}	1.01466×10^{-17}
-0.9	0.0303673	0.0303708	3.57024×10^{-6}
-0.8	0.0598669	0.059874	7.05155×10^{-6}
-0.7	0.0877185	0.0877288	1.03576×10^{-5}
-0.6	0.113236	0.113249	1.3407×10^{-5}
-0.5	0.135826	0.135842	1.61247×10^{-5}
-0.4	0.154988	0.155006	1.84444×10^{-5}
-0.3	0.170314	0.170335	2.03097×10^{-5}
-0.2	0.181488	0.181509	2.16752×10^{-5}
-0.1	0.188282	0.188305	2.2508×10^{-5}
0.	0.190562	0.190585	2.27878×10^{-5}
0.1	0.188282	0.188305	2.25079×10^{-5}
0.2	0.181488	0.181509	2.16751×10^{-5}
0.3	0.170314	0.170335	2.03095×10^{-5}
0.4	0.154988	0.155006	1.84442×10^{-5}
0.5	0.135826	0.135842	1.61245×10^{-5}
0.6	0.113236	0.113249	1.34067×10^{-5}
0.7	0.0877185	0.0877289	1.03574×10^{-5}
0.8	0.059867	0.059874	7.05141×10^{-6}
0.9	0.0303673	0.0303709	3.57017×10^{-6}
1.	0.	-8.61145×10^{-20}	1.09648×10^{-17}

Tab. 4. Comparison of OHAM & HPM for temperature distribution, when $\gamma = 0.03, A = 0.04, B = 0.1, \alpha = 135, M = 0.00005$, $G = 0.1, \Gamma = 0.01, a = 0.2$.

y	Θ_{OHAM}	Θ_{HPM}	Absolute Difference
-1.	-1.06962×10^{-19}	-9.06168×10^{-19}	5.96311×10^{-19}
-0.9	0.0503963	0.0503956	7.16137×10^{-7}
-0.8	0.10079	0.100788	1.42771×10^{-6}

-0.7	0.151173	0.151171	2.12385×10^{-6}
-0.6	0.201537	0.201534	2.78821×10^{-6}
-0.5	0.25187	0.251867	3.40088×10^{-6}
-0.4	0.302162	0.302158	3.9404×10^{-6}
-0.3	0.352401	0.352397	4.38587×10^{-6}
-0.2	0.402579	0.402574	4.71879×10^{-6}
-0.1	0.452689	0.452684	4.92468×10^{-6}
0.	0.502726	0.502721	4.99437×10^{-6}
0.1	0.552689	0.552684	4.92468×10^{-6}
0.2	0.602579	0.602574	4.71879×10^{-6}
0.3	0.652401	0.652397	4.38587×10^{-6}
0.4	0.702162	0.702158	3.9404×10^{-6}
0.5	0.75187	0.751867	3.40087×10^{-6}
0.6	0.801537	0.801534	2.78821×10^{-6}
0.7	0.851173	0.851171	2.12385×10^{-6}
0.8	0.90079	0.900788	1.42771×10^{-6}
0.9	0.950396	0.950396	7.16137×10^{-7}
1.	1.	1.	1.62664×10^{-17}

Tab. 5. Calculations for $\Theta'(-1)$ and $\Theta'(1)$ against different values of γ keeping $G = 0.001, a = 0.2, M = 0.00015, \alpha = 155, A = 0.02, \Gamma = 0.02, B = 0.3$ fixed.

γ	$\Theta'_{OHAM}(-1)$	$\Theta'_{HPM}(-1)$	Absolute Difference	$\Theta'_{OHAM}(1)$	$\Theta'_{HPM}(1)$	Absolute Difference
-1.	0.499496	0.499499	2.79823×10^{-6}	0.500504	0.500501	2.79821×10^{-6}
-0.877	0.499558	0.499561	2.45405×10^{-6}	0.500442	0.500439	2.45403×10^{-6}
-0.754	0.49962	0.499622	2.10986×10^{-6}	0.50038	0.500378	2.10985×10^{-6}
-0.631	0.499682	0.499684	1.76568×10^{-6}	0.500318	0.500316	1.76567×10^{-6}
-0.508	0.499744	0.499745	1.4215×10^{-6}	0.500256	0.500255	1.42149×10^{-6}
-0.385	0.499806	0.499807	1.07732×10^{-6}	0.500194	0.500193	1.07731×10^{-6}
-0.262	0.499868	0.499869	7.33135×10^{-7}	0.500132	0.500131	7.33132×10^{-7}
-0.139	0.49993	0.49993	3.88954×10^{-7}	0.50007	0.50007	3.88952×10^{-7}
-0.016	0.499992	0.499992	4.47716×10^{-8}	0.500008	0.500008	4.47714×10^{-8}
0.107	0.500054	0.500054	2.9941×10^{-7}	0.499946	0.499946	2.99409×10^{-7}
0.23	0.500116	0.500115	6.43592×10^{-7}	0.499884	0.499885	6.43589×10^{-7}
0.353	0.500178	0.500177	9.87774×10^{-7}	0.499822	0.499823	9.8777×10^{-7}
0.476	0.50024	0.500239	1.33196×10^{-6}	0.49976	0.499761	1.33195×10^{-6}
0.599	0.500302	0.5003	1.67614×10^{-6}	0.499698	0.4997	1.67613×10^{-6}
0.722	0.500364	0.500362	2.02032×10^{-6}	0.499636	0.499638	2.02031×10^{-6}
0.845	0.500426	0.500423	2.3645×10^{-6}	0.499574	0.499577	2.36449×10^{-6}
0.968	0.500488	0.500485	2.70868×10^{-6}	0.499512	0.499515	2.70867×10^{-6}

Fig. 2. Velocity Profile for $B = 2.5, \gamma = 2, A = 1, a = 0.1, \alpha = 165, M = 0.002, \Gamma = 0.2$ using OHAM

Fig. 3. Velocity Profile for $B = 2.5, \gamma = 2, A = 1, a = 0.1, \alpha = 165, M = 0.002, \Gamma = 0.2$ using HPM

Fig. 4. Velocity Profile for $B = 1.4, \gamma = 0.1, G = 1, \alpha = 135, a = 3, M = 1.15, \Gamma = 0.2$ using OHAM
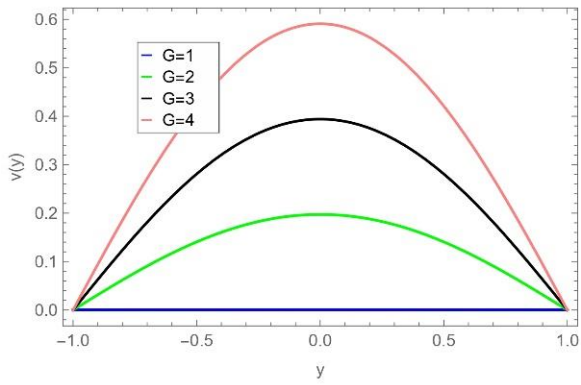
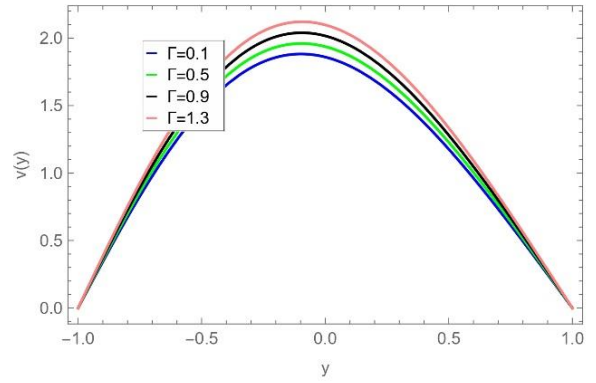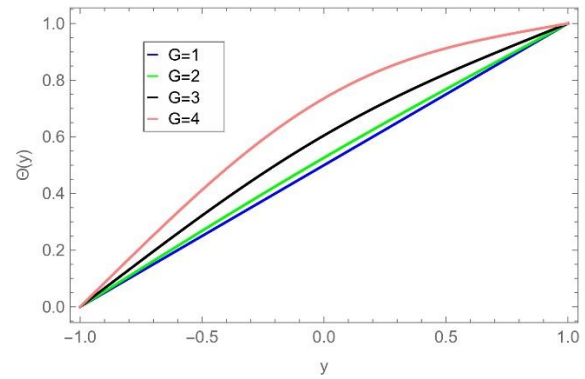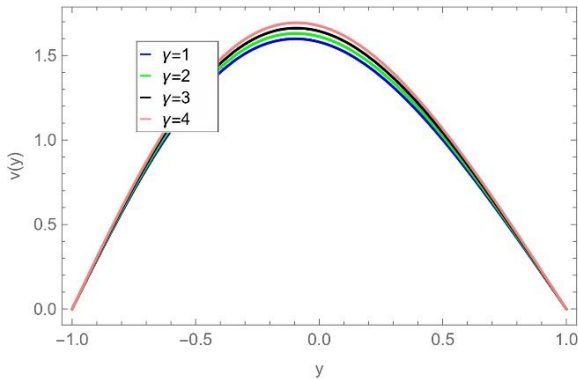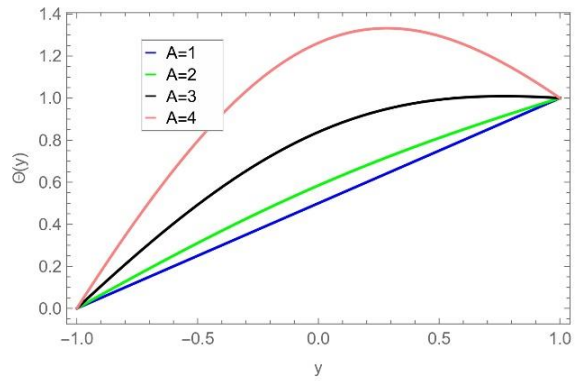**Fig. 5.** Velocity Profile for $B = 1.4, \gamma = 0.1, G = 1, \alpha = 135, a = 3, M = 1.15, \Gamma = 0.2$ using HPM
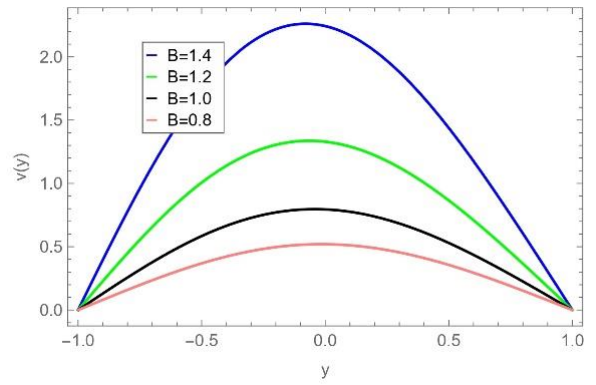
Fig. 6. Velocity Profile for $G = 1, \gamma = 3, A = -1, a = 3 \, \alpha = 135, M = 1.15, \Gamma = 0.2$ using OHAM
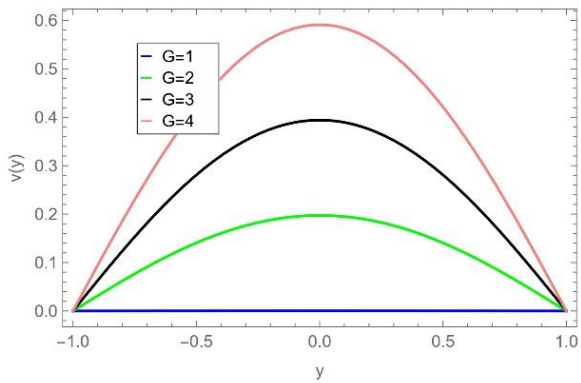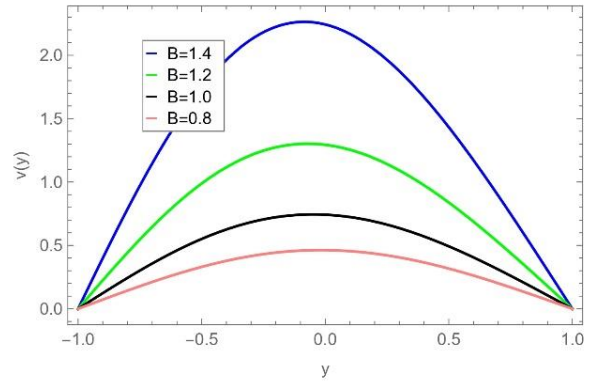
Fig. 7. Velocity Profile for $G = 1, \gamma = 3, A = -1, a = 3 \, \alpha = 135, M = 1.15, \Gamma = 0.2$ using HPM

Fig. 8. Velocity Profile for $G = 1, \gamma = 0.1, A = -1, a = 3, \alpha = 135, M = 1.15, B = 1.4$ using OHAM

Fig. 9. Velocity Profile for $G = 1, \gamma = 0.1, A = -1, a = 3, \alpha = 135, M = 1.15, B = 1.4$ using HPM

sciendo

Rashid Nawaz, Muhammad Farooq, Siddiq Ur Rehman, Gamal M. Ismail, Hijaz Ahmad   DOI 10.2478/ama-2024-0058
Analytical Analysis of Thermal Conductivity by Laminar Flow under the Influence of MHD using Vogel's Model
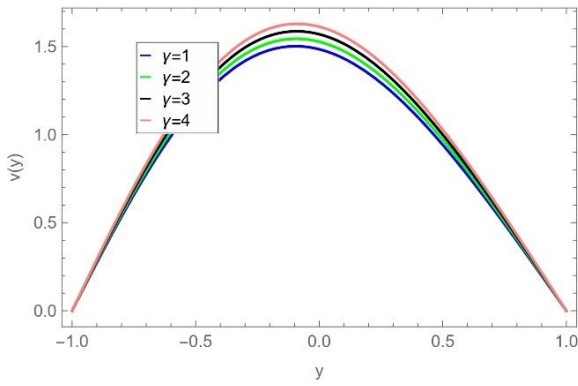
Fig. 10. Velocity Profile for $B = 2.1, A = 1.1, G = 1, \alpha = 13, a = 2, M = 1.15, \Gamma = 0.2$ using OHAM

Fig. 11. Velocity Profile for $B = 2.1, A = 1.1, G = 1, \alpha = 13, a = 2, M = 1.15, \Gamma = 0.2$ using HPM

Figures 2-11 for the velocity profile in the case of inclined Poiseuille flow are presented to explore the impacts of various parameters. As we can see in these graphs, changes in parameter values create changes in fluid velocity. Both Figures 4 and 5 show that the pressure gradient A is inversely proportional to the fluid's velocity. The results of parameter B's influence on the flow velocity are shown in Figures 6 and 7. If you raise B, the fluid's velocity goes up, and if you lower B, it slows down. The impact of the MHD parameter Γ on the velocity field is seen in Figures 8 and 9. The fluid's velocity and Γ are directly related to one another. The MHD parameter controls the relationship between magnetic fields and fluid velocity on an inclined plane. Increased parameter values modify velocity profiles and flow stability by amplifying magnetic effects.

Fig. 12. Temperature distribution for $B = 2.5, \gamma = 2, A = 1, a = 0.1, \alpha = 165, M = 0.002, \Gamma = 0.2$ using OHAM

Fig. 13. Temperature distribution for $B = 2.5, \gamma = 2, A = 1, a = 0.1, \alpha = 165, M = 0.002, \Gamma = 0.2$ using HPM

Fig. 14. Temperature distribution for $B = 0.3, \gamma = 0.1, G = 1, \alpha = 165, a = 3, M = 0.015, \Gamma = 0.2$ using OHAM
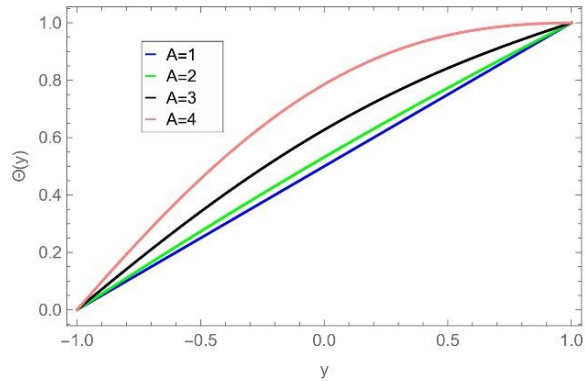
Fig. 15. Temperature distribution for $B = 0.3, \gamma = 0.1, G = 1, \alpha = 165, a = 3, M = 0.015, \Gamma = 0.2$ using HPM
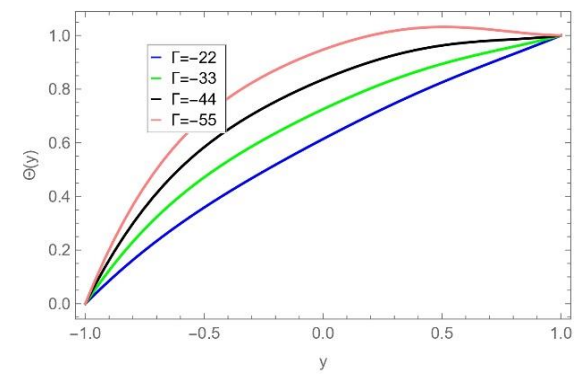
Fig. 16. Temperature distribution for $G = 2, \gamma = 0.2, A = 0.3, a = 3, \alpha = 135, M = 1.15, B = 1.4$ using OHAM
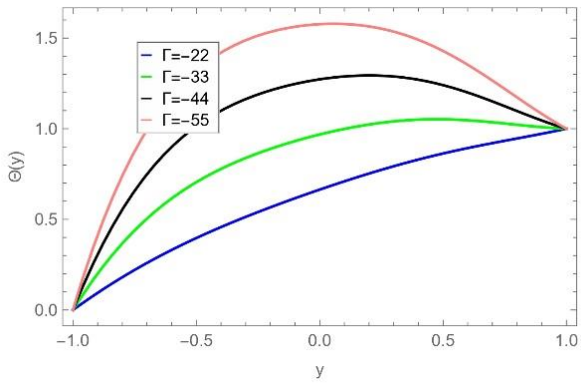
Fig. 17. Temperature distribution for $G = 2, \gamma = 0.2, A = 0.3, a = 3, \alpha = 135, M = 1.15, B = 1.4$ using HPM

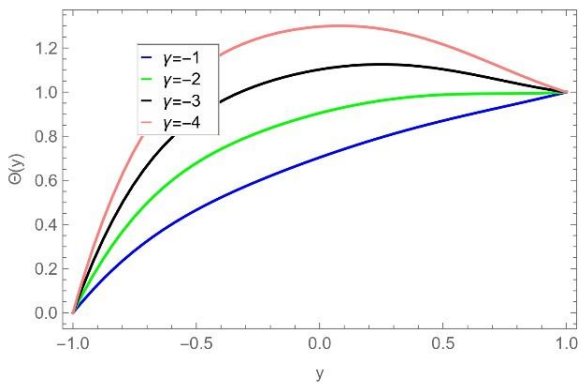**Fig. 18.** Temperature distribution for $B = 2.1, A = 1.1, G = 1, \alpha = 135, a = 2, M = 1.15, \Gamma = 0.2$ using OHAM

Fig. 19. Temperature distribution for $B = 2.1, A = 1.1, G = 1, \alpha = 135, a = 2, M = 1.15, \Gamma = 0.2$ using HPM
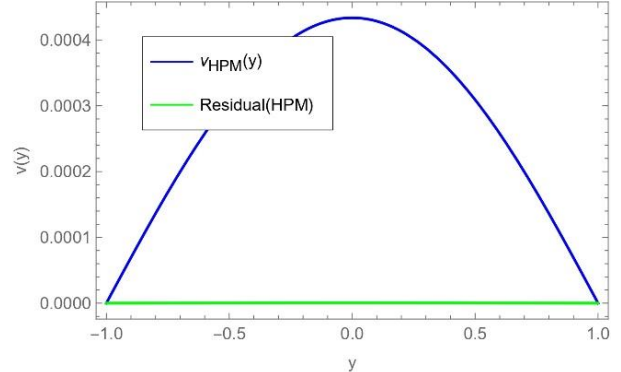

Fig. 20. Shear Stress for $\gamma = 2, M = 0.002, a = 0.1, \alpha = 165, B = 2.5, \Gamma = 0.2, A = 1$ using OHAM
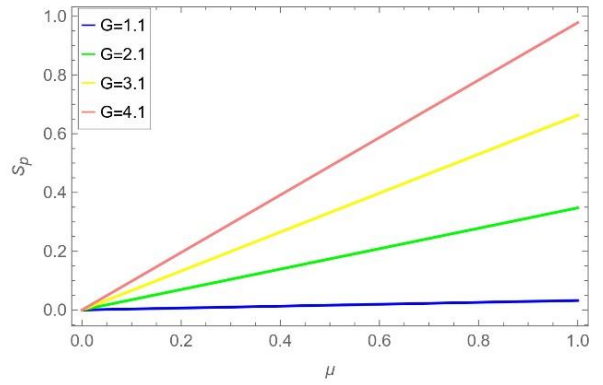
Fig. 21. Shear Stress for $\gamma = 2, M = 0.002, a = 0.1, \alpha = 165, B = 2.5, \Gamma = 0.2, A = 1$ using HPM
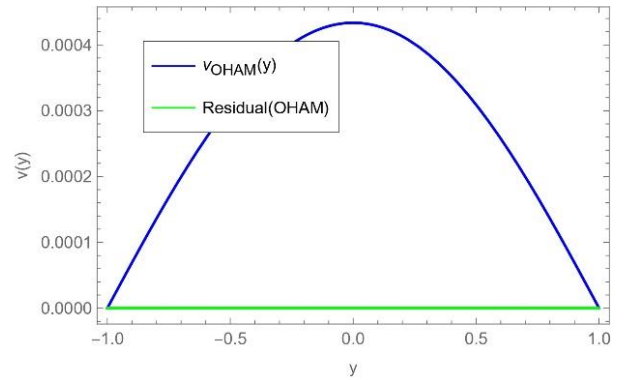
Fig. 22. Error graph of vOHAM for $\gamma = 2, M = 0.002, a = 0.1, \alpha = 165, B = 2.5, \Gamma = 0.2, A = 1$ and $G = 1$

Fig. 23. Error graph of vHPM for $\gamma = 2, M = 0.002, a = 0.1, \alpha = 165, B = 2.5, \Gamma = 0.2, A = 1$ and $G = 1$
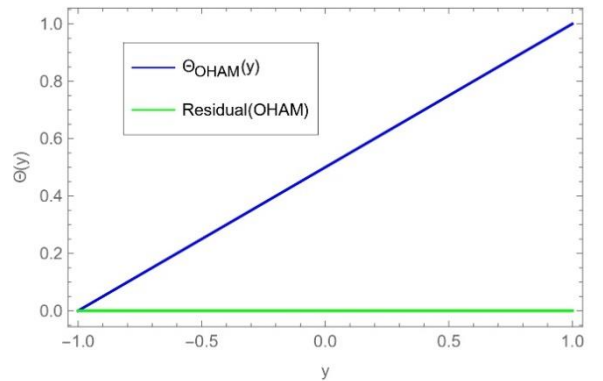
Fig. 24. Error graph of ΘOHAM for $\gamma = 2, M = 0.002, a = 0.1, \alpha = 165, B = 2.5, \Gamma = 0.2, A = 1$ and $G = 1$
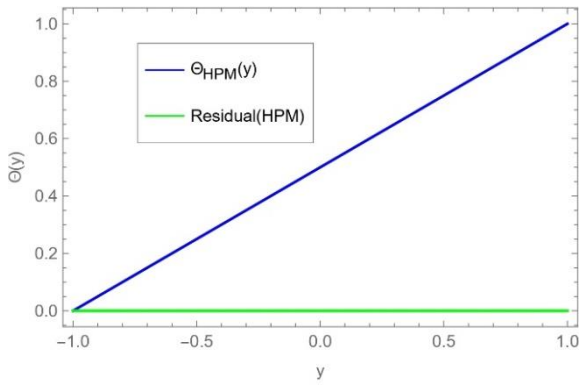
Fig. 25. Error graph of ΘHPM for $\gamma = 2, M = 0.002, a = 0.1, \alpha = 165, B = 2.5, \Gamma = 0.2, A = 1$ and $G = 1$

In Figures 12-19 we observed the effect of numerous parameters on temperature distribution Θ. The increase or decrease in the values of these parameters causes change in the temperature of the fluid. In Figures 14 and 15 we have seen that increase in the value of A causes increase in the temperature of the fluid. The parameter A is directly related to the temperature distribution Θ. In figures 16 and 17 a reduction in the MHD parameter enhances magnetic influence and encourages the conversion of magnetic energy into thermal energy. Due to increased heat production and decreased magnetic field resistance, this raises the fluid temperature. The influence of parameter γ on temperature distribution is visualized using graphs 18-19. Since fluid mixing is improved and internal thermal gradients are minimized as the Brinkman number rises, fluid momentum prevails over thermal conduction, resulting in a more uniform temperature distribution. This implies that viscous heating of the fluid produces more heat than heat transfer from the heated wall to the fluid. By varying the values of G, we can examine how shear stress S_p behaves in an inclined Poiseuille flow in Figures 20-21. The visual representation of error graphs is shown in Figures 22-25.

6. CONCLUSION

In this work, we have investigated the flow of couple stress fluids as heat was transferred between two inclined fixed plates that were parallel to each another. Analytical solutions of the problem have been obtained using HPM and OHAM. The results for velocity and temperature have been plotted graphically and discussed in detail. The important outcomes of the graphical analysis of the problem are as follows:

- MHD parameter Γ produces a Lorentz force, which suppresses turbulence and alters the velocity profile to produce a faster fluid flow. while, decreased temperature distribution in the fluid results from reduced energy loss due to magnetic drag reduction.
- Greater forced convection, thinner boundary layers, and enhanced fluid motion are all benefits of increased Brinkman number γ. Because of greater thermal mixing, increased heat transfer efficiency causes a reduction in fluid temperature distribution.

Nomenclature:

η	Couple Stress Parameter
$\frac{D}{Dt}$	Material Derivative
κ	Thermal Conductivity
L	Gradient of Velocity
μ	Viscosity Coefficient
ρ	Constant Density
Y	Velocity
A	Rivlin Ericksen tensor
B	Magnetic Induction
C	Unit Tensor
D	Current Density
T	Cauchy Stress Tensor
Θ	Temperature
c_p	Specific Heat
g	Body force
p	Dynamic Pressure
S	Extra Stress Tensor
Γ	MHD Parameter
γ	Brinkman number

REFERENCES

1. Stokes VK, Stokes VK. Couple stresses in fluids. Theories of Fluids with Microstructure: An Introduction. 1984;34-80.
2. Stokes, Baumann. Theories of Fluids with Microstructure. Springer Berlin Heidelberg. 1984.
3. Devakar M, Iyengar TK. Run up flow of a couple stress fluid between parallel plates. Nonlinear Analysis: Modelling and Control. 2010;15(1):29-37.
4. Devakar M, Iyengar TK. Stokes' problems for an incompressible couple stress fluid. Nonlinear Analysis: Modelling and Control. 2008;13(2):181-90.
5. Hayat T, Mustafa M, Iqbal Z, Alsaedi A. Stagnation-point flow of couple stress fluid with melting heat transfer. Applied Mathematics and Mechanics. 2013;34:167-76.
6. Akbar NS, Nadeem S. Intestinal flow of a couple stress nanofluid in arteries. IEEE transactions on nanobioscience. 2013;12(4):332-9.
7. Srinivasacharya D, Srinivasacharyulu N, Odelu O. Flow and heat transfer of couple stress fluid in a porous channel with expanding and contracting walls. International Communications in heat and mass Transfer. 2009;36(2):180-5.
8. Muthuraj R, Srinivas S, Shukla AK, Immaculate DL. Non-Darcian and thermal radiation effects on Magnetoconvection flow of Two-immiscible fluids with heat transfer. InProceedings of the 24th National and 2nd International ISHMT-ASTFE Heat and Mass Transfer Conference (IHMTC-2017) 2017. Begel House Inc..
9. Srinivasacharya D, Kaladhar K. Analytical solution of MHD free convective flow of couple stress fluid in an annulus with Hall and Ion-slip effects. Nonlinear Analysis: Modelling and Control. 2011;16(4):477-87.
10. Tsai CY, Novack M, Roffe G. Rheological and heat transfer characteristics of flowing coal-water mixtures. General Applied Science Labs. Inc. Ronkonkoma. NY (USA). 1988.
11. Yürüsoy M, Pakdemirli M. Approximate analytical solutions for the flow of a third-grade fluid in a pipe. International Journal of Non-Linear Mechanics. 2002;37(2):187-95.
12. Makinde OD. Laminar falling liquid film with variable viscosity along an inclined heated plate. Applied Mathematics and Computation. 2006;175(1):80-8.
13. Makinde OD. Thermal criticality for a reactive gravity driven thin film flow of a third-grade fluid with adiabatic free surface down an inclined plane. Applied Mathematics and Mechanics. 2009;30(3):373-80.

14. Makinde OD. Thermodynamic second law analysis for a gravity-driven variable viscosity liquid film along an inclined heated plate with convective cooling. Journal of Mechanical Science and Technology. 2010;24:899-908.

15. Qayyum M, Khan H, Rahim MT, Ullah I. Analysis of unsteady axisymmetric squeezing fluid flow with slip and no-slip boundaries using OHAM. Mathematical Problems in Engineering. 2015.

16. Ullah I, Rahim MT, Khan H, Qayyum M. Analysis of various semi-numerical schemes for magnetohydrodynamic (MHD) squeezing fluid flow in porous medium. Propulsion and Power Research. 2019;8(1):69-78.

17. Li YM, Ullah I, Ameer Ahammad N, Ullah I, Muhammad T, Asiri SA. Approximation of unsteady squeezing flow through porous space with slip effect: DJM approach. Waves in Random and Complex Media. 2022;1-5.

18. Ullah I, Rahim MT, Khan H, Qayyum M. Homotopy analysis solution for magnetohydrodynamic squeezing flow in porous medium. Advances in Mathematical Physics. 2016.

19. Qayyum M, Khan H, Rahim MT, Ullah I. Analysis of unsteady axisymmetric squeezing fluid flow with slip and no-slip boundaries using OHAM. Mathematical Problems in Engineering. 2015.

20. Ullah I, Rahim MT, Khan H. Application of Daftardar Jafari method to first grade MHD squeezing fluid flow in a porous medium with slip boundary condition. InAbstract and Applied Analysis 2014. Hindawi.

21. Abouelregal AE, Ahmad H, Yao SW, Abu-Zinadah H. Thermo-viscoelastic orthotropic constraint cylindrical cavity with variable thermal properties heated by laser pulse via the MGT thermoelasticity model. Open Physics. 2021;19(1):504-18.

22. Abouelregal AE, Ahmad H, Yao SW. Functionally graded piezoelectric medium exposed to a movable heat flow based on a heat equation with a memory-dependent derivative. Materials. 2020;13(18):3953.

23. Hussain A, Arshad M, Rehman A, Hassan A, Elagan SK, Ahmad H, Ishan A. Three-dimensional water-based magneto-hydrodynamic rotating nanofluid flow over a linear extending sheet and heat transport analysis: A numerical approach. Energies. 2021;14(16):5133.

24. Abouelregal AE, Ahmad H. A modified thermoelastic fractional heat conduction model with a single-lag and two different fractional-orders. Journal of Applied and Computational Mechanics. 2021;7(3):1676-86.

25. Anjum A, Mir NA, Farooq M, Javed M, Ahmad S, Malik MY, Alshomrani AS. Physical aspects of heat generation/absorption in the second grade fluid flow due to Riga plate: application of Cattaneo-Christov approach. Results in Physics. 2018;9:955-60.

26. Saleem S, Awais M, Nadeem S, Sandeep N, Mustafa MT. Theoretical analysis of upper-convected Maxwell fluid flow with Cattaneo–Christov heat flux model. Chinese journal of physics. 2017;55(4): 1615-1625..

27. He JH. Homotopy perturbation method for bifurcation of nonlinear problems. International Journal of Nonlinear Sciences and Numerical Simulation. 2005;6(2):207-8.

28. He JH. Approximate analytical solution for seepage flow with fractional derivatives in porous media. Computer methods in applied mechanics and engineering. 1998;167(1-2):57-68.

29. He JH. Homotopy perturbation method: a new nonlinear analytical technique. Applied Mathematics and computation. 2003;135(1):73-9.

30. He JH. Application of homotopy perturbation method to nonlinear wave equations. Chaos, Solitons & Fractals. 2005;26(3):695-700.

31. El-Shahed M. Application of He's homotopy perturbation method to Volterra's integro-differential equation. International Journal of Nonlinear Sciences and Numerical Simulation. 2005;6(2):163-8.

32. Farooq M, Rahim MT, Islam S, Siddiqui AM. Steady Poiseuille flow and heat transfer of couple stress fluids between two parallel inclined plates with variable viscosity. Journal of the Association of Arab Universities for Basic and Applied Sciences. 2013;14(1):9-18.

33. Mahian O, Mahmud S, Pop I. Analysis of first and second laws of thermodynamics between two isothermal cylinders with relative rotation in the presence of MHD flow. International Journal of Heat and Mass Transfer. 2012;55(17-18):4808-16.

Rashid Nawaz: https://orcid.org/0000-0002-4773-8446

Muhammad Farooq: https://orcid.org/0000-0003-3392-101X

Siddiq Ur Rehman: https://orcid.org/0009-0003-8596-4895

Gamal M. Ismail: https://orcid.org/0000-0002-9060-4371

Hijaz Ahmad: https://orcid.org/0000-0002-5438-5407

A NOVEL ANALYTICAL METHOD FOR THE EXACT SOLUTION
OF THE FRACTIONAL-ORDER BIOLOGICAL POPULATION MODEL

Tarig M. ELZAKI*⊙**, Mohamed Z. MOHAMED**/*****⊙

*Mathematics Department, College of Sciences and Arts, Alkamel, University of Jeddah, Saudi Arabia
**Mathematics Department, Academy of Engineering and Medical Sciences, Khartoum 12045, Sudan
***Mathematics Department, Prince Muqrin University, Almadinah Almunawwarah, Saudi Arabia

tarig.alzaki@gmail.com, M.zain198327@hotmail.com

Abstract: In this research, we develop a new analytical technique based on the Elzaki transform (ET) to solve the fractional-order biological population model (FBPM) with initial and boundary conditions (ICs and BCs). This approach can be used to locate both the closed approximate solution and the exact solution of a differential equation. The usefulness and validity of this strategy for managing the solution of FBPM are demonstrated using a few real-world scenarios. The dependability of the suggested strategy is also shown using a table and a few graphs. The approximate solutions that were achieved and the convergence analysis are shown in numerical simulations in a range of fractional orders. From the numerical simulations, it can be seen that the population density increases with increasing fractional order, whereas the population density drops with decreasing fractional order.

Key words: fractional biological population model, novel analytical method, Elzaki transform, Mittag-Leffler

1. INTRODUCTION

Although fractional derivatives have a long mathematical history, science did not use them frequently for a very long time. One possible explanation for the unpopularity of fractional derivatives is the prevalence of various non-equivalent definitions of them [1–3]. Furthermore, due to their non-locality, fractional derivatives lack a precise geometrical interpretation [1]. Over the past 10 years, however, fractional calculus has begun to attract the attention of mathematicians and engineers much more. It was discovered that fractional derivatives can effectively imitate a variety of applications, notably interdisciplinary ones. Fractional derivatives can be used to explain a variety of phenomena, such as the non-linear oscillation of earthquakes [3], Kilbas et al. [1] provide an overview of a few fractional derivative applications in continuous mechanics and statistical mechanics.

Many authors have researched the analytical findings on the existence and distinctiveness of fractional differential equation [FDE] solutions. Various techniques, including Adomian decomposition (ADM), Homotopy analysis [5], and many more, have been used in recent years to solve FDEs, FPDEs, and dynamic systems incorporating fractional derivatives. Fractional operators can be used to effectively represent phenomena with the memory effect since they are non-local. We stress that a particular fractional operator can change a PDE from a local to a non-local one by substituting it for the classical derivative with respect to time.

In this essay, the FBPM will be resolved using a novel approach called the Elzaki transform (ET) approach. ET and its variations are used to tackle boundary value problems. The recommended approach [10, 11, 21, 22, 24] presents the solution in a finite series form that is straightforward to compute, but the real strategy offers greater precision because different starting approximations are employed in any iterations. This class of equations

including linear fractional differential equations did not have an analytical solution method prior to the 17th century. Linear and non-linear population problems were handled in [8, 9] using the VIM and HPM. Akinfe and Loyinmi [13] have examined other earlier research attempts on the current fractional biological population model and its applications in quantum physics, optics, fluid modelling employing the ET and a solitary wave solution to the generalised Burgers-Fisher's equation using an improved differential transform method in [14–17, 28]. Additionally, the fractional order model is examined in [18–20] with regard to Esmehan Ucar et al.

The Caputo fractional derivative was used for this study because it enables the formulation of the physical problems to include conventional initial and boundary conditions (BCs). Provide a few other crucial properties of fractional derivatives. A few instances of the identified difficulties are addressed using the general description of the suggested solution. Finding analytical solutions to FBPM using initial conditions (ICs) and BCs is fairly difficult. The current study uses a relatively simple and straightforward methodology to obtain closed-form analytical answers for the FBPM. We'll talk about the fractional biological population model in this article, and this strategy is a potent method for resolving the functional equations that arise from modelling various systems analytically.

The plan of our paper is as follows: Brief definitions of fractional calculus are given in Section 1. Some theorems of the ET are given in Section 2. The novel analytical method is presented in Section 3. The convergence analysis is presented in Section 4. In Section 5, three numerical examples are given to illustrate the applicability of the considered method. Numeric results are presented in Section 6. Section 7 is devoted to the conclusions of the work.

The generalised time-fractional non-linear biological popula-

tion equation we suggest in this paper is as follows:

$$D_t^\alpha \Phi(x,y,t) = \frac{\partial^2}{\partial x^2}(\Phi^2) + \frac{\partial^2}{\partial y^2}(\Phi^2) + f(\Phi),$$
$$0 < \alpha \le 1, \quad t > 0,$$

(1)

Given ICs and BCs, and based on Verhulst and Malthusian law, we explore a more generic version of,

$$f(\Phi) = h\,\Phi^a(1 - r\Phi^b), \quad h, a, b, r \in \mathbb{R},$$

They switch to Verhulst and Malthusian laws when choosing exceptional values.

Definition 1: The following definition for the Riemann–Liouville (R–L) [23], fractional integral (FI) operator, of the order

$\alpha > 0$, of $f \in C_\mu$, $\mu \ge -1$, is given:

$$J^\alpha f(\eta) = \frac{1}{\Gamma(\alpha)} \int_0^\eta (\eta - v)^{\alpha-1} f(v)\, dv, \ \alpha > 0. J^0 f(\eta) = f(\eta).$$

Some properties of J^α, for, $f^n \in C_\mu, n \in N, \alpha, \beta \ge 0$ and $\gamma \ge -1$:

(a) $\quad J^\alpha J^\beta f(\eta) = J^{\alpha+\beta} f(\eta)$

(b) $\quad J^\alpha \eta^\gamma = \frac{\Gamma(\gamma+1)}{\Gamma(\gamma+\alpha+1)} \eta^{\alpha+\gamma}$

Definition 2: According to Caputo, the fractional derivative of $f(\eta)$, is: $D^\alpha f(\eta) = J^{m-\alpha} D^m f(\eta)$. For $m-1 < \alpha \le m$, $m \in N, \eta > 0$, and $f \in C_{-1}^m$.

Caputo's fractional derivative (CFD) computes an ordinary derivative first, then a FI to determine the right order of a fractional derivative. The FI operator of RL and the integer order integration are both linear operations:

$$J^\alpha \left(\sum_{i=1}^n c_i f_i(\eta) \right) = \sum_{i=1}^n c_i J^\alpha f_i(\eta),$$

where, $\{c_i\}_{i=1}^n$ are constants.

Fractional derivatives are interpreted as having a Caputo, meaning, in the current investigation, provides the reason for using the Caputo definition.

2. ELZAKI TRANSFORM

Here is a short explanation of the modified Sumudu transform, also known as the ET of the function $\Phi(\eta)$,

$$E[\Phi(\eta)] = p \int_0^\infty \Phi(\eta)\, e^{-\frac{\eta}{p}} d\eta = T(p), \quad \eta > 0,$$

where, p is a complex value.

Tarig M. Elzaki has shown in [10, 11, 5] that PDEs, ODEs, systems of PDEs, and Euler-Bernoulli Beam's can all be solved using the modified Sumudu transform, or ET. When Sumudu and Laplace transforms are unsuccessful in solving DEs with variable coefficients, ET can be effectively used [12].

Theorem 1: [2] If $\Phi = \Phi(x,y,t)$, then the partial derivatives are transformed by ET as follows:

$$E\left[\frac{\partial \Phi}{\partial t}\right] = \frac{1}{p} T(x,y,p) - p\, \Phi(x,y,0),$$

$$E\left[\frac{\partial^2 \Phi}{\partial t^2}\right] = \frac{1}{p^2} T(x,y,p) - \Phi(x,y,0) - p\frac{\partial \Phi(x,y,0)}{\partial t},$$

$$E\left[\frac{\partial \Phi}{\partial x}\right] = \frac{d}{dx}[T(x,y,p)], \qquad E\left[\frac{\partial^2 \Phi}{\partial x^2}\right] = \frac{d^2}{dx^2}[T(x,y,p)],$$

$$E\left[\frac{\partial \Phi}{\partial y}\right] = \frac{d}{dy}[T(x,y,p)], \qquad E\left[\frac{\partial^2 \Phi}{\partial y^2}\right] = \frac{d^2}{dy^2}[T(x,y,p)].$$

ET of some functions:

$\Phi(\eta)$	$E[\Phi(\eta)] = T(p)$
1	p^2
η	p^3
η^n	$n!\ p^{n+2}$
$e^{a\eta}$	$\dfrac{p^2}{1-ap}$
$\sin a\eta$	$\dfrac{ap^3}{1+a^2p^2}$
$\cos a\eta$	$\dfrac{p^2}{1+a^2p^2}$

Here, we show some lemmas that can be applied to extract the function $\Phi(\eta)$, from its ET.

Lemma 1: ET of R–L FI operator of order $\alpha > 0$, is represented as:

$$E[J^\alpha \Phi(\eta)] = p^\alpha T(p).$$

Proof: We begin by:

$$E[J^\alpha \Phi(\eta)] = E\left[\frac{1}{\Gamma(\alpha)} \int_0^\eta (\eta - \alpha)^{\alpha-1} \Phi(\eta)\, d\eta\right]$$
$$= \frac{1}{\Gamma(\alpha)}\ \frac{1}{p}\ T(p)\ G(p) = p^\alpha T(p),$$

where

$$G(p) = E[\eta^{\alpha-1}] = p^{\alpha+1} \Gamma(\alpha).$$

Lemma 2: ET of CFD for $\alpha > 0$, $m-1 < \alpha \le m$, $m \in N$, is;

$$E[D_t^\alpha \Phi] = p^{m-\alpha}\left[\begin{array}{l} \frac{T(x,y,p)}{p^m} - \frac{\Phi(x,y,0)}{p^{m-2}} - \frac{\frac{\partial \Phi(x,y,0)}{\partial t}}{p^{m-3}} \\ - \cdots - p\frac{\partial^{m-1}\Phi(x,y,0)}{\partial t^{m-1}} \end{array} \right],$$

or

$$E[D_t^\alpha \Phi] = \frac{1}{p^\alpha} E[\Phi] - \sum_{k=0}^{m-1} \frac{\partial^k \Phi(x,y,0)}{\partial t^k} p^{2-\alpha+k}, m-1 < \alpha \le m,$$

The following is the definition of the normal and generalized Mittag-Leffler functions:

$$\in_\alpha (\eta) = \sum_{n=0}^\infty \frac{\eta^n}{\Gamma(n\alpha+1)}, \in_{\alpha,\beta} (\eta) = \sum_{n=0}^\infty \frac{\eta^n}{\Gamma(n\alpha+\beta)}.$$

Lemma 3: If, $\alpha, \beta > 0$, $a \in C$ and $\frac{1}{p^\alpha} > |a|$, the formula for inverse ET is as follows:

$$E^{-1}\left[\frac{p^{\beta+1}}{1+a\,p^\alpha}\right] = \eta^{\beta-1} \in_{\alpha,\beta} (-a\,\eta^\alpha).$$

Proof:

$$\frac{p^{\beta+1}}{1+a\,p^\alpha} = p^{\beta+1} \frac{1}{1+a\,p^\alpha} = p^{\beta+1} \sum_{n=0}^\infty (-a)^n (p^\alpha)^n$$
$$= \sum_{n=0}^\infty (-a)^n\, p^{n\alpha+\beta+1},$$

Tarig M. Elzaki, Mohamed Z. Mohamed
A Novel Analytical Method for Tthe Exact Solution of the Fractional-Order Biological Population Model

DOI 10.2478/ama-2024-0059

Then:

$$E^{-1}\left[\frac{p^{\beta+1}}{1+ap^{\alpha}}\right] = E^{-1}\left[\sum_{n=0}^{\infty}(-a)^n \, p^{n\alpha+\beta+1}\right]$$

$$= \sum_{n=0}^{\infty}\frac{(-a)^n \eta^{n\alpha+\beta-1}}{\Gamma(n\alpha+\beta)} = A^{\beta-1}\sum_{n=0}^{\infty}\frac{(-a\eta^{\alpha})^n}{\Gamma(n\alpha+\beta)} = \eta^{\beta-1}\in_{\alpha,\beta}(-a\eta^{\alpha}).$$

3. THE NOVEL ANALYTICAL METHOD

We explain the basic tenets of the suggested approach in this section. Let's look at the fractional non-linear non-homogeneous PDE,

$$D_t^{\alpha}\Phi + R[\Phi] + N[\Phi] = g(x,y,t),$$
$$0 < \alpha \leq 2, \; 0 \leq x \leq d, \; 0 \leq y \leq j, \; t \geq 0, \; d,j \in \mathbb{R}, \quad (2)$$

with the ICs, and BCs,

$$\Phi(x,y,0) = h_1(x,y), \quad \Phi_t(x,y,0) = h_2(x,y),$$
$$\Phi(0,y,t) = k_1(y,t), \quad \Phi(d,y,t) = k_2(y,t),$$
$$\Phi(x,0,t) = k_3(x,t), \quad \Phi(x,j,t) = k_4(x,t), \quad (3)$$

R, N are linear and non-linear operators, and g, inhomogeneous terms.

Considering that, Φ_n^* the new method was employed to calculate the new solution,

$$\Phi_n^* = \Phi_n + (d-x)(k_1 - \Phi_n(0,y,t)) + x(k_2 - \Phi_n(d,y,t)) +$$
$$(j-y)(k_3 - \Phi_n(x,0,t)) + y(k_4 - \Phi_n(x,j,t)), \quad (4)$$

where $n = 0, 1, \cdots$.

It is obvious that Φ_n^* will meet both the ICs' and BCs' standards. We can resolve Eq. (2) by using the ET to deduce Eq. (3).

$$E[D_t^{\alpha}\Phi] + E\{R[\Phi] + N[\Phi]\} = E[g],$$
$$E(\Phi) = p^2 h_1(x,t) + p^3 h_2(x,t) - p^{\alpha}E\{R[\Phi] + N[\Phi] - g\}. \quad (5)$$

It is believed that the solution to Eq. (2) has the following series form:

$$\Phi = \sum_{n=0}^{\infty}\Phi_n, \quad (6)$$

As a result of applying Eq. (6) and the inverse of ET to Eq. (5), we can now determine,

$$\sum_{n=0}^{\infty}\Phi = G - E^{-1}\{p^{\alpha}E[R[\Phi] + N[\Phi]]\}, \quad (7)$$

When, G an expression is made from a source word and the necessary ICs,

This method relies on how we choose the initial iteration Φ_0 that offers the exact solution in a constrained number of steps.

To discover the solution iteratively, apply the relations listed below.

$$\Phi_{n+1} = E^{-1}\{p^{\alpha}E[R[\Phi_n^*] + N[\Phi_n^*]]\},$$

$$\Phi_0 = G, \quad (8)$$

From Eqs. (8) and (4), we can infer that:

$$\Phi_0, \; \Phi_1, \; \Phi_2, \dots,$$

The solution can then be inferred from Eq. (6).

We demonstrate that FBPMs that are under the ICs, BCs, and ET may be resolved using the suggested strategy.

4. CONVERGENCE ANALYSIS CONSIDERED PROBLEM

This section examines the FBPM's convergence for the specified problem, as stated in Eq. (1). To do this, we apply the operator's Eq. (1) as:

$$T(\Phi) = D_t\Phi = (D_x^2 + D_y^2)\Phi^2 + h\Phi^a - rh\Phi^{a+b},$$

Let $H \in L^2[T], \; \forall\Phi \in H,$ [25] where

$$H \in L_{\Phi}^2[(m,n) \times [0,T]], \text{ such that,}$$

$$\Phi := [(m,n) \times [0,T]] \to R^3,$$

with $m << 0$ and $B = [(m,n) \times [0,T]],$

where $\|\Phi\|_H^2 = \int_B \Phi^2 dxdydt$, then $E_t^{-1}\{E_t[\Phi(x,y,t)]\} < \infty$

We now assume the following in order to demonstrate T, to be semi-continuous [25]:

Assumption:

H_1 for $\sigma > 0$, exist a constant $\beta > 0$, and $\forall\Phi_1, \Phi_2 \in H$, with $k\|\Phi_1 + \Phi_2\| \leq \sigma,$

we obtain

$$\|T(\Phi_1) - T(\Phi_2)\| \leq \beta\|\Phi_1 - \Phi_2\|, \forall\Phi_1, \Phi_2 \in H.$$

Theorem 2: (Convergence Condition)[26]. Without initial and BCs convergent to a specific solution, the problem under consideration is examined in Eq. (1).

Making use of the above Assumption for operator $T(\Phi)$, in Eq. (1), to obtain,

$$T(\Phi_1) - T(\Phi_2) = (D_x^2 + D_y^2)\Phi_1^2 + h\Phi_1^a - rh\Phi_1^{a+b}$$
$$-\{(D_x^2 + D_y^2)\Phi_2^2 + h\Phi_2^a - rh\Phi_2^{a+b}\} = D_x^2(\Phi_1^2 - \Phi_2^2)$$
$$+D_y^2(\Phi_1^2 - \Phi_2^2) + h(\Phi_1^a - \Phi_2^a) - rh(\Phi_1^{a+b} - \Phi_2^{a+b}),$$

By using the norm, we can get:

$$\|T(\Phi_1) - T(\Phi_2)\| = \|D_x^2(\Phi_1 - \Phi_2)(\Phi_1 + \Phi_2)\|$$
$$+\|D_y^2(\Phi_1 - \Phi_2)(\Phi_1 + \Phi_2)\|$$
$$+h\|(\Phi_1^a - \Phi_2^a)\| - rh\|(\Phi_1^{a+b} - \Phi_2^{a+b})\|,$$

By utilizing the conditions on the operators D_x^2, D_y^2, in H, $\exists \varsigma_1, \varsigma_2 > 0$, and if $a = b = 1$, we can define,

$$D_x^2(\Phi_1 - \Phi_2)(\Phi_1 + \Phi_2) \leq \varsigma_1\|\Phi_1 - \Phi_2\|,$$
$$D_y^2(\Phi_1 - \Phi_2)(\Phi_1 + \Phi_2) \leq \varsigma_2\|\Phi_1 - \Phi_2\|,$$

Therefore,

$$\|T(\Phi_1) - T(\Phi_2)\| \leq \varsigma_1\|\Phi_1 - \Phi_2\| + \varsigma_2\|\Phi_1 - \Phi_2\|$$
$$+h\|\Phi_1 - \Phi_2\| - r\sigma\|\Phi_1 - \Phi_2\| \Rightarrow$$
$$\|T(\Phi_1) - T(\Phi_2)\| \leq (\varsigma_1 + \varsigma_2 + h - r\sigma)\|\Phi_1 - \Phi_2\|,$$

Let,

$$d = \varsigma_1 + \varsigma_2 + h - r\sigma > 0, \text{ then we can write,}$$

$$\|T(\Phi_1) - T(\Phi_2)\| \leq d\|\Phi_1 - \Phi_2\|.$$

Thus, the assumption is met. As a result, the suggested approach converges.

5. ILLUSTRATIVE EXAMPLES

In this section, we will use three numerical examples to illustrate the efficiency and dependability of the method.

Example 1: Take a look at FBPM in one dimension,

$$D_t^\alpha \Phi(x,t) = \frac{\partial^2}{\partial x^2}(\Phi^2) + \Phi\left(1 - \frac{4}{9}\Phi\right), 0 < \alpha \le 1, \tag{9}$$
$$0 \le x \le d, \quad t \ge 0, \quad d,j \in \mathbb{R},$$

With the IC and BCs,

$$\Phi(x,0) = e^{\frac{x}{3}}, \quad \Phi(0,t) = \in_\alpha(t^\alpha), \quad \Phi(d,t) = e^{\frac{d}{3}} \in_\alpha(t^\alpha), \tag{10}$$

By combining the IC, ET, and Eq. (9), the following result is produced:

$$\frac{1}{p^\alpha}E[\Phi(x,t)] - \Phi(x,0)p^{2-\alpha} - E[\Phi(x,t)] =$$
$$E\left[\frac{\partial^2}{\partial x^2}(\Phi^2) - \frac{4}{9}\Phi^2\right],$$
$$\Rightarrow E[\Phi(x,t)] = \frac{p^2 e^{\frac{x}{3}}}{1-p^\alpha} + \frac{p^\alpha}{1-p^\alpha}E\left[\frac{\partial^2}{\partial x^2}(\Phi^2) - \frac{4}{9}\Phi^2\right],$$

Inverse ET suggests that:

$$\Phi((x,t)) = e^{\frac{x}{3}} \in_\alpha(t^\alpha) + E^{-1}\left\{\frac{p^\alpha}{1-p^\alpha}E\left[\frac{\partial^2}{\partial x^2}(\Phi^2) - \frac{4}{9}\Phi^2\right]\right\},$$

The following diagram illustrates the iteration formula using an initial approximation.

$$\Phi_{n+1}(x,t) = E^{-1}\left\{\frac{p^\alpha}{1-p^\alpha}E\left[\frac{\partial^2}{\partial x^2}(\Phi_n^*)^2 - \frac{4}{9}(\Phi_n^*)^2\right]\right\}, \tag{11}$$

with, $\Phi_0(x,t) = e^{\frac{x}{3}} \in_\alpha(t^\alpha)$,

Utilize the BCs in Eq. (4) and $n = 0$, to ascertain:

$$\Phi_0^*(x,t) = \Phi_0(x,t) + (d-x)(\Phi(0,t) - \Phi_0(0,t))$$
$$+x\big(\Phi(d,t) - \Phi_0(d,t)\big) = e^{\frac{x}{3}} \in_\alpha(t^\alpha),$$

Eq. (11), give:

$$\Phi_1(x,t) = E^{-1}\left\{\frac{p^\alpha}{1-p^\alpha}E\left[\frac{\partial^2}{\partial x^2}(\Phi_0^*)^2 - \frac{4}{9}(\Phi_0^*)^2\right]\right\} = 0,$$

Then,

$$\Phi_1 = 0, \quad \Phi_2 = 0, \quad \Phi_3 = 0, \quad \dots,$$

Using Eq. (6), to find the solution of Eq. (9),

$$\Phi(x,t) = \sum_{n=0}^\infty \Phi_n(x,t) = e^{\frac{x}{3}} \in_\alpha(t^\alpha), \text{ if } \alpha = 1, \text{ then: } \Phi(x,t) = e^{\frac{x}{3}+t}.$$

Example 2: Consider the FBPM in two-dimensional,

$$D_t^\alpha \Phi = \frac{\partial^2}{\partial x^2}(\Phi^2) + \frac{\partial^2}{\partial y^2}(\Phi^2) + \Phi, \quad 0 < \alpha \le 1, 0 \le x,y \le \pi, \ t \ge 0, \tag{12}$$
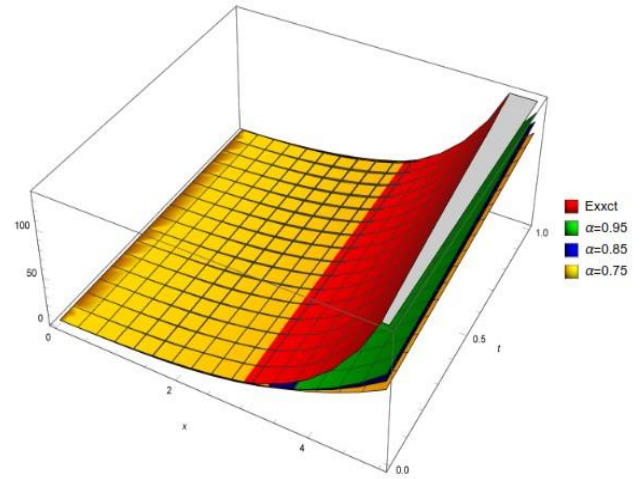
With the IC and BCs,

$$\Phi(x,y,0) = \sqrt{\sin x \sinh y}, \quad \Phi(0,y,t) = 0, \quad \Phi(\pi,y,t) = 0,$$
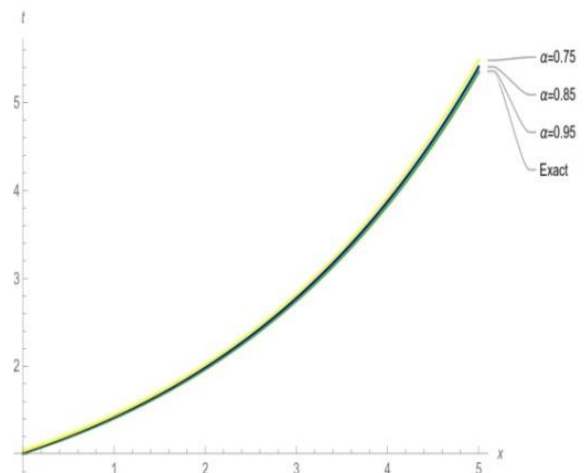$$\Phi(x,0,t) = 0, \quad \Phi(x,\pi,t) = \sqrt{\sin x \sinh \pi}. \in_\alpha(t^\alpha), \tag{13}$$

By combining the IC, ET, and Eq. (12), the following result is produced:

$$\frac{1}{p^\alpha}E[\Phi] - \Phi(x,y,0)p^{2-\alpha} - E[\Phi(x,y,t)] =$$
$$E\left[\frac{\partial^2}{\partial x^2}(\Phi^2) + \frac{\partial^2}{\partial y^2}(\Phi^2)\right],$$
$$\Rightarrow E[\Phi] = \frac{p^2}{1-p^\alpha}\sqrt{\sin x \sinh y} + \frac{p^\alpha}{1-p^\alpha}E\left[\frac{\partial^2}{\partial A^2}(\Phi^2) + \frac{\partial^2}{\partial y^2}(\Phi^2)\right],$$

(1)

(2)

Fig. 1,2. 3D and 2D graph representations of the exact and approximate solutions for example 1, when $t = 0.01, \alpha = 0.95, 0.85, 0.75$

Tab. 1. The numirical outcome of example 1 is determinted by comparing the exact and approximate solutions for the two-terms approximation

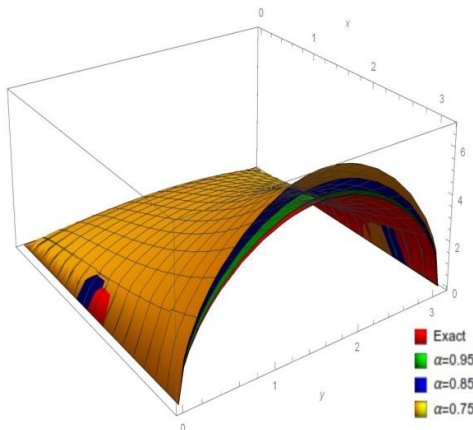|	t	x	$\alpha = 0,75$	$\alpha = 0,85$	$\alpha = 0,95$	Exact
		0	1.03517	1.02136	1.01293	1.01005
		0.1	1.07026	1.05598	1.04727	1.04429
		0.2	1.10654	1.09177	1.08277	1.07968
		0.3	1.14404	1.12878	1.11947	1.11628
		0.4	1.18282	1.16704	1.15741	1.15411
$\Phi(x,t)$	0.01	0.5	1.22291	1.20659	1.19664	1.19323
		0.6	1.26436	1.24749	1.2372	1.23368
		0.7	1.30722	1.28978	1.27914	1.27549
		0.8	1.35153	1.33349	1.32249	1.31873
		0.9	1.39734	1.37869	1.36732	1.36343
		1	1.4447	1.42542	1.41366	1.40964

Inverse ET suggests that:

$$\Phi = \in_\alpha (t^\alpha)\sqrt{sinxsinhy} + E^{-1}\left\{\frac{p^\alpha}{1-p^\alpha}E\left[\frac{\partial^2}{\partial x^2}(\Phi^2) + \frac{\partial^2}{\partial y^2}(\Phi^2)\right]\right\},$$

The following diagram illustrates the iteration formula using an initial approximation

$$\Phi_{n+1} = E^{-1}\left\{\frac{p^\alpha}{1-p^\alpha}E\left[\frac{\partial^2}{\partial x^2}(\Phi_n^*)^2 + \frac{\partial^2}{\partial y^2}(\Phi_n^*)^2\right]\right\}, \tag{14}$$

With, $\Phi_0 = \in_\alpha (t^\alpha)\sqrt{sinxsinhy}$, use the BCs in Eq. (4) and $n = 0$ to find:

$$\Phi_0^* = \Phi_0 + (\pi - x)(\Phi(0,y,t) - \Phi_0(0,y,t))$$
$$+x\big(\Phi(\pi,y,t) - \Phi_0(\pi,y,t)\big)$$
$$+(\pi - y)(\Phi(x,0,t) - \Phi_0(x,0,t)) + y(\Phi(x,\pi,t)$$
$$-\Phi_0(x,\pi,t)) = \in_\alpha (t^\alpha)\sqrt{sinxsinhy},$$

From Eq. (14), we get:

$$\Phi_1 = E^{-1}\left\{\frac{p^\alpha}{1-p^\alpha}E\left[\frac{\partial^2}{\partial x^2}(\Phi_0^*)^2 + \frac{\partial^2}{\partial y^2}(\Phi_0^*)^2\right]\right\} = 0,$$

Then, $\Phi_1 = 0$, $\Phi_2 = 0$, $\Phi_3 = 0$, ... ,

The solution to Eq. (12) can then be found by applying Eq. (6),
$\Phi = \sum_{n=0}^{\infty}\Phi_n = \in_\alpha (t^\alpha)\sqrt{sinxsinhy},$

when, $\alpha = 1$, then $\Phi = e^t\sqrt{sinxsinhy}$.

(3)

(4)

Fig. 3,4: 3D and 2D graph representations of the exact and approximate solutions for example 2, when $t = 0.5, \alpha = 0.95, 0.85, 0.75$

Tab. 2. The numerical outcames of example 2 is determinted by comparing the exact and approximate solutions for the two-terms approximation.

	t	y	x	$\alpha = 0,75$	$\alpha = 0,85$	$\alpha = 0,95$	Exact
			0	0.	0.	0.	0.
			$\pi/6$	4.85575	4.43446	4.1024	3.96186
			$\pi/3$	6.39052	5.83608	5.39906	5.21411
$\Phi(y,t)$	0,5	π	$\pi/2$	6.86706	6.27127	5.80167	5.60292
			$2\pi/3$	6.39052	5.83608	5.39906	5.21411
			$5\pi/6$	4.85575	4.43446	4.1024	3.96186
			π	0.	0.	0.	0.

Example 3: Think of the FBPM in two dimensions,

$$D_t^\alpha\Phi(x,y,t) = \frac{\partial^2}{\partial x^2}(\Phi^2) + \frac{\partial^2}{\partial y^2}(\Phi^2) + k\Phi, \ 0 < \alpha \le 1,$$
$$0 \le x \le d, 0 \le y \le j, \ t \ge 0, \ k, d, j \in \mathbb{R}, \tag{15}$$

With the IC and BCs,

$$\Phi(x,y,0) = \sqrt{xy}, \Phi(0,y,t) = 0,$$
$$\Phi(d,y,t) = \sqrt{dy} \in_\alpha ((kt)^\alpha),$$
$$\Phi(x,0,t) = 0, \ \Phi(x,j,t) = \sqrt{jx} \in_\alpha ((kt)^\alpha), \tag{16}$$

The recurrence relationship will continue to exist in the same manner, using an initial approximation, as shown in the following:

$$\Phi_{n+1} = E^{-1}\left\{\frac{p^\alpha}{1-p^\alpha}E\left[\frac{\partial^2}{\partial x^2}(\Phi_n^*)^2 + \frac{\partial^2}{\partial y^2}(\Phi_n^*)^2\right]\right\}, \tag{17}$$

With, $\Phi_0 = \in_\alpha ((kt)^\alpha)\sqrt{xy}$, use the BCs in Eq. (4) and but, $n = 0$ to get:

$$\Phi_0^* = \Phi_0 + (d - x)(\Phi(0,y,t) - \Phi_0(0,y,t))$$

$$+x(\Phi(d,y,t) - \Phi_0(d,y,t)) +$$
$$(j - y)(\Phi(x,0,t) - \Phi_0(x,0,t)) + y(\Phi(x,j,t)$$
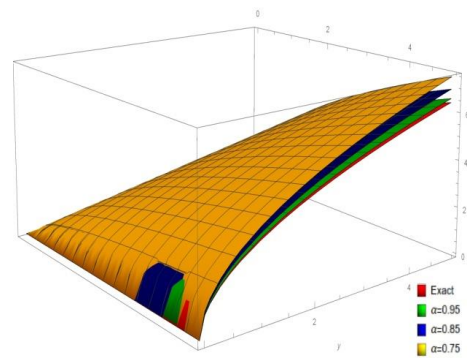$$-\Phi_0(x,j,t)) = \in_\alpha ((kt)^\alpha)\sqrt{xy},$$

The following is taken from Eq. (17):

$$\Phi_1 = E^{-1}\left\{\frac{p^\alpha}{1-p^\alpha}E[0]\right\} = 0$$
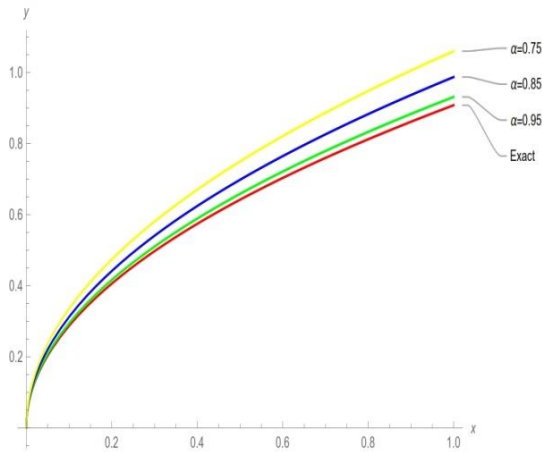
then, $\Phi_1 = 0$, $\Phi_2 = 0$, $\Phi_3 = 0$, ... ,

The solution to Eq. (15) can therefore be found using Eq. (6).

$$\Phi = \sum_{n=0}^{\infty}\Phi_n = \in_\alpha \left((kt)^\alpha\right)\sqrt{xy}, \text{ if, } \alpha = 1, \text{ then: } \Phi = e^{kt}\sqrt{xy}.$$

(5)

(6)

Fig. 5,6. 3D and 2D graph representations of the exact and approximate solutions for example 3, when $t = 0.5, \alpha = 0.95, 0.85, 0.75$.

Tab. 3. The numerical outcames of example 3 is determined by comparing the exact and approximate solutions for the two-terms approximation.

	t	y	x	$\alpha = 0,75$	$\alpha = 0,85$	$\alpha = 0,95$	Exact
			0	1.03517	1.02136	1.01293	1.01005
			0.1	1.07026	1.05598	1.04727	1.04429
			0.2	1.10654	1.09177	1.08277	1.07968
			0.3	1.14404	1.12878	1.11947	1.11628
			0.4	1.18282	1.16704	1.15741	1.15411
$\Phi(y,t)$	0.5	0.5	0.5	1.22291	1.20659	1.19664	1.19323
			0.6	1.26436	1.24749	1.2372	1.23368
			0.7	1.30722	1.28978	1.27914	1.27549
			0.8	1.35153	1.33349	1.32249	1.31873
			0.9	1.39734	1.37869	1.36732	1.36343
			1	1.4447	1.42542	1.41366	1.40964

6. NUMERICAL RESULTS

This work resolves the FBPM with ICs and BCs using a novel analytical method based on ET. Results obtained using earlier methods are not equivalent to those obtained with current methods. Multiple parameter values are given in Eqs. (9), (12), and (15) to offer a variety of solutions. A variety of solutions can be produced by permitting the arbitrary parameters to have different values in the solutions. The responses gathered are categorized. Visual representations in 2D and 3D are also developed. The following information can be used to describe these plots: Figs. 1, 2, and 3 depict lone waves in different arrangements. Fig. 1 was produced for the values in Eq. (9), showing: at $\alpha = 0.95, 0.85, 0.75$, $t = 0.01$. This mixture falls within the periodic category. Fig. 2 was created using the data $\alpha = 0.95$, $0.85, 0.75$, $t = 0.5$, in Eq. (12). Fig. 3 was created using $\alpha = 0.95, 0.85, 0.75, t = 0.5$, in Eq. (15).

The exact solution, which was determined by comparing the values of the exact and approximate solutions of FBPM discovered in this issue for various values of the variables, $0 < x \le 1, t = 0.01$, is provided in Tab.1. Tab.2 contains the approximate solutions, which were ascertained by comparing the values of the exact solutions of FBPM for various values of the variables,

$0 < x \le \pi, y = \pi, t = 0.5.$ $0 < x \le 1, y = 0.5, t = 0.5.$ We were able to determine the exact solution, which is presented in Tab. 3. The suggested approaches were found to be both workable and efficient. Using the Wolfram Mathematica program, the simulations were run and the outcomes were examined.

7. CONCLUSION

In this paper, analytical evaluations of non-linear fractional biological population models with ICs and BCs are carried out. The Riemann–Liouville FI operator generates partially specified fractional derivatives. For the speedy and effective resolution of numerous difficulties, a novel approach based on ET is proposed. Three examples are offered to show the value of the suggested approach. The solutions might be handled in a very simple way. The outstanding ability of the approach to solve non-linear FBPM utilising ICs and BCs allows it to be modified to address a variety of boundary value problems. In addition, 2D and 3D graphs were employed to demonstrate how the suggested approach contributed to the outcomes.

Figures and tables demonstrate that the shapes of the solutions discovered using the suggested method is similar to those of the precise solution when the same parameters were chosen. The proposed method can also be extended to solve additional FPDEs that arise in applied research, according to the ease with which it can be put into practice. In the future, we propose adapting the considered novel approach with the ET scheme for the analysis of nonlinear partial differential equations and certain advanced integral-related problems in fluid dynamics and elasticity and investigating the proposed method's stability and error analysis in forthcoming articles. Finally, we affirm that, subject to ICs and BCs, the proposed technique is valid and applicable to all non-linear FPDEs.

REFERENCES

1. Kilbas AA, Srivastava HM, Trujillo JJ. Theory and Applications of Fractional Differential Equations. Elsevier. San Diego. 2006.
2. Momani S, Shawagfeh NT. Decomposition method for solving fractional Riccati differential equations. Appl.Math. Comput. 2006; 182:1083-1092.
3. Gejji VD, Jafari H. Solving a multi-order fractional differential equation, Appl. Math. Comput. 2007;189:541-548.
4. Hilal EMA, Elzaki TM. Solution of Nonlinear Partial Differential Equations by New Laplace Variational Iteration Method, Journal of Function Spaces. 2014; 1-5. http://dx.doi.org/10.1155/2014/790714.
5. Elzaki TM, Biazar J. Homotopy Perturbation Method and Elzaki Transform for Solving System of Nonlinear Partial Differential Equations. World Applied Sciences Journal.
DOI: 10.5829/idosi.wasj.2013.24.07.1041
6. Elzaki TM, Ishag AA. Modified Laplace Transform and Ordinary Differential Equations with Variable Coefficients, World Engineering & Applied Sciences Journal. 2019; 10 (3): 79-84.
DOI:10.5829/idosi.weasj.2019.79.84
7. Srivastava VK, Awasthi MK, Kumar S. Analytical approximations of two and three dimensional time fractional telegraphic equation by reduced differential transform method. Egypt J Basic Appl Sci.
https:// dx.doi.org/10.1016/j.ejbas.2014.01.002
8. Shakeri F, Dehghan M. Numerical solution of a biological population model using He's variational iteration method. Comput Math Appl.2006;54:1197-209.
9. Roul P. Application of homotopy perturbation method to biological population model. Appl Appl Math. 2010;10:1369-78.

10. Duz M, Elzaki TM. Solution Of Constant Coeffients Partial Derivative Equations With Elzaki Transform Method. Twms J. App. Eng. Math. 2019;9(3):563-570.

11. Ike C, Elzaki TM. Elzaki Transform Method for Natural Frequency Analysis of Euler-Bernoulli Beams. Engineering and Technology Journal. 2023; 1-12. DOI: 10.30684/etj.2023.140211.1456

12. Elzaki TM, Ishag AA. Modified Laplace Transform and Ordinary Differential Equations with Variable Coefficients. World Engineering & Applied Sciences Journal. 2019;10(3):79-84. DOI: 10.5829/idosi.weasj.2019.79.84

13. Akinfe KT, Loyinmi AC. The implementation of an improved differential transform scheme on the Schrodinger equation governing wave-particle duality in quantum physics and optics. Results in Physics. 2022.

14. Akinfe KT. A reliable analytic technique for the modified prototypical Kelvin–Voigt viscoelastic fluid model by means of the hyperbolic tangent function, Partial Differential Equations in Applied Mathematics 2023;7:100523.

15. Akinfe KT, Loyinmi AC. An improved differential transform scheme implementation on the generalized Allen–Cahn equation governing oil pollution dynamics in oceanography, Partial Differential Equations in Applied Mathematics. 2022;6:100416.

16. Akinfe KT, Loyinmi AC. Exact solutions to the family of Fisher's reaction-diffusion equation using Elzaki homotopy transformation perturbation method. Wiley. 2019. DOI:10.1002/eng2.12084

17. Akinfe KT, Loyinmi AC. An algorithm for solving the Burgers–Huxley equation using the Elzaki transform. SN Appl. Sci. 2020;2(7). https://doi.org/10.1007/s42452-019-1653-3

18. Uçar E, Özdemir N. A fractional model of cancer-immune system with Caputo and Caputo–Fabrizio derivatives. Eur. Phys. J. Plus 2021;136(43). https://doi.org/10.1140/epjp/s13360-020-00966-9

19. Ucar E, Özdemir N, Altun E. Fractional order model of immune cells influenced by cancer cells Math. Model. Nat. Phenom. 2019; 14(3):308. DOI: https://doi.org/10.1051/mmnp/2019002

20. Ozdemir N, Uçar S, Eroglu BBI. Dynamical Analysis of Fractional Order Model for Computer Virus Propagation with Kill Signals. International Journal of Nonlinear Sciences and Numerical Simulation. 2019.

21. Hassaballa AA, Elzaki TM. Applications of the Improved G /G Expansion Method for Solve Burgers-Fisher Equation. Journal of Computational and Theoretical Nanoscience, 2017;14: 4664–4668.

22. Elzaki TM, Elzaki SM, Elnour EA. Applications of New Transform Elzaki Transform to Mechanics, Electrical Circuits and Beams Problems. Global Journal of Mathematical Sciences: Theory and Practical. 2012;4(1):25-34.

23. Mohamed M, hamza A, Elzaki TM, Algolam M, Elhussein S. Solution of Fractional Heat-Like and Fractional Wave-Like Equation by Using Modern Strategy. Acta Mechanica et Automatica 2023;17(3):372-380. https://doi.org/10.2478/ama-2023-0042.

24. Elzaki TM, Shams EA, Areshi M, Chamekh M. Fractional partial differential equations and novel double integral transform, Journal of King Saud University. 2022;34(3):101832.

25. Gadain H.E. Application of double Laplace decomposition method for solving singular one dimensional a system of hyperbolic equations. J. Nonlinear Sci. Appl. 2017;10:111–121.

26. Kaya D, Inan IE. A convergence analysis of the ADM and an application. Appl. Math. Comput. 2005;161:1015–1025.

27. Rahman MU, Althobaiti A, Riaz MB, Al-Duais FS. A Theoretical and Numerical Study on Fractional Order Biological Models with Caputo Fabrizio Derivative. Fractal Fract. 2022;6:446. https://doi.org/10.3390/ fractalfract6080446

28. Akinfe KT, Loyinmi AC. A solitary wave solution to the generalized Burgers-Fisher's equation using an improved differential transform method: A hybrid scheme approach. https://doi.org/10.1016/j.heliyon.2021.e07001

Tarig M. Elzaki: https://orcid.org/0000-0002-6946-9267

Mohamed Z. Mohamed: https://orcid.org/0000-0002-8041-5841

A NOVEL ANALYTICAL METHOD FOR THE EXACT SOLUTION
OF THE FRACTIONAL-ORDER BIOLOGICAL POPULATION MODEL

Tarig M. ELZAKI*, Mohamed Z. MOHAMED**/***

*Mathematics Department, College of Sciences and Arts, Alkamel, University of Jeddah, Saudi Arabia
**Mathematics Department, Academy of Engineering and Medical Sciences, Khartoum 12045, Sudan
***Mathematics Department, Prince Muqrin University, Almadinah Almunawwarah, Saudi Arabia

tarig.alzaki@gmail.com, M.zain198327@hotmail.com

Abstract: In this research, we develop a new analytical technique based on the Elzaki transform (ET) to solve the fractional-order biological population model (FBPM) with initial and boundary conditions (ICs and BCs). This approach can be used to locate both the closed approximate solution and the exact solution of a differential equation. The usefulness and validity of this strategy for managing the solution of FBPM are demonstrated using a few real-world scenarios. The dependability of the suggested strategy is also shown using a table and a few graphs. The approximate solutions that were achieved and the convergence analysis are shown in numerical simulations in a range of fractional orders. From the numerical simulations, it can be seen that the population density increases with increasing fractional order, whereas the population density drops with decreasing fractional order.

Key words: fractional biological population model, novel analytical method, Elzaki transform, Mittag-Leffler

1. INTRODUCTION

Although fractional derivatives have a long mathematical history, science did not use them frequently for a very long time. One possible explanation for the unpopularity of fractional derivatives is the prevalence of various non-equivalent definitions of them [1–3]. Furthermore, due to their non-locality, fractional derivatives lack a precise geometrical interpretation [1]. Over the past 10 years, however, fractional calculus has begun to attract the attention of mathematicians and engineers much more. It was discovered that fractional derivatives can effectively imitate a variety of applications, notably interdisciplinary ones. Fractional derivatives can be used to explain a variety of phenomena, such as the non-linear oscillation of earthquakes [3], Kilbas et al. [1] provide an overview of a few fractional derivative applications in continuous mechanics and statistical mechanics.

Many authors have researched the analytical findings on the existence and distinctiveness of fractional differential equation [FDE] solutions. Various techniques, including Adomian decomposition (ADM), Homotopy analysis [5], and many more, have been used in recent years to solve FDEs, FPDEs, and dynamic systems incorporating fractional derivatives. Fractional operators can be used to effectively represent phenomena with the memory effect since they are non-local. We stress that a particular fractional operator can change a PDE from a local to a non-local one by substituting it for the classical derivative with respect to time.

In this essay, the FBPM will be resolved using a novel approach called the Elzaki transform (ET) approach. ET and its variations are used to tackle boundary value problems. The recommended approach [10, 11, 21, 22, 24] presents the solution in a finite series form that is straightforward to compute, but the real strategy offers greater precision because different starting approximations are employed in any iterations. This class of equations

including linear fractional differential equations did not have an analytical solution method prior to the 17th century. Linear and non-linear population problems were handled in [8, 9] using the VIM and HPM. Akinfe and Loyinmi [13] have examined other earlier research attempts on the current fractional biological population model and its applications in quantum physics, optics, fluid modelling employing the ET and a solitary wave solution to the generalised Burgers-Fisher's equation using an improved differential transform method in [14–17, 28]. Additionally, the fractional order model is examined in [18–20] with regard to Esmehan Ucar et al.

The Caputo fractional derivative was used for this study because it enables the formulation of the physical problems to include conventional initial and boundary conditions (BCs). Provide a few other crucial properties of fractional derivatives. A few instances of the identified difficulties are addressed using the general description of the suggested solution. Finding analytical solutions to FBPM using initial conditions (ICs) and BCs is fairly difficult. The current study uses a relatively simple and straightforward methodology to obtain closed-form analytical answers for the FBPM. We'll talk about the fractional biological population model in this article, and this strategy is a potent method for resolving the functional equations that arise from modelling various systems analytically.

The plan of our paper is as follows: Brief definitions of fractional calculus are given in Section 1. Some theorems of the ET are given in Section 2. The novel analytical method is presented in Section 3. The convergence analysis is presented in Section 4. In Section 5, three numerical examples are given to illustrate the applicability of the considered method. Numeric results are presented in Section 6.Section 7 is devoted to the conclusions of the work.

The generalised time-fractional non-linear biological popula-

tion equation we suggest in this paper is as follows:

$$D_t^\alpha \Phi(x,y,t) = \frac{\partial^2}{\partial x^2}(\Phi^2) + \frac{\partial^2}{\partial y^2}(\Phi^2) + f(\Phi),$$
$$0 < \alpha \le 1, \quad t > 0, \tag{1}$$

Given ICs and BCs, and based on Verhulst and Malthusian law, we explore a more generic version of,

$$f(\Phi) = h\,\Phi^a(1 - r\Phi^b), \quad h,a,b,r \in \mathbb{R},$$

They switch to Verhulst and Malthusian laws when choosing exceptional values.

Definition 1: The following definition for the Riemann–Liouville (R–L) [23], fractional integral (FI) operator, of the order

$\alpha > 0$,of $f \in C_\mu, \mu \ge -1$, is given:

$$J^\alpha f(\eta) = \frac{1}{\Gamma(\alpha)}\int_0^\eta (\eta - v)^{\alpha-1}f(v)\,dv, \ \alpha > 0. J^0 f(\eta) = f(\eta).$$

Some properties of J^α, for, $f^n \in C_\mu, n \in N, \alpha, \beta \ge 0$ and $\gamma \ge -1$:

(a) $J^\alpha J^\beta f(\eta) = J^{\alpha+\beta}f(\eta)$

(b) $J^\alpha \eta^\gamma = \frac{\Gamma(\gamma+1)}{\Gamma(\gamma+\alpha+1)}\,\eta^{\alpha+\gamma}$

Definition 2: According to Caputo, the fractional derivative of $f(\eta)$, is: $D^\alpha f(\eta) = J^{m-\alpha}D^m f(\eta)$. For $m-1 < \alpha \le m$, $m \in N, \eta > 0$,and $f \in C_{-1}^m$.

Caputo's fractional derivative (CFD) computes an ordinary derivative first, then a FI to determine the right order of a fractional derivative. The FI operator of RL and the integer order integration are both linear operations:

$$J^\alpha\left(\sum_{i=1}^n c_i f_i(\eta)\right) = \sum_{i=1}^n c_i J^\alpha f_i(\eta),$$

where,$\{c_i\}_{i=1}^n$are constants.

Fractional derivatives are interpreted as having a Caputo, meaning, in the current investigation, provides the reason for using the Caputo definition.

2. ELZAKI TRANSFORM

Here is a short explanation of the modified Sumudu transform, also known as the ET of the function $\Phi(\eta)$,

$$E[\Phi(\eta)] = p\int_0^\infty \Phi(\eta)\,e^{-\frac{\eta}{p}}d\eta = T(p)\,, \quad \eta > 0,$$

where, p is a complex value.

Tarig M. Elzaki has shown in [10, 11, 5] that PDEs, ODEs, systems of PDEs, and Euler-Bernoulli Beam's can all be solved using the modified Sumudu transform, or ET. When Sumudu and Laplace transforms are unsuccessful in solving DEs with variable coefficients, ET can be effectively used [12].

Theorem 1: [2] If $\Phi = \Phi(x,y,t)$, then the partial derivatives are transformed by ET as follows:

$$E\left[\frac{\partial\Phi}{\partial t}\right] = \frac{1}{p}T(x,y,p) - p\,\Phi(x,y,0)\,,$$

$$E\left[\frac{\partial^2\Phi}{\partial t^2}\right] = \frac{1}{p^2}T(x,y,p) - \Phi(x,y,0) - p\frac{\partial\Phi(x,y,0)}{\partial t}\,,$$

$$E\left[\frac{\partial\Phi}{\partial x}\right] = \frac{d}{dx}[T(x,y,p)], \qquad E\left[\frac{\partial^2\Phi}{\partial x^2}\right] = \frac{d^2}{dx^2}[T(x,y,p)],$$

$$E\left[\frac{\partial\Phi}{\partial y}\right] = \frac{d}{dy}[T(x,y,p)], \qquad E\left[\frac{\partial^2\Phi}{\partial y^2}\right] = \frac{d^2}{dy^2}[T(x,y,p)].$$

ET of some functions:

$\Phi(\eta)$	$E[\Phi(\eta)] = T(p)$
1	p^2
η	p^3
η^n	$n!\ p^{n+2}$
$e^{a\eta}$	$\dfrac{p^2}{1-ap}$
$\sin a\eta$	$\dfrac{ap^3}{1+a^2p^2}$
$\cos a\eta$	$\dfrac{p^2}{1+a^2p^2}$

Here, we show some lemmas that can be applied to extract the function $\Phi(\eta)$, from its ET.

Lemma 1: ET of R–L FI operator of order $\alpha > 0$, is represented as:

$$E[J^\alpha\,\Phi(\eta)] = p^\alpha\,T(p).$$

Proof: We begin by:

$$E[J^\alpha\,\Phi(\eta)] = E\left[\frac{1}{\Gamma(\alpha)}\int_0^\eta(\eta-\alpha)^{\alpha-1}\Phi(\eta)\,d\eta\right]$$
$$= \frac{1}{\Gamma(\alpha)}\,\frac{1}{p}\,T(p)\,G(p) = p^\alpha\,T(p)$$

where

$$G(p) = E[\eta^{\alpha-1}] = p^{\alpha+1}\Gamma(\alpha).$$

Lemma 2: ET of CFD for $\alpha > 0$, $m - 1 < \alpha \le m$, $m \in N$, is;

$$E[D_t^\alpha\Phi] = p^{m-\alpha}\left[\begin{array}{l}\frac{T(x,y,p)}{p^m} - \frac{\Phi(x,y,0)}{p^{m-2}} - \frac{\frac{\partial\Phi(x,y,0)}{\partial t}}{p^{m-3}}\\ - \cdots - p\,\frac{\partial^{m-1}\Phi(x,y,0)}{\partial t^{m-1}}\end{array}\right],$$

or

$$E[D_t^\alpha\Phi] = \frac{1}{p^\alpha}E[\Phi] - \sum_{k=0}^{m-1}\frac{\partial^k\Phi(x,y,0)}{\partial t^k}p^{2-\alpha+k}, m-1 < \alpha \le m,$$

The following is the definition of the normal and generalized Mittag-Leffler functions:

$$\in_\alpha(\eta) = \sum_{n=0}^\infty \frac{\eta^n}{\Gamma(n\alpha+1)}, \in_{\alpha,\beta}(\eta) = \sum_{n=0}^\infty \frac{\eta^n}{\Gamma(n\alpha+\beta)}..$$

Lemma 3: If, $\alpha,\beta > 0$, $a \in C$ and $\frac{1}{p^\alpha} > |a|$, the formula for inverse ET is as follows:

$$E^{-1}\left[\frac{p^{\beta+1}}{1+a\,p^\alpha}\right] = \eta^{\beta-1}\in_{\alpha,\beta}(-a\,\eta^\alpha).$$

Proof:

$$\frac{p^{\beta+1}}{1+a\,p^\alpha} = p^{\beta+1}\frac{1}{1+a\,p^\alpha} = p^{\beta+1}\sum_{n=0}^\infty(-a)^n(p^\alpha)^n$$
$$= \sum_{n=0}^\infty(-a)^n\,p^{n\alpha+\beta+1},$$

Tarig M. Elzaki, Mohamed Z. Mohamed
A Novel Analytical Method for Tthe Exact Solution of the Fractional-Order Biological Population Model

DOI 10.2478/ama-2024-0059

Then:

$$E^{-1}\left[\frac{p^{\beta+1}}{1+a\,p^{\alpha}}\right] = E^{-1}\left[\sum_{n=0}^{\infty}(-a)^n\,p^{n\,\alpha+\beta+1}\right]$$

$$= \sum_{n=0}^{\infty}\frac{(-a)^n\eta^{n\alpha+\beta-1}}{\Gamma(n\alpha+\beta)} = A^{\beta-1}\sum_{n=0}^{\infty}\frac{(-a\eta^{\alpha})^n}{\Gamma(n\alpha+\beta)} = \eta^{\beta-1}\in_{\alpha,\beta}(-a\eta^{\alpha}).$$

3. THE NOVEL ANALYTICAL METHOD

We explain the basic tenets of the suggested approach in this section. Let's look at the fractional non-linear non-homogeneous PDE,

$$D_t^{\alpha}\Phi + R[\Phi] + N[\Phi] = g(x,y,t),$$
$$0 < \alpha \le 2,\ 0 \le x \le d,\ 0 \le y \le j,\ t \ge 0,\ d,j \in \mathbb{R}, \qquad (2)$$

with the ICs, and BCs,

$$\Phi(x,y,0) = h_1(x,y),\quad \Phi_t(x,y,0) = h_2(x,y),$$
$$\Phi(0,y,t) = k_1(y,t),\quad \Phi(d,y,t) = k_2(y,t),$$
$$\Phi(x,0,t) = k_3(x,t),\quad \Phi(x,j,t) = k_4(x,t), \qquad (3)$$

R, N are linear and non-linear operators, and g, inhomogeneous terms.

Considering that, Φ_n^* the new method was employed to calculate the new solution,

$$\Phi_n^* = \Phi_n + (d-x)(k_1 - \Phi_n(0,y,t)) + x(k_2 - \Phi_n(d,y,t)) + (j-y)(k_3 - \Phi_n(x,0,t)) + y(k_4 - \Phi_n(x,j,t)), \qquad (4)$$

where $n = 0, 1, \cdots$.

It is obvious that Φ_n^* will meet both the ICs' and BCs' standards. We can resolve Eq. (2) by using the ET to deduce Eq. (3).

$$E[D_t^{\alpha}\Phi] + E\{R[\Phi] + N[\Phi]\} = E[g],$$
$$E(\Phi) = p^2 h_1(x,t) + p^3 h_2(x,t) - p^{\alpha}E\{R[\Phi] + N[\Phi] - g\}. \qquad (5)$$

It is believed that the solution to Eq. (2) has the following series form:

$$\Phi = \sum_{n=0}^{\infty}\Phi_n, \qquad (6)$$

As a result of applying Eq. (6) and the inverse of ET to Eq. (5), we can now determine,

$$\sum_{n=0}^{\infty}\Phi = G - E^{-1}\{p^{\alpha}E[R[\Phi] + N[\Phi]]\}, \qquad (7)$$

When, G an expression is made from a source word and the necessary ICs,

This method relies on how we choose the initial iteration Φ_0 that offers the exact solution in a constrained number of steps.

To discover the solution iteratively, apply the relations listed below.

$$\Phi_{n+1} = E^{-1}\{p^{\alpha}E[R[\Phi_n^*] + N[\Phi_n^*]]\},$$

$$\Phi_0 = G, \qquad (8)$$

From Eqs. (8) and (4), we can infer that:

$$\Phi_0,\ \Phi_1,\ \Phi_2,\cdots,$$

The solution can then be inferred from Eq. (6).

We demonstrate that FBPMs that are under the ICs, BCs, and ET may be resolved using the suggested strategy.

4. CONVERGENCE ANALYSIS CONSIDERED PROBLEM

This section examines the FBPM's convergence for the specified problem, as stated in Eq. (1). To do this, we apply the operator's Eq. (1) as:

$$T(\Phi) = D_t\Phi = \left(D_x^2 + D_y^2\right)\Phi^2 + h\Phi^a - rh\Phi^{a+b},$$

Let $H \in L^2[T],\ \forall\Phi \in H$, [25] where

$$H \in L_{\Phi}^2[(m,n) \times [0,T]],\ \text{such that,}$$

$$\Phi := [(m,n) \times [0,T]] \to R^3,$$

with $m << 0$ and $B = [(m,n) \times [0,T]]$,

where $\|\Phi\|_H^2 = \int_B \Phi^2 dxdydt$, then $E_t^{-1}\{E_t[\Phi(x,y,t)]\} < \infty$

We now assume the following in order to demonstrate T, to be semi-continuous [25]:

Assumption:

H_1 for $\sigma > 0$, exist a constant $\beta > 0$, and $\forall\Phi_1,\Phi_2 \in H$, with $k\|\Phi_1 + \Phi_2\| \le \sigma$,

we obtain

$$\|T(\Phi_1) - T(\Phi_2)\| \le \beta\|\Phi_1 - \Phi_2\|, \forall\Phi_1,\Phi_2 \in H.$$

Theorem 2: (Convergence Condition)[26]. Without initial and BCs convergent to a specific solution, the problem under consideration is examined in Eq. (1).

Making use of the above Assumption for operator $T(\Phi)$, in Eq. (1), to obtain,

$$T(\Phi_1) - T(\Phi_2) = \left(D_x^2 + D_y^2\right)\Phi_1^2 + h\Phi_1^a - rh\Phi_1^{a+b}$$
$$-\{\left(D_x^2 + D_y^2\right)\Phi_2^2 + h\Phi_2^a - rh\Phi_2^{a+b}\} = D_x^2(\Phi_1^2 - \Phi_2^2)$$
$$+D_y^2(\Phi_1^2 - \Phi_2^2) + h(\Phi_1^a - \Phi_2^a) - rh(\Phi_1^{a+b} - \Phi_2^{a+b}),$$

By using the norm, we can get:

$$\|T(\Phi_1) - T(\Phi_2)\| = \|D_x^2(\Phi_1 - \Phi_2)(\Phi_1 + \Phi_2)\|$$
$$+\|D_y^2(\Phi_1 - \Phi_2)(\Phi_1 + \Phi_2)\|$$
$$+h\|(\Phi_1^a - \Phi_2^a)\| - rh\|(\Phi_1^{a+b} - \Phi_2^{a+b})\|,$$

By utilizing the conditions on the operators D_x^2, D_y^2, in H, $\exists \varsigma_1, \varsigma_2 > 0$, and if $a = b = 1$, we can define,

$$D_x^2(\Phi_1 - \Phi_2)(\Phi_1 + \Phi_2) \le \varsigma_1\|\Phi_1 - \Phi_2\|,$$
$$D_y^2(\Phi_1 - \Phi_2)(\Phi_1 + \Phi_2) \le \varsigma_2\|\Phi_1 - \Phi_2\|,$$

Therefore,

$$\|T(\Phi_1) - T(\Phi_2)\| \le \varsigma_1\|\Phi_1 - \Phi_2\| + \varsigma_2\|\Phi_1 - \Phi_2\|$$
$$+h\|\Phi_1 - \Phi_2\| - r\sigma\|\Phi_1 - \Phi_2\| \Rightarrow$$
$$\|T(\Phi_1) - T(\Phi_2)\| \le (\varsigma_1 + \varsigma_2 + h - r\sigma)\|\Phi_1 - \Phi_2\|,$$

Let,

$d = \varsigma_1 + \varsigma_2 + h - r\sigma > 0$, then we can write,

$$\|T(\Phi_1) - T(\Phi_2)\| \le d\|\Phi_1 - \Phi_2\|.$$

Thus, the assumption is met. As a result, the suggested approach converges.

5. ILLUSTRATIVE EXAMPLES

In this section, we will use three numerical examples to illustrate the efficiency and dependability of the method.

Example 1: Take a look at FBPM in one dimension,

$$D_t^\alpha \Phi(x,t) = \frac{\partial^2}{\partial x^2}(\Phi^2) + \Phi\left(1 - \frac{4}{9}\Phi\right), 0 < \alpha \le 1, \quad (9)$$
$$0 \le x \le d, \quad t \ge 0, \quad d,j \in \mathbb{R},$$

With the IC and BCs,

$$\Phi(x,0) = e^{\frac{x}{3}}, \quad \Phi(0,t) = \in_\alpha(t^\alpha), \quad \Phi(d,t) = e^{\frac{d}{3}} \in_\alpha(t^\alpha), \quad (10)$$

By combining the IC, ET, and Eq. (9), the following result is produced:

$$\frac{1}{p^\alpha}E[\Phi(x,t)] - \Phi(x,0)p^{2-\alpha} - E[\Phi(x,t)] =$$
$$E\left[\frac{\partial^2}{\partial x^2}(\Phi^2) - \frac{4}{9}\Phi^2\right],$$
$$\Rightarrow E[\Phi(x,t)] = \frac{p^2 e^{\frac{x}{3}}}{1-p^\alpha} + \frac{p^\alpha}{1-p^\alpha}E\left[\frac{\partial^2}{\partial x^2}(\Phi^2) - \frac{4}{9}\Phi^2\right],$$

Inverse ET suggests that:

$$\Phi((x,t)) = e^{\frac{x}{3}} \in_\alpha(t^\alpha) + E^{-1}\left\{\frac{p^\alpha}{1-p^\alpha}E\left[\frac{\partial^2}{\partial x^2}(\Phi^2) - \frac{4}{9}\Phi^2\right]\right\},$$

The following diagram illustrates the iteration formula using an initial approximation.

$$\Phi_{n+1}(x,t) = E^{-1}\left\{\frac{p^\alpha}{1-p^\alpha}E\left[\frac{\partial^2}{\partial x^2}(\Phi_n^*)^2 - \frac{4}{9}(\Phi_n^*)^2\right]\right\}, \quad (11)$$

with, $\Phi_0(x,t) = e^{\frac{x}{3}} \in_\alpha(t^\alpha)$,

Utilize the BCs in Eq. (4) and $n = 0$, to ascertain:

$$\Phi_0^*(x,t) = \Phi_0(x,t) + (d-x)(\Phi(0,t) - \Phi_0(0,t))$$
$$+x\left(\Phi(d,t) - \Phi_0(d,t)\right) = e^{\frac{x}{3}} \in_\alpha(t^\alpha),$$

Eq. (11), give:

$$\Phi_1(x,t) = E^{-1}\left\{\frac{p^\alpha}{1-p^\alpha}E\left[\frac{\partial^2}{\partial x^2}(\Phi_0^*)^2 - \frac{4}{9}(\Phi_0^*)^2\right]\right\} = 0,$$

Then,

$$\Phi_1 = 0, \quad \Phi_2 = 0, \quad \Phi_3 = 0, \quad \dots,$$

Using Eq. (6), to find the solution of Eq. (9),

$$\Phi(x,t) = \sum_{n=0}^\infty \Phi_n(x,t) = e^{\frac{x}{3}} \in_\alpha(t^\alpha), \text{ if } \alpha = 1, \text{ then: } \Phi(x,t) = e^{\frac{x}{3}+t}.$$

Example 2: Consider the FBPM in two-dimensional,

$$D_t^\alpha \Phi = \frac{\partial^2}{\partial x^2}(\Phi^2) + \frac{\partial^2}{\partial y^2}(\Phi^2) + \Phi, \quad 0 < \alpha \le 1, 0 \le x,y \le \pi, \quad t \ge 0, \quad (12)$$

With the IC and BCs,
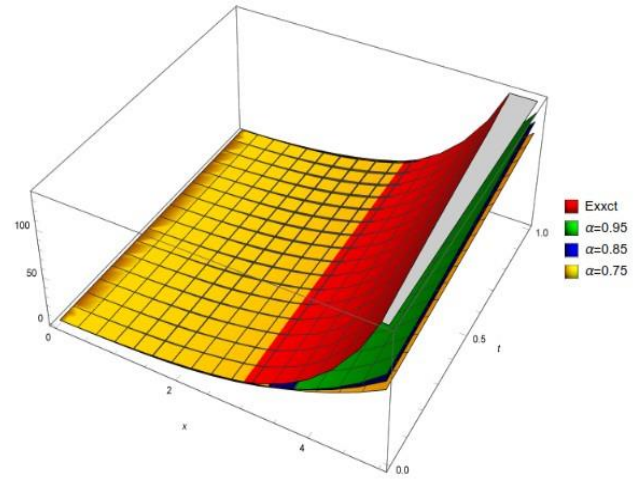
$$\Phi(x,y,0) = \sqrt{sinxsinhy}, \quad \Phi(0,y,t) = 0, \quad \Phi(\pi,y,t) = 0,$$
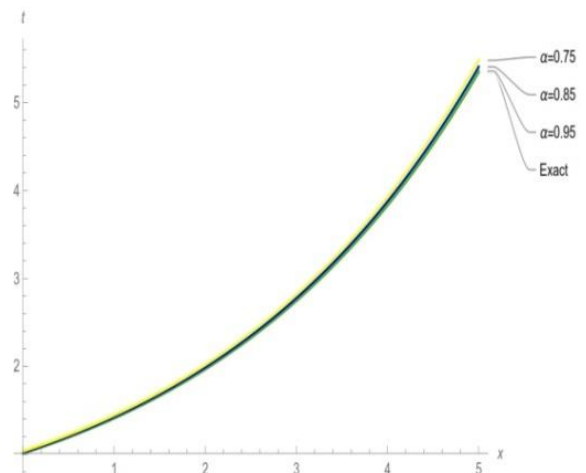$$\Phi(x,0,t) = 0, \quad \Phi(x,\pi,t) = \sqrt{sinxsinh\pi}. \in_\alpha(t^\alpha), \quad (13)$$

By combining the IC, ET, and Eq. (12), the following result is produced:

$$\frac{1}{p^\alpha}E[\Phi] - \Phi(x,y,0)p^{2-\alpha} - E[\Phi(x,y,t)] =$$
$$E\left[\frac{\partial^2}{\partial x^2}(\Phi^2) + \frac{\partial^2}{\partial y^2}(\Phi^2)\right],$$
$$\Rightarrow E[\Phi] = \frac{p^2}{1-p^\alpha}\sqrt{sinxsinhy} + \frac{p^\alpha}{1-p^\alpha}E\left[\frac{\partial^2}{\partial A^2}(\Phi^2) + \frac{\partial^2}{\partial y^2}(\Phi^2)\right],$$

(1)

(2)

Fig. 1,2. 3D and 2D graph representations of the exact and approximate solutions for example 1, when $t = 0.01, \alpha = 0.95, 0.85, 0.75$

Tab. 1. The numirical outcome of example 1 is determinted by comparing the exact and approximate solutions for the two-terms approximation

	t	x	$\alpha = 0,75$	$\alpha = 0,85$	$\alpha = 0,95$	Exact
		0	1.03517	1.02136	1.01293	1.01005
		0.1	1.07026	1.05598	1.04727	1.04429
		0.2	1.10654	1.09177	1.08277	1.07968
		0.3	1.14404	1.12878	1.11947	1.11628
		0.4	1.18282	1.16704	1.15741	1.15411
$\Phi(x,t)$	0.01	0.5	1.22291	1.20659	1.19664	1.19323
		0.6	1.26436	1.24749	1.2372	1.23368
		0.7	1.30722	1.28978	1.27914	1.27549
		0.8	1.35153	1.33349	1.32249	1.31873
		0.9	1.39734	1.37869	1.36732	1.36343
		1	1.4447	1.42542	1.41366	1.40964

Tarig M. Elzaki, Mohamed Z. Mohamed
A Novel Analytical Method for Tthe Exact Solution of the Fractional-Order Biological Population Model

DOI 10.2478/ama-2024-0059

Inverse ET suggests that:

$$\Phi = \in_\alpha (t^\alpha)\sqrt{sinxsinhy} + E^{-1}\left\{\frac{p^\alpha}{1-p^\alpha}E\left[\frac{\partial^2}{\partial x^2}(\Phi^2) + \frac{\partial^2}{\partial y^2}(\Phi^2)\right]\right\},$$

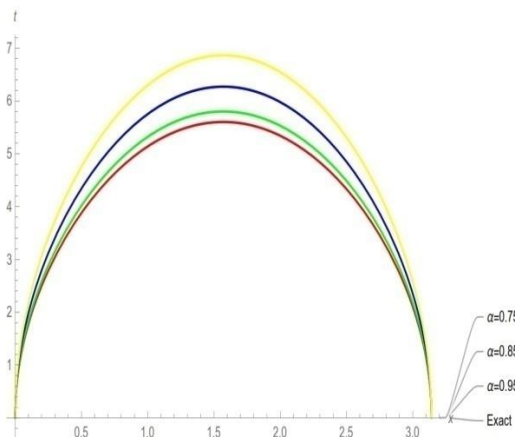The following diagram illustrates the iteration formula using an initial approximation

$$\Phi_{n+1} = E^{-1}\left\{\frac{p^\alpha}{1-p^\alpha}E\left[\frac{\partial^2}{\partial x^2}(\Phi_n^*)^2 + \frac{\partial^2}{\partial y^2}(\Phi_n^*)^2\right]\right\}, \qquad (14)$$

With, $\Phi_0 = \in_\alpha (t^\alpha)\sqrt{sinxsinhy}$, use the BCs in Eq. (4) and $n = 0$ to find:

$$\Phi_0^* = \Phi_0 + (\pi - x)(\Phi(0,y,t) - \Phi_0(0,y,t))$$
$$+x(\Phi(\pi,y,t) - \Phi_0(\pi,y,t))$$
$$+(\pi - y)(\Phi(x,0,t) - \Phi_0(x,0,t)) + y(\Phi(x,\pi,t)$$
$$-\Phi_0(x,\pi,t)) = \in_\alpha (t^\alpha)\sqrt{sinxsinhy},$$

From Eq. (14), we get:

$$\Phi_1 = E^{-1}\left\{\frac{p^\alpha}{1-p^\alpha}E\left[\frac{\partial^2}{\partial x^2}(\Phi_0^*)^2 + \frac{\partial^2}{\partial y^2}(\Phi_0^*)^2\right]\right\} = 0,$$

Then, $\Phi_1 = 0$, $\Phi_2 = 0$, $\Phi_3 = 0$, ... ,

The solution to Eq. (12) can then be found by applying Eq. (6),
$\Phi = \sum_{n=0}^{\infty}\Phi_n = \in_\alpha (t^\alpha)\sqrt{sinxsinhy}$,

when, $\alpha = 1$, then $\Phi = e^t\sqrt{sinxsinhy}$.

(3)

(4)

Fig. 3,4: 3D and 2D graph representations of the exact and approximate solutions for example 2, when t = 0.5, α = 0.95, 0.85, 0.75

Tab. 2. The numerical outcames of example 2 is determinted by comparing the exact and approximate solutions for the two-terms approximation.

	t	y	x	α = 0,75	α = 0,85	α = 0,95	Exact
			0	0.	0.	0.	0.
			$\pi/6$	4.85575	4.43446	4.1024	3.96186
			$\pi/3$	6.39052	5.83608	5.39906	5.21411
$\Phi(y,t)$	0,5	π	$\pi/2$	6.86706	6.27127	5.80167	5.60292
			$2\pi/3$	6.39052	5.83608	5.39906	5.21411
			$5\pi/6$	4.85575	4.43446	4.1024	3.96186
			π	0.	0.	0.	0.

Example 3: Think of the FBPM in two dimensions,

$$D_t^\alpha\Phi(x,y,t) = \frac{\partial^2}{\partial x^2}(\Phi^2) + \frac{\partial^2}{\partial y^2}(\Phi^2) + k\Phi, \ 0 < \alpha \le 1,$$

$$0 \le x \le d, 0 \le y \le j, \ t \ge 0, \ k, d, j \in \mathbb{R}, \qquad (15)$$

With the IC and BCs,

$$\Phi(x,y,0) = \sqrt{xy}, \Phi(0,y,t) = 0,$$
$$\Phi(d,y,t) = \sqrt{dy} \in_\alpha ((kt)^\alpha),$$
$$\Phi(x,0,t) = 0, \ \Phi(x,j,t) = \sqrt{jx} \in_\alpha ((kt)^\alpha), \qquad (16)$$

The recurrence relationship will continue to exist in the same manner, using an initial approximation, as shown in the following:

$$\Phi_{n+1} = E^{-1}\left\{\frac{p^\alpha}{1-p^\alpha}E\left[\frac{\partial^2}{\partial x^2}(\Phi_n^*)^2 + \frac{\partial^2}{\partial y^2}(\Phi_n^*)^2\right]\right\}, \qquad (17)$$

With, $\Phi_0 = \in_\alpha ((kt)^\alpha)\sqrt{xy}$, use the BCs in Eq. (4) and but, $n = 0$ to get:

$$\Phi_0^* = \Phi_0 + (d - x)(\Phi(0,y,t) - \Phi_0(0,y,t))$$

$$+x(\Phi(d,y,t) - \Phi_0(d,y,t)) +$$
$$(j - y)(\Phi(x,0,t) - \Phi_0(x,0,t)) + y(\Phi(x,j,t)$$
$$-\Phi_0(x,j,t)) = \in_\alpha ((kt)^\alpha)\sqrt{xy},$$

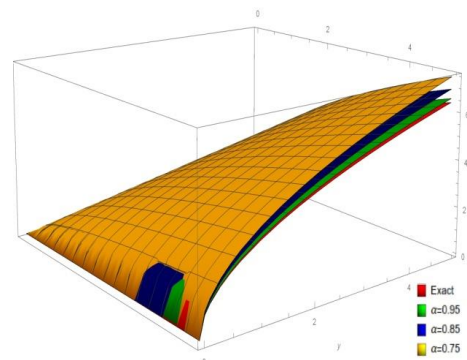The following is taken from Eq. (17):

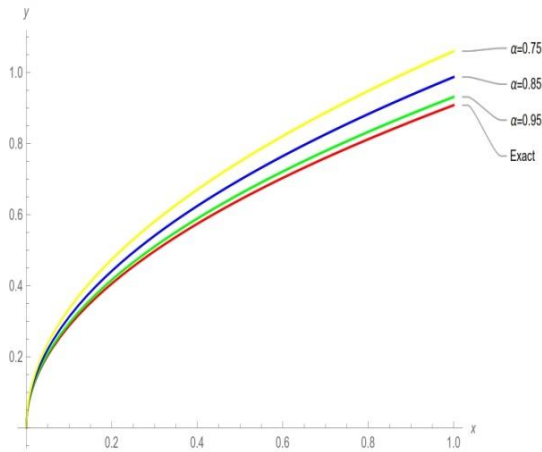$$\Phi_1 = E^{-1}\left\{\frac{p^\alpha}{1-p^\alpha}E[0]\right\} = 0$$

then, $\Phi_1 = 0$, $\Phi_2 = 0$, $\Phi_3 = 0$, ... ,

The solution to Eq. (15) can therefore be found using Eq. (6).

$$\Phi = \sum_{n=0}^{\infty}\Phi_n = \in_\alpha \left((kt)^\alpha\right)\sqrt{xy}, \text{ if, } \alpha = 1, \text{ then: } \Phi = e^{kt}\sqrt{xy}.$$

(5)

(6)

Fig. 5,6. 3D and 2D graph representations of the exact and approximate solutions for example 3, when $t = 0.5, \alpha = 0.95, 0.85, 0.75$.

Tab. 3. The numerical outcames of example 3 is determinted by comparing the exact and approximate solutions for the two-terms approximation.

	t	y	x	$\alpha = 0,75$	$\alpha = 0,85$	$\alpha = 0,95$	Exact
			0	1.03517	1.02136	1.01293	1.01005
			0.1	1.07026	1.05598	1.04727	1.04429
			0.2	1.10654	1.09177	1.08277	1.07968
			0.3	1.14404	1.12878	1.11947	1.11628
			0.4	1.18282	1.16704	1.15741	1.15411
$\Phi(y,t)$	0.5	0.5	0.5	1.22291	1.20659	1.19664	1.19323
			0.6	1.26436	1.24749	1.2372	1.23368
			0.7	1.30722	1.28978	1.27914	1.27549
			0.8	1.35153	1.33349	1.32249	1.31873
			0.9	1.39734	1.37869	1.36732	1.36343
			1	1.4447	1.42542	1.41366	1.40964

6. NUMERICAL RESULTS

This work resolves the FBPM with ICs and BCs using a novel analytical method based on ET. Results obtained using earlier methods are not equivalent to those obtained with current methods. Multiple parameter values are given in Eqs. (9), (12), and (15) to offer a variety of solutions. A variety of solutions can be produced by permitting the arbitrary parameters to have different values in the solutions. The responses gathered are categorized. Visual representations in 2D and 3D are also developed. The following information can be used to describe these plots: Figs. 1, 2, and 3 depict lone waves in different arrangements. Fig. 1 was produced for the values in Eq. (9), showing: at $\alpha = 0.95, 0.85, 0.75$, $t = 0.01$.This mixture falls within the periodic category. Fig. 2 was created using the data $\alpha = 0.95, 0.85, 0.75$, $t = 0.5$, in Eq. (12). Fig. 3 was created using $\alpha = 0.95, 0.85, 0.75, t = 0.5$, in Eq. (15).

The exact solution, which was determined by comparing the values of the exact and approximate solutions of FBPM discovered in this issue for various values of the variables,$0 < x \le 1, t = 0.01$, is provided in Tab.1. Tab.2 contains the approximate solutions, which were ascertained by comparing the values of the exact solutions of FBPM for various values of the variables,

$0 < x \le \pi, y = \pi, t = 0.5. \ 0 < x \le 1, y = 0.5, t = 0.5.$ We were able to determine the exact solution, which is presented in Tab. 3. The suggested approaches were found to be both workable and efficient. Using the Wolfram Mathematica program, the simulations were run and the outcomes were examined.

7. CONCLUSION

In this paper, analytical evaluations of non-linear fractional biological population models with ICs and BCs are carried out. The Riemann–Liouville FI operator generates partially specified fractional derivatives. For the speedy and effective resolution of numerous difficulties, a novel approach based on ET is proposed. Three examples are offered to show the value of the suggested approach. The solutions might be handled in a very simple way. The outstanding ability of the approach to solve non-linear FBPM utilising ICs and BCs allows it to be modified to address a variety of boundary value problems. In addition, 2D and 3D graphs were employed to demonstrate how the suggested approach contributed to the outcomes.

Figures and tables demonstrate that the shapes of the solutions discovered using the suggested method is similar to those of the precise solution when the same parameters were chosen. The proposed method can also be extended to solve additional FPDEs that arise in applied research, according to the ease with which it can be put into practice. In the future, we propose adapting the considered novel approach with the ET scheme for the analysis of nonlinear partial differential equations and certain advanced integral-related problems in fluid dynamics and elasticity and investigating the proposed method's stability and error analysis in forthcoming articles. Finally, we affirm that, subject to ICs and BCs, the proposed technique is valid and applicable to all non-linear FPDEs.

REFERENCES

1. Kilbas AA, Srivastava HM, Trujillo JJ. Theory and Applications of Fractional Differential Equations. Elsevier. San Diego. 2006.
2. Momani S, Shawagfeh NT. Decomposition method for solving fractional Riccati differential equations. Appl.Math. Comput. 2006; 182:1083-1092.
3. Gejji VD, Jafari H. Solving a multi-order fractional differential equation, Appl. Math. Comput. 2007;189:541-548.
4. Hilal EMA, Elzaki TM. Solution of Nonlinear Partial Differential Equations by New Laplace Variational Iteration Method, Journal of Function Spaces. 2014; 1-5. http://dx.doi.org/10.1155/2014/790714.
5. Elzaki TM, Biazar J. Homotopy Perturbation Method and Elzaki Transform for Solving System of Nonlinear Partial Differential Equations. World Applied Sciences Journal.
DOI: 10.5829/idosi.wasj.2013.24.07.1041
6. Elzaki TM, Ishag AA. Modified Laplace Transform and Ordinary Differential Equations with Variable Coefficients, World Engineering & Applied Sciences Journal. 2019; 10 (3): 79-84.
DOI:10.5829/idosi.weasj.2019.79.84
7. Srivastava VK, Awasthi MK, Kumar S. Analytical approximations of two and three dimensional time fractional telegraphic equation by reduced differential transform method. Egypt J Basic Appl Sci.
https:// dx.doi.org/10.1016/j.ejbas.2014.01.002
8. Shakeri F, Dehghan M. Numerical solution of a biological population model using He's variational iteration method. Comput Math Appl.2006;54:1197-209.
9. Roul P. Application of homotopy perturbation method to biological population model. Appl Appl Math. 2010;10:1369-78.

10. Duz M, Elzaki TM. Solution Of Constant Coeffients Partial Derivative Equations With Elzaki Transform Method. Twms J. App. Eng. Math. 2019;9(3):563-570.

11. Ike C, Elzaki TM. Elzaki Transform Method for Natural Frequency Analysis of Euler-Bernoulli Beams. Engineering and Technology Journal. 2023; 1-12. DOI: 10.30684/etj.2023.140211.1456

12. Elzaki TM, Ishag AA. Modified Laplace Transform and Ordinary Differential Equations with Variable Coefficients. World Engineering & Applied Sciences Journal. 2019;10(3):79-84. DOI: 10.5829/idosi.weasj.2019.79.84

13. Akinfe KT, Loyinmi AC. The implementation of an improved differential transform scheme on the Schrodinger equation governing wave-particle duality in quantum physics and optics. Results in Physics. 2022.

14. Akinfe KT. A reliable analytic technique for the modified prototypical Kelvin–Voigt viscoelastic fluid model by means of the hyperbolic tangent function, Partial Differential Equations in Applied Mathematics 2023;7:100523.

15. Akinfe KT, Loyinmi AC. An improved differential transform scheme implementation on the generalized Allen–Cahn equation governing oil pollution dynamics in oceanography, Partial Differential Equations in Applied Mathematics. 2022;6:100416.

16. Akinfe KT, Loyinmi AC. Exact solutions to the family of Fisher's reaction-diffusion equation using Elzaki homotopy transformation perturbation method. Wiley. 2019. DOI:10.1002/eng2.12084

17. Akinfe KT, Loyinmi AC. An algorithm for solving the Burgers–Huxley equation using the Elzaki transform. SN Appl. Sci. 2020;2(7). https://doi.org/10.1007/s42452-019-1653-3

18. Uçar E, Özdemir N. A fractional model of cancer-immune system with Caputo and Caputo–Fabrizio derivatives. Eur. Phys. J. Plus 2021;136(43). https://doi.org/10.1140/epjp/s13360-020-00966-9

19. Ucar E, Özdemir N, Altun E. Fractional order model of immune cells influenced by cancer cells Math. Model. Nat. Phenom. 2019; 14(3):308. DOI: https://doi.org/10.1051/mmnp/2019002

20. Ozdemir N, Uçar S, Eroglu BBI. Dynamical Analysis of Fractional Order Model for Computer Virus Propagation with Kill Signals. International Journal of Nonlinear Sciences and Numerical Simulation. 2019.

21. Hassaballa AA, Elzaki TM. Applications of the Improved G /G Expansion Method for Solve Burgers-Fisher Equation. Journal of Computational and Theoretical Nanoscience, 2017;14: 4664–4668.

22. Elzaki TM, Elzaki SM, Elnour EA. Applications of New Transform Elzaki Transform to Mechanics, Electrical Circuits and Beams Problems. Global Journal of Mathematical Sciences: Theory and Practical. 2012;4(1):25-34.

23. Mohamed M, hamza A, Elzaki TM, Algolam M, Elhussein S. Solution of Fractional Heat-Like and Fractional Wave-Like Equation by Using Modern Strategy. Acta Mechanica et Automatica 2023;17(3):372-380. https://doi.org/10.2478/ama-2023-0042.

24. Elzaki TM, Shams EA, Areshi M, Chamekh M. Fractional partial differential equations and novel double integral transform, Journal of King Saud University. 2022;34(3):101832.

25. Gadain H.E. Application of double Laplace decomposition method for solving singular one dimensional a system of hyperbolic equations. J. Nonlinear Sci. Appl. 2017;10:111–121.

26. Kaya D, Inan IE. A convergence analysis of the ADM and an application. Appl. Math. Comput. 2005;161:1015–1025.

27. Rahman MU, Althobaiti A, Riaz MB, Al-Duais FS. A Theoretical and Numerical Study on Fractional Order Biological Models with Caputo Fabrizio Derivative. Fractal Fract. 2022;6:446. https://doi.org/10.3390/ fractalfract6080446

28. Akinfe KT, Loyinmi AC. A solitary wave solution to the generalized Burgers-Fisher's equation using an improved differential transform method: A hybrid scheme approach. https://doi.org/10.1016/j.heliyon.2021.e07001

Tarig M. Elzaki: https://orcid.org/0000-0002-6946-9267

Mohamed Z. Mohamed: https://orcid.org/0000-0002-8041-5841

NOVEL INTEGRAL TRANSFORM
TREATING SOME Ψ-FRACTIONAL DERIVATIVE EQUATIONS

Mourad CHAMEKH[*] , **Mohamed Ali LATRACH**[**] , **Tarig M. ELZAKI**[***]

*Department of Mathematics, University of Jeddah, College of Science and Arts El Kamel, Jeddah, Saudi Arabia
**University of Tunis El Manar, National Engineering School at Tunis, LAMSIN, B.P. 37, 1002, Tunis-Belvédère, Tunisia
***Department of Mathematics, University of Jeddah, College of Science and Arts El Kamel, Jeddah, Saudi Arabia

chamekhmourad1@gmail.com, mohamedali.latrach@enit.utm.tn, tarig.alzaki@gmail.com

Abstract: The paper deals with a new integral transformation method called Ψ-Elzaki transform (PETM) in order to analyze some Ψ-fractional differential equations. The proposed method uses a modification of the Elzaki transform that is well adapted to deal with Ψ-fractional operators. To solve the nonlinear Ψ-fractional differential equations, we combine the PETM by an iterative method to overcome this nonlinearity. To validate the accuracy and efficiency of this approach, we compare the results of the discussed numerical examples with the exact solutions.

Key words: Ψ-Integral transform, Ψ-Hilfer derivative, Ψ-Caputo derivative, Ψ-Hilfer fractional derivative equation.

1. INTRODUCTION

In the literature, we find a several applications of the fractional derivatives such as studying the undernutrition problems in pregnant women and predicting their intricacies, and investigating the behaviour of neural networks under challenging circumstances by modelling of Hindmarsh-Rose neuron (HRN) in the biomedical field [1, 2]. We also find the modelling of wind turbine system and their chaotic permanent magnet synchronous generator in the physical field [3], and describing the chemical reactions and activation energy by the artificial neural networks (ANN) in the chemical field [4]. The fractional derivatives can also present models for many applications in stochastic analysis [5], and can be used to extend the physical results obtained in [6−10].

Fractional derivatives and fractional integral operators have been well developed in the scientific literature, especially in recent years. Much of the results obtained are compatible with the needs of physical and biological problems. For example, we find that the mathematical models studied in biology which intervene in many epidemiological phenomena can lead to fractional systems as in [11], which relies on the stochastic perturbation technique to simulate the propagation evolution of the Lassa fever. Moreover, researchers have never lost sight of the fundamental need to ensure better models that describe certain problems more realistically (see for example [12]). Therefore, new methods and definitions of fractional sums/differences are developed for discrete fractional calculus (DFC) [13−15], and new methods and definitions of fractional derivatives are developed to give an even better description of some dynamical problems [16] and some chaotic systems [17]. Integral inequalities were also studied by Rashid et al., which are based on fractional calculus [18].

The classical derivatives of fractional order have been known by certain approaches such as Riemann-Liouville, Caputo, GrAnwald-Letnikov...etc. Despite the diversity of these notions, researchers have continued to develop some generalized notions that place these fractional derivatives into a more general concept. This is perhaps the main purpose of the birth of the Ψ-fractional derivative. This concept of the fractional derivative with respect to another function has existed since 1964 when Érdlyi started the discussion in [19] of this generalized derivative. Next, Olser in 1970 gave a precise definition of the Ψ-fractional derivative [20]. Since then, many scholars have developed this notion, including examples from Almeida [21, 22], Sousa, and Oliveira [23]. But the interest of these notions was not purely mathematical. Conversely, recent works show more and more the efficiency of certain models based on the type of Ψ-fractional differential equation. For example, the modeling taking into account the relaxation and/or the law of deformation of certain bodies is based on this type of generalized fractional equations Yang [24]. Concerning the study of Ψ-fractional differential equation by the integral transform method, we can cite the works of Jarad and Abdeljawad that proposed a Ψ-Laplace transformation to treat the type of the equations [25]. Using Elzaki transform Singh et al. in [26] has developed the Hilfer-Prabhakar fractional derivative to study the free electron laser equation. Motivated by these physical applications, we propose in this article to treat some equations comprising fractional derivatives respecting another function. However, the Ψ-fractional derivative problems are not easy to solve. Given their obvious interest, we propose a new technique called Ψ-Elzaki transform based on the Elzaki transform (ET) developed by Elzaki [27], which can enrich the research work on approximate solutions. Furthermore, we will be extending our technique to nonlinear problems, proposing a combination of the Ψ-Elzaki transform and an iterative procedure to overcome the nonlinearity.

sciendo

Mourad Chamekh, Mohamed Ali Latrach, Tarig M. Elzaki
Novel Integral Transform Treating Some Ψ-Fractional Derivative Equations

DOI 10.2478/ama-2024-0060

2. PRELIMINARY

In this section, we recall certain definitions and results. For more details concerning these definitions and the proofs of the properties, we can consult for example [3, 21, 28, 29].

2.1. Definitions

Let u an integrable function on $[\delta_1, \delta_1] \to \mathbb{R}, (\alpha, p) \in \mathbb{R} \times \mathbb{N}$ and $\Psi \in C^p([\delta_1, \delta_1])$ such that $\Psi'(t) > 0$. We have the following definitions:

– We define the right Ψ-Riemann-Liouville fractional integral of order $\alpha > 0$ as follow:

$$I_{\delta_1^+}^{\alpha,\Psi} u(t) = \frac{1}{\Gamma(\alpha)} \int_{\delta_1}^t \Psi'(s) u(s) \big(\Psi(t) - \Psi(s)\big)^{\alpha-1} ds. \qquad (1)$$

– For $\alpha > 0$, $n - 1 < \alpha \le n$, the left and right Ψ-Reimann-Liouville fractional derivative is given as:

$$D_{\delta^\pm}^{n-\alpha,\Psi} u(t) = \left(\frac{1}{\Psi'(t)}\frac{d}{dt}\right)^n I_{\delta^\pm}^{\alpha,\Psi} u(t).$$

– For $n - 1 < \alpha \le n$, the left and right Ψ-Caputo fractional derivative of order α is defined as:

$${}^C D_{\delta^\pm}^{\alpha,\Psi} u(t) = I_{\delta^\pm}^{\alpha,\Psi} u_\Psi^{[n]}(t), \text{ with } u_\Psi^{[n]}(t) = \left(\frac{1}{\Psi'(t)}\frac{d}{dt}\right)^n u(t).$$

A more general definition that covers Ψ-Riemann-Liouville and Ψ-Caputo fractional derivatives, this is the following fractional derivative of Ψ-Hilfer.

Definition 1 ([30]): Let u be an integrable function defined on $[\delta_1, \delta_2]$, and $\Psi \in C^1([\delta_1, \delta_2], \mathbb{R})$ be an increasing function such that $\Psi'(t) \ne 0$, for all $t \in [\delta_1, \delta_2]$. The Ψ-Hilfer fractional derivative right-sided of order $\mu > 0$, $n = \lceil \mu \rceil$ and of type $0 \le \alpha \le 1$ is defined by:

$${}^H D_{\delta_1^+}^{\mu,\alpha,\Psi} u(t) = I_{\delta_1^+}^{\alpha(n-\mu),\Psi} \left(\frac{1}{\Psi'(t)}\frac{d}{dt}\right)^n I_{\delta_1^+}^{(1-\alpha)(n-\mu),\Psi} u(t).$$

In the rest, we denote $\Lambda_{\delta_1^+}$, $\Lambda_{\delta_1^-}$ respectively by Λ_{δ_1}, Λ_{δ_2} where Λ denotes the fractional integral or derivative operators defined in this paper. For these definitions, we have the following properties.

2.2. Properties

If $u(t) = \big(\Psi(t) - \Psi(\delta_1)\big)^\eta$ where $\eta > n$ and $\alpha > 0$ then,

$${}^C D_{\delta_1}^{\alpha,\Psi} u(t) = \frac{\Gamma(\eta+1)}{\Gamma(\eta-\alpha+1)} \big(\Psi(t) - \Psi(\delta_1)\big)^{\eta-\alpha}. \qquad (2)$$

$$I_{\delta_1}^{\alpha,\Psi} u(t) = \frac{\Gamma(\eta+1)}{\Gamma(\eta+\alpha+1)} \big(\Psi(t) - \Psi(\delta_1)\big)^{\eta+\alpha}. \qquad (3)$$

$${}^C D_{\delta_1}^{\alpha,\Psi} I_{\delta_1}^{\alpha,\Psi} u(t) = u(t). \qquad (4)$$

$$I_{\delta_1}^{n,\Psi} u_\Psi^{[n]}(t) = u(t) - \sum_{k=0}^{n-1} \frac{u_\Psi^{[n]}(\delta_1)}{k!} \big(\Psi(t) - \Psi(\delta_1)\big)^k. \qquad (5)$$

$$I_{\delta_1}^{\alpha,\Psi} {}^C D_{\delta_1}^{\alpha,\Psi} u(t) =$$
$$u(t) - \sum_{k=0}^{n-1} \frac{u_\Psi^{[n]}(\delta_1)}{k!} \big(\Psi(t) - \Psi(\delta_1)\big)^k. \qquad (6)$$

Proof. For to prove the formula (2), we have

$$u_\Psi^{[n]}(t) = \frac{\Gamma(\eta+1)}{\Gamma(\eta-n+1)} u(t)^{\frac{\eta-n}{\eta}}.$$

Using the definition of the Ψ-Caputo fractional derivative, we can write

$${}^C D_{\delta_1}^{\alpha,\Psi} u(t) = \frac{\Gamma(\eta+1)}{\Gamma(\eta-n+1)} I_{\delta_1}^{n-\alpha,\Psi} u(t)^{\frac{\eta-n}{\eta}} =$$

$$\frac{\Gamma(\eta+1)}{\Gamma(n-\alpha)\Gamma(\eta-n+1)} \int_{\delta_1}^t \Psi'(t) \big(\Psi(t) - \Psi(s)\big)^{n-\alpha-1} u(s)^{\frac{\eta-n}{\eta}} ds =$$

$$\Omega(t) \int_{\delta_1}^t \frac{u'(s)}{\eta} \left(\frac{(u(s))^{1-\eta}}{u(t)}\right)^{\frac{1}{\eta}} \left(1 - \left(\frac{u(s)}{u(t)}\right)^{\frac{1}{\eta}}\right)^{n-\alpha-1} \left(\frac{u(s)}{u(t)}\right)^{\frac{\eta-n}{\eta}} ds,$$

where $\Omega(t) = \frac{\Gamma(\eta+1)u(t)^{\frac{\eta-\alpha}{\eta}}}{\Gamma(n-\alpha)\Gamma(\eta-n+1)}$.

With the change of variables $v(s) = \left(\frac{u(s)}{u(t)}\right)^{\frac{1}{\eta}}$, we obtain

$${}^C D_{\delta_1}^{\alpha,\Psi} u(t) = \Omega(t) \int_0^1 (1-v)^{n-\alpha-1} v^{\eta-n} dv = \Omega(t) B(n-\alpha, \eta-n+1),$$

where B is the Beta function that satisfies $B(x,y) = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)}$ which ends the proof of the formula (2).

In the formula proof (2), we have already implied the formula proof (3) where we have

$$\frac{\Gamma(\eta+1)}{\Gamma(\eta-n+1)} I_{\delta_1}^{n-\alpha,\Psi} u(t)^{\frac{\eta-n}{\eta}} = \frac{\Gamma(\eta+1)}{\Gamma(\eta-\alpha+1)} u(t)^{\frac{\eta-\alpha}{\eta}}.$$

Then

$$I_{\delta_1}^{n-\alpha,\Psi} \big(\Psi(t) - \Psi(\delta_1)\big)^{\eta-\alpha} =$$

$$\frac{\Gamma(\eta-n+1)}{\Gamma(\eta-\alpha+1)} \big(\Psi(t) - \Psi(\delta_1)\big)^{\eta-\alpha}.$$

With simple changes of variables for $n - \alpha$ and $\eta - n$, we complete the formula proof (3).

Formula (4) is a direct consequence of formulas (2) and (3) and, formulas (5) and (6) have been well proven in [29].

3. NOVEL Ψ-INTEGRAL TRANSFORM METHOD

This section discusses a type of method that has been and still is useful for solving certain differential equations, namely integral transforms. This method is useful because it can turn a complicated problem into a simpler one. In addition to their clarity, integral transformations have also helped us to develop some formulas in fractional calculus. We can cite for example Laplace transforms, Sumudu transform, Jafari transform, Elzaki transform, etc. [31−33].

3.1. Elzaki transform

Considering the following Elzaki transform:
Definition 2 Let a function u with respect to t, let $s_1, s_2 > 0$, it exists $M > 0$ such that $|u(t)| < M \exp\left(\frac{|t|}{s_j}\right)$, for $t \in (-1)^j \times [0, \infty)$. Elzaki transform is defined by:

$$U(s) = \mathcal{E}[u(t)] = \mathcal{E}(s) = s \int_0^\infty u(t)\, exp\left(-\frac{t}{s}\right) dt, \qquad (7)$$

where s is a complex number, $t > 0$ and $|s| \in [s_1, s_2]$.

By using the properties of this proposed integral transform, we can have an efficient tool to deal with linear fractional equations. Moreover, we can reach nonlinear cases by combining ET with other techniques. For this, let us recall the following formulas dealing with the integral and the classical fractional derivative by ET.

$$\mathcal{E}[^{RL}I_0^\alpha u] = s^\alpha U(s). \qquad (8)$$

$$\mathcal{E}[^{RL}D_0^\alpha u] = \frac{U(s)}{s^\alpha} - \sum_{k=0}^{n-1} s^{2-n+k}\, {}^{RL}I_0^{n-k-\alpha} u(0). \qquad (9)$$

$$\mathcal{E}[^{C}D_0^\alpha u] = \frac{U(s)}{s^\alpha} - \sum_{k=0}^{n-1} s^{2-\alpha+k} u^{[k]}(0). \qquad (10)$$

3.2. Generalized Elzaki transform

The aim of this study is to show the relevance of this novel transform integral and its efficiency in solving certain Ψ-fractional equations. We will consider the novel following definition of generalized integral transform:

Definition 3 Let a function u on $[0, \infty)$ with respect to t, and Ψ an increasing function such that $\Psi(0) = 0$. Let $s_1, s_2 > 0$, if it exists $M > 0$ such that $|u(t)| < M \, exp\left(\frac{|t|}{s_j}\right)$, for $t \in (-1)^j \times [0, \infty)$. Then, the Elzaki transform of u with respect to Ψ is defined by:

$$U_\Psi(s) = \mathcal{E}_\Psi[u(t)](s) = s \int_0^\infty \Psi'(t) u(t)\, exp\left(-\frac{\Psi(t)}{s}\right) dt, \qquad (11)$$

where s is a complex number, $t > 0$ and $|s| \in [s_1, s_2]$.

Jarad and Abdeljawad in [25] have defined the Ψ-Laplace transform. As we have Duality's ownership between ET and LT, we can easily check duality between the Ψ-ELzaki and Ψ-Laplace transforms. The duality relation is given by the following relations:

$$\mathcal{E}_\Psi[u(t)](s) = s \mathcal{L}_\Psi[u(t)]\left(\frac{1}{s}\right), \qquad (12)$$

$$\mathcal{L}_\Psi[u(t)](s) = s \mathcal{E}_\Psi[u(t)]\left(\frac{1}{s}\right). \qquad (13)$$

Using this duality between the two transformations, we can make certain results developed in [25, 34] for the Ψ-Laplace transform into Ψ-Elzaki transform.

3.3. Convergence of PET and properties

Theorem 1: Let Ψ an increasing function with $\Psi(0) = 0 \in \mathbb{R}$, and u is continuous on $[0, +\infty[$ and a Ψ-exponentially function bounded order $\frac{1}{d} > 0$ (i.e. $\|u\|_\infty \leq M \, exp\left(\frac{\Psi(t)}{d}\right)$; with M a positive constant). Then, the Ψ-Elzaki transform of u exists for $s > d$.

Proof: We have

$$|\mathcal{E}_\Psi[u(t)]| = \left| s \int_0^\infty \Psi'(t) u(t)\, exp\left(-\frac{\Psi(t)}{s}\right) dt \right|$$

$$\leq s \int_0^\infty \Psi'(t)\, exp\left(-\frac{\Psi(t)}{s}\right) \|u\|_\infty dt$$

$$\leq sM \int_0^\infty \Psi'(t)\, exp\left(-\frac{\Psi(t)}{s} + \frac{\Psi(t)}{d}\right) dt.$$

Using the fact that $s > d$, so the primitive limit is cancels for infinity, we get

$$|\mathcal{E}_\Psi[u(t)]| \leq \frac{s^2 M d}{s - d}.$$

Then the PET is convergent.

We can also define the inverse of PET by:

$$\mathcal{E}_\Psi^{-1}[U(s)](t) = \frac{1}{2\pi j} \int_{a-j\infty}^{a+j\infty} U\left(\frac{1}{s}\right) exp\left(-\frac{\Psi(t)}{s}\right) s\, dt. \qquad (14)$$

Remark 1: We get the immediate result: If

$$\mathcal{E}_\Psi[u(t)](s) = U(s), \text{ then } \mathcal{E}_\Psi[u(\Psi(t))](s) = U(s). \qquad (15)$$

i.e. we have $U_\Psi(s) = U(s)$.

Using Remark. 1 and the duality propriety of ET and LT, and the results proved in [25] we can easily get the following proprieties:

Properties:

$$\mathcal{E}_\Psi[exp(\lambda\Psi(t))] = \frac{s^2}{1 - \lambda s}. \qquad (16)$$

$$\mathcal{E}_\Psi\left[(\Psi(t))^\beta\right] = \Gamma(\beta + 1) s^{\beta+2}. \qquad (17)$$

3.4. PET and Ψ-Fractional operators

In this section, we will discuss some results of PET that can be useful for solving fractional equations of the type Ψ-Caputo. But these results can also be extended to Ψ-Hilfer or Ψ-Riemann-Liouville Fractional equations type and other equations based on Ψ-fractional differential operators.

Proposition 1: If $\alpha \in \,]n-1, n]$, $n \in \mathbb{N}$ then

$$\mathcal{E}_\Psi[^{RL}I_0^{\alpha,\Psi}u] = s^\alpha U_\Psi(s). \qquad (18)$$

$$\mathcal{E}_\Psi[^{RL}D_0^{\alpha,\Psi}u] = \frac{U_\Psi(s)}{s^\alpha} - \sum_{k=0}^{n-1} s^{2-n+k}\, {}^{RL}I_0^{n-k-\alpha,\Psi}u(0). \qquad (19)$$

$$\mathcal{E}_\Psi[^{C}D_0^{\alpha,\Psi}u] = \frac{U_\Psi(s)}{s^\alpha} - \sum_{k=0}^{n-1} s^{2-\alpha+k}\, {}^{RL}D_0^{k,\Psi}u(0). \qquad (20)$$

$$\mathcal{E}_\Psi[^{H}D_0^{\mu,\alpha,\Psi}u] = \frac{U_\Psi(s)}{s^\mu} - \sum_{k=0}^{n-1} s^{2-n(1-\alpha)+k}\, I_0^{(1-\alpha)(n-\mu)-k,\Psi}u(0). \qquad (21)$$

Proof: For to prove the formula (18):
Using the duality between ET and LT, we can write

$$\mathcal{E}_\Psi[^{RL}I_0^{\alpha,\Psi}u](s) = s\mathcal{L}_\Psi[^{RL}I_0^{\alpha,\Psi}u]\left(\frac{1}{s}\right).$$

With Laplace transform, the Ψ-Riemann-Liouville integral is given by

$$\mathcal{L}_\Psi[^{RL}I_0^{\alpha,\Psi}u](s) = s^{-\alpha}\mathcal{L}_\Psi[u](s).$$

Using both previous formulas, we get

$$\mathcal{E}_\Psi[^{RL}I_0^{\alpha,\Psi}u](s) = \left(\frac{1}{s}\right)^{-\alpha} s\mathcal{L}_\Psi[u]\left(\frac{1}{s}\right) = s^\alpha U_\Psi(s).$$

The remainders of the results can be easily demonstrated in the same way.

3.5. PET and some special functions

Special functions play an important role in the definitions of certain fractional derivatives (such as fractional derivatives involving the kernel of the Mittag-Leffler function (ML) [24]) or even in the resolution of certain fractional differential equations. For this, it is useful to present some results concerning these functions.

Definition 4: We consider the following generalized version of the Mittag-Leffler function defined in [24], recognized as the Prabhakar function

$$E_{\alpha,\beta}^{\gamma}(z) = \sum_{k=0}^{\infty} \frac{\gamma_k}{k!\,\Gamma(\alpha k+\beta)} z^k, \qquad (22)$$

where $\gamma_k = \frac{\Gamma(\gamma+k)}{\Gamma(\gamma)}$; $(\alpha,\beta,\gamma,z) \in \mathbb{C}^4$, with $Re(\alpha) > 0$.

In fractional differences, the authors use the Kernel of the discrete Mittag-Leffler function, see [14, 15] for more details.

By PET, we have the following results:

Proposition 2 Assume that $Re(\alpha) > 0$ and $|a\,s^\beta| < 1$. Applying the PET, we get for ML function with one parameter

$$\mathcal{E}_\Psi\left[E_\beta\left(a(\Psi(t))^\beta\right)\right] = \frac{s^2}{1-as^\beta}. \qquad (23)$$

For two parameter ML function we have

$$\mathcal{E}_\Psi\left[(\Psi(t))^{\beta-1} E_{\beta,\beta}\left(a(\Psi(t))^\beta\right)\right] = \frac{s^{\beta+1}}{1-as^\beta}. \qquad (24)$$

For Prabhakar function, we get

$$\mathcal{E}_\Psi\left[(\Psi(t))^{\alpha-1} E_{\beta,\alpha}^{\gamma}\left(a(\Psi(t))^\beta\right)\right] = \mathcal{E}\left[t^{\alpha-1} E_{\beta,\alpha}^{\gamma}\left(at^\beta\right)\right]$$
$$= \frac{s^{\alpha+1}}{(1-as^\beta)^\gamma}. \qquad (25)$$

We have

$$^{C}D_0^{\eta,\Psi}\,(\Psi(t))^{\alpha-1} E_{\beta,\alpha}^{\gamma}\left(a(\Psi(t))^\beta\right)$$
$$= (\Psi(t))^{\beta-\eta-1} E_{\beta,\alpha-\eta}^{\gamma}\left(a(\Psi(t))^\beta\right) \qquad (26)$$

The proof of these properties is immediate if Remark.1 is taken into account with the properties of ET developed in the scientific literature.

4. SOLUTION FOR Ψ-FRACTIONAL DERIVATIVE EQUATIONS

Although the concept of Ψ-fractional operator has been discussed for decades, the notion of Ψ-fractional derivative equation has only been discussed frequently in recent years. For example, the random differential equations via the fractional derivative Ψ-Hilfer [37], Ψ-fractional differential equations [29], and for modeling taking into account the relaxation of certain bodies [24]. Guided by these motivations, we propose to discuss the solution to the following types of equations:

$$\begin{cases} ^{C}D_{\delta_1}^{\alpha,\Psi} u(t) = g(t,u), & t \in [\delta_1,\delta_2], \quad n = \lceil\alpha\rceil, \\ u^{(k)}(\delta_1) = u_k, u^{(n-1)}(\delta_2) = u_{n-1}, & k = 0,1,\cdots,n-1. \end{cases}$$
$$(27)$$

Applying the PET on the (27), and using (20), we get

$$\mathcal{E}_\Psi[u] = s^\alpha \mathcal{E}_\Psi[g(t,u)] + \sum_{k=0}^{n-1} s^{2+k} D_0^{k,\Psi} u_k. \qquad (28)$$

By the inverse of PET, we obtain

$$u(t) = \mathcal{E}_\Psi^{-1}\{s^\alpha \mathcal{E}[g(t,u)]\} + \sum_{k=0}^{n-1} \frac{(\Psi(t)-\Psi(0))^k}{k!} D_0^{k,\Psi} u_k. \qquad (29)$$

To overcome the nonlinear term of this equation, one can use a direct iterative method or the one described for example by [33].

Assume that have the following equation:

$$u = N(u) + f, \qquad (30)$$

with N a non linear operator, and f a given function. We can solving (30) by the following iterative method

$$\begin{cases} u_0 = f, \\ u_{p+1} = \mathcal{N}(u_p). \end{cases} \qquad (31)$$

With $\mathcal{N}$ is approximate of N. Then, the solution is

$$u = \sum_{k=0}^{\infty} u_k.$$

In the situation of problem (27), we can pose

$$u_0 = \sum_{k=0}^{n-1} \frac{t^k}{k!} D_0^{k,\Psi} u(0). \qquad (32)$$

$$N(u) = \mathcal{E}_\Psi^{-1}\{s^\alpha \mathcal{E}[g(t,u)]\}. \qquad (33)$$

Remark 2 If N is linear the iterative procedure in (31) is as follows: $u_{p+1} = N(u_p)$.

5. NUMERICAL EXAMPLES

Based on the Ψ-Elzaki transformation, we present some numerical examples to discuss the efficiency of the proposed method. The main objective is to compare the results obtained with the exact solution in each of the considered examples.

Example 1: We begin with an illustrative example that seeks the eigenfunction of the Ψ-Caputo fractional derivative operator. This is a relaxation model based on a fractional derivative with respect to another function. The model is given by:

$$\begin{cases} ^{C}D_0^{\alpha,\Psi} u(t) = \lambda u(t), & for\ t \in [0,T], T > 0, \alpha \in \,]0,1], \\ u(0) = 1. \end{cases}$$
$$(34)$$

By the formula (26), we can find the exact solution to the problem

$$u(t) = E_\alpha(\lambda(\Psi(t) - \Psi(s))^\alpha).$$

Applying the PET to the problem (34), we obtain

$$u(t) = \mathcal{E}_\Psi^{-1}\{s^\alpha \mathcal{E}[\lambda u]\} + u(0).$$

In this case, the iterative procedure reduces to the following system

$$u_0 = 1,$$
$$u_{p+1} = \lambda \mathcal{E}_\Psi^{-1}\{s^\alpha \mathcal{E}[u_p]\}.$$

Then,

$$u_1 = \lambda \mathcal{E}_\Psi^{-1}\{s^\alpha \mathcal{E}[u_0]\} = \lambda \mathcal{E}_\Psi^{-1}\{s^{\alpha+2}\} = \lambda \frac{(\Psi(t)-\Psi(0))^\alpha}{\Gamma(\alpha+1)},$$

$$u_2 = \lambda \mathcal{E}_\Psi^{-1}\{s^\alpha \mathcal{E}[u_1]\} = \frac{\lambda^2}{\Gamma(\alpha+1)} \mathcal{E}_\Psi^{-1}\{s^{2\alpha+2}\Gamma(\alpha+1)\} =$$

$\lambda^2 \frac{\left(\Psi(t)-\Psi(0)\right)^{2\alpha}}{\Gamma(2\alpha+1)}$, and so on.

We can write

$$u_p = \lambda^p \frac{\left(\Psi(t) - \Psi(0)\right)^{p\alpha}}{\Gamma(p\alpha + 1)}.$$

The solution is given by

$$u = \sum_{k=0}^{\infty} \lambda^p \frac{\left(\Psi(t)-\Psi(0)\right)^{p\alpha}}{\Gamma(p\alpha+1)} = E_\alpha\left(\lambda\left(\Psi(t) - \Psi(0)\right)^{\alpha}\right).$$

Example 2: We choose an inhomogeneous relaxation model based on the Ψ-fractional derivative. This model has been solved numerically in [38] having only the Caputo fractional derivative.

$$\begin{cases} {}^{C}D_0^{\alpha,\Psi} u(t) + u(t) = \frac{\left(\Psi(t)-\Psi(0)\right)^{4-\alpha}}{\Gamma(5-\alpha)}, \ t \geq 0, \alpha \in \ ]0,1]. \\ u(0) = 0. \end{cases}$$

(35)

Applying (29), we obtain

$$u(t) = -\mathcal{E}_\Psi^{-1}\{s^\alpha \mathcal{E}[u]\} + \mathcal{E}_\Psi^{-1}\left\{s^\alpha \mathcal{E}\left[\frac{\left(\Psi(t)-\Psi(0)\right)^{4-\alpha}}{\Gamma(5-\alpha)}\right]\right\}.$$

Then

$$\begin{cases} u_0 = \mathcal{E}_\Psi^{-1}\left\{s^\alpha \mathcal{E}\left[\frac{\left(\Psi(t)-\Psi(0)\right)^{4-\alpha}}{\Gamma(5-\alpha)}\right]\right\}, \\ u_{p+1} = -\mathcal{E}_\Psi^{-1}\{s^\alpha \mathcal{E}[u_p]\}, \quad p \geq 0. \end{cases}$$

Putting

$$u_0 = \mathcal{E}_\Psi^{-1}\left\{s^\alpha \mathcal{E}\left[\frac{\left(\Psi(t)-\Psi(0)\right)^{4-\alpha}}{\Gamma(5-\alpha)}\right]\right\} = \frac{\left(\Psi(t)-\Psi(0)\right)^{4}}{\Gamma(5)}.$$

$$u_1 = -\mathcal{E}_\Psi^{-1}\left\{s^\alpha \mathcal{E}\left[\frac{\left(\Psi(t)-\Psi(0)\right)^{4}}{\Gamma(5)}\right]\right\} = -\frac{\left(\Psi(t)-\Psi(0)\right)^{\alpha+4}}{\Gamma(\alpha+5)}.$$

$$u_2 = -\mathcal{E}_\Psi^{-1}\left\{s^\alpha \mathcal{E}\left[-\frac{\left(\Psi(t)-\Psi(0)\right)^{\alpha+4}}{\Gamma(\alpha+5)}\right]\right\} = \frac{\left(\Psi(t)-\Psi(0)\right)^{2\alpha+4}}{\Gamma(2\alpha+5)}$$

$$u_3 = -\frac{\left(\Psi(t)-\Psi(0)\right)^{3\alpha+4}}{\Gamma(3\alpha+5)}.$$

Then

$$u(t) = \left(\Psi(t) - \Psi(0)\right)^{4}\left\{\frac{1}{\Gamma(5)} - \frac{\left(\Psi(t)-\Psi(0)\right)^{\alpha}}{\Gamma(\alpha+5)} + \frac{\left(\Psi(t)-\Psi(0)\right)^{2\alpha}}{\Gamma(2\alpha+5)} - \right.$$

$$\left.\frac{\left(\Psi(t)-\Psi(0)\right)^{3\alpha}}{\Gamma(3\alpha+5)} + \dots\right\}$$

$$= \left(\Psi(t) - \Psi(0)\right)^{4}E_{\alpha,5}\left[-\left(\Psi(t) - \Psi(0)\right)^{\alpha}\right].$$

For the first performance test we set $\alpha = 0.2$ and $\Psi = t^2$. This involves comparing the exact solution with a numerical solution estimated after 3 iterations of PET method. We notice in Fig.1 that the numerical solution is almost identical to the exact solution.

Fig.2 shows the variation of the solution according to the values of the fractional derivative α. Knowing that in this case, we took $\Psi = t^2$.

The main objective of defining Ψ-fractional operators is to obtain a generalized notion that gathers some fractional derivatives in a single definition. In Fig.3, we present the solutions according to different values of the functions Ψ.
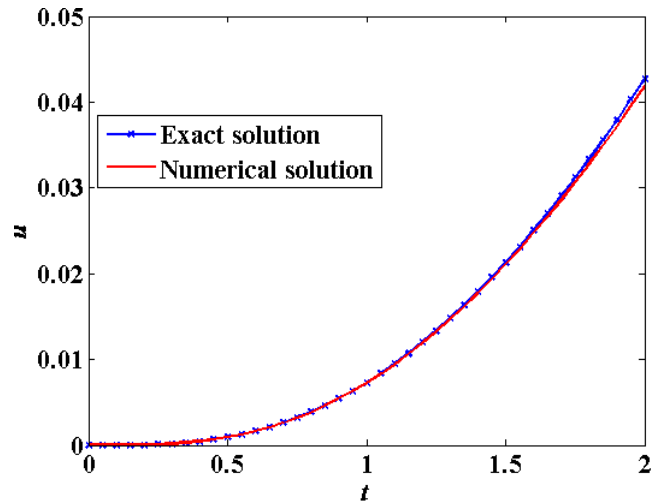

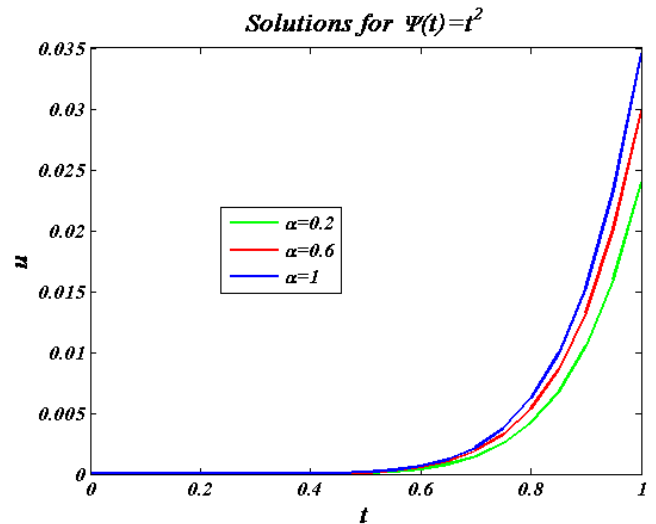**Fig. 1.** Comparison of the exact solution and the numerical solution after 3 iterations of PET method

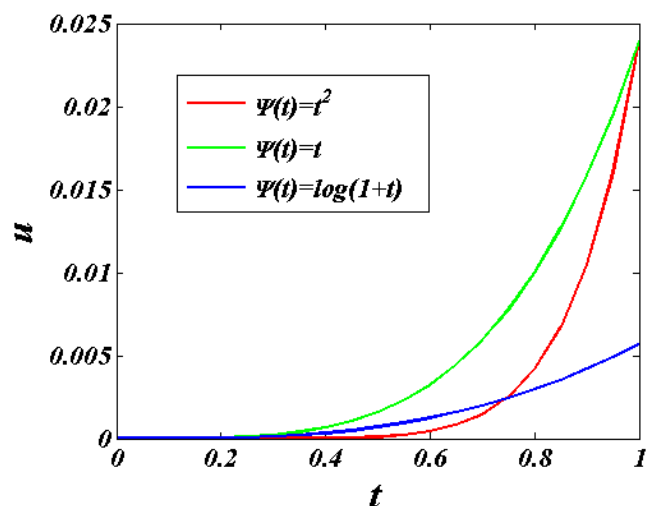**Fig. 2.** Solutions for different values of α and for $\Psi(t) = t^2$

Fig. 2. Solutions for different values of Ψ

Mourad Chamekh, Mohamed Ali Latrach, Tarig M. Elzaki
Novel Integral Transform Treating Some Ψ-Fractional Derivative Equations

DOI 10.2478/ama-2024-0060

Example 3: Considering the following inhomogeneous Ψ-fractional derivative equation:

$$\begin{cases} {}^{C}D_0^{\alpha,\Psi} u(t) - \lambda u(t) = C, \ t \geq 0, and \ 0 < \alpha \leq 1. \\ u(0) = 0. \end{cases} \quad (36)$$

Applying (29), we obtain

$$u(t) = \lambda \mathcal{E}_\Psi^{-1}\{s^\alpha \mathcal{E}[u]\} + C\mathcal{E}_\Psi^{-1}\{s^{\alpha+2}\}.$$

Then

$$\begin{cases} u_0 = C\mathcal{E}_\Psi^{-1}\{s^{\alpha+2}\} = C\frac{(\Psi(t)-\Psi(0))^\alpha}{\Gamma(\alpha+1)}, \\ u_{p+1} = \lambda \mathcal{E}_\Psi^{-1}\{s^\alpha \mathcal{E}[u_p]\}, \quad p \geq 0. \end{cases}$$

Then

$$u_1 = \lambda \mathcal{E}_\Psi^{-1}\{s^\alpha \mathcal{E}[u_0]\} = \lambda C\mathcal{E}_\Psi^{-1}\left\{s^\alpha \mathcal{E}\left[\frac{(\Psi(t)-\Psi(0))^\alpha}{\Gamma(\alpha+1)}\right]\right\}$$

$$= \lambda C\frac{(\Psi(t)-\Psi(0))^{2\alpha}}{\Gamma(2\alpha+1)}.$$

$$u_2 = \lambda \mathcal{E}_\Psi^{-1}\{s^\alpha \mathcal{E}[u_1]\} = \lambda^2 C\frac{(\Psi(t)-\Psi(0))^{3\alpha}}{\Gamma(3\alpha+1)}.$$

And

$$u_3 = \lambda^3 C\frac{(\Psi(t)-\Psi(0))^{4\alpha}}{\Gamma(4\alpha+1)}.$$

Then

$$u(t) = \frac{C}{\lambda}\left\{\lambda\frac{(\Psi(t)-\Psi(0))^\alpha}{\Gamma(\alpha+1)} + \lambda^2 C\frac{(\Psi(t)-\Psi(0))^{2\alpha}}{\Gamma(2\alpha+1)} \right.$$
$$\left. + \lambda^3 C\frac{(\Psi(t)-\Psi(0))^{3\alpha}}{\Gamma(3\alpha+1)} + \cdots\right\},$$

$$= \frac{C}{\lambda}\left(E_\alpha\left(\lambda(\Psi(t)-\Psi(0))^\alpha\right) - 1\right).$$

Example 4: With Ψ-Caputo fractional derivative, we consider the non-linear fractional boundary value problem

$$\begin{cases} {}^{C}D_0^{\alpha,\Psi} u(t) + u^2(t) = 1, \qquad t \geq 0 \ and \ 0 < \alpha \leq 1. \\ u(0) = 0. \end{cases}$$
$$(37)$$

Applying (29), we obtain

$$u(t) = -\mathcal{E}_\Psi^{-1}\{s^\alpha \mathcal{E}[u^2]\} + \mathcal{E}_\Psi^{-1}\{s^{\alpha+2}\}$$

$$= -\mathcal{E}_\Psi^{-1}\{s^\alpha \mathcal{E}[u^2]\} + \frac{(\Psi(t)-\Psi(0))^\alpha}{\Gamma(\alpha+1)}.$$

Using for this example the Adomian decomposition method

$$\begin{cases} u_0 = \frac{(\Psi(t)-\Psi(0))^\alpha}{\Gamma(\alpha+1)}, \\ u_{p+1} = -\mathcal{E}_\Psi^{-1}\{s^\alpha \mathcal{E}[A_p]\}, \quad p \geq 0. \end{cases}$$

With A_p is p-th order of Adomian polynomial of the term u^2. Putting

$$u_0 = \frac{(\Psi(t)-\Psi(0))^\alpha}{\Gamma(\alpha+1)}.$$

$$u_1 = -\mathcal{E}_\Psi^{-1}\{s^\alpha \mathcal{E}[A_0]\} = -\mathcal{E}_\Psi^{-1}\left\{s^\alpha \mathcal{E}\left[\frac{(\Psi(t)-\Psi(0))^{2\alpha}}{\Gamma^2(\alpha+1)}\right]\right\}$$

$$= -\frac{\Gamma(2\alpha+1)}{\Gamma^2(\alpha+1)}\mathcal{E}_\Psi^{-1}\{s^{3\alpha+2}\}$$

$$= -\frac{\Gamma(2\alpha+1)}{\Gamma^2(\alpha+1)\Gamma(3\alpha+1)}\left(\Psi(t)-\Psi(0)\right)^{3\alpha}.$$

$$u_2 = -\mathcal{E}_\Psi^{-1}\{s^\alpha \mathcal{E}[A_1]\} = -\mathcal{E}_\Psi^{-1}\{s^\alpha \mathcal{E}[2u_0 u_1]\}$$

$$= \frac{2\Gamma(2\alpha+1)\Gamma(4\alpha+1)}{\Gamma^3(\alpha+1)\Gamma(3\alpha+1)\Gamma(5\alpha+1)}\left(\Psi(t)-\Psi(0)\right)^{5\alpha}.$$

And

$$u_3 = -\mathcal{E}_\Psi^{-1}\{s^\alpha \mathcal{E}[A_2]\} = -\mathcal{E}_\Psi^{-1}\{s^\alpha \mathcal{E}[u_0 u_2 + u_1^2]\}$$

$$= -\frac{4\Gamma(2\alpha+1)\Gamma(4\alpha+1)\Gamma(6\alpha+1)}{\Gamma^4(\alpha+1)\Gamma(3\alpha+1)\Gamma(5\alpha+1)\Gamma(7\alpha+1)}\left(\Psi(t)-\Psi(0)\right)^{7\alpha}$$

$$- \frac{\Gamma^2(2\alpha+1)\Gamma(6\alpha+1)}{\Gamma^4(\alpha+1)\Gamma^2(3\alpha+1)\Gamma(7\alpha+1)}\left(\Psi(t)-\Psi(0)\right)^{7\alpha}.$$

Note that at $\alpha = 1$ the solution of (37) is given by

$$u(t) = \left(\Psi(t)-\Psi(0)\right) - \frac{\left(\Psi(t)-\Psi(0)\right)^3}{3} + \frac{2\left(\Psi(t)-\Psi(0)\right)^5}{15}$$
$$- \frac{17(\Psi(t)-\Psi(0))^7}{315} + \cdots =$$

$$\tanh\left(\Psi(t)-\Psi(0)\right).$$

6. CONCLUSIONS

This paper discusses and develops a generalized integral transform PET to find exact or approximate solutions to the linear or nonlinear Ψ-fractional differential equations. Contemplating the preliminary, several existing details of fractional calculus can be identified in the literature. The Ψ-fractional differential equations thus proposed in numerical tests were processed by the PETM combined with an iterative method, and this combination was very suitable for dealing with the fractional derivative involving a function Ψ. We conclude the effectiveness and success of this method for solving certain type of linear or nonlinear Ψ-fractional differential equations. Therefore, the proposed method in this paper and its numerical results can stimulates to work for applying on other operators fractional calculus.

REFERENCES

1. Chu Y-M, Rashid S, Karim S, Khalid A, Elagan S-K. Deterministic-stochastic analysis of fractional differential equations malnutrition model with random perturbations and crossover effects. Sci Rep. 2023;13(1):14824.
https://doi: 10.1038/s41598-023-41861-4
2. Al-Qurashi M, Asif Q. U-A, Chu Y-M, Rashid S, Elagan SK. Complexity analysis and discrete fractional difference implementation of the Hindmarsh–Rose neuron system. Results in Physics. 2023;51 106627:2211-3797. https://doi.org/10.1016/j.rinp.2023.106627
3. Alsharidi AK, Rashid S, Elagan SK. Short-memory discrete fractional difference equation wind turbine model and its inferential control of a chaotic permanent magnet synchronous transformer in time-scale analysis. AIMS Mathematics. 2023;8(8):19097-19120.
https://doi.10.3934/math.2023975
4. Kanan M, Ullah H, Raja M-A. Z, Fiza M, Ullah H, Shoaib M., Akgül A, Asad J. Intelligent computing paradigm for second-grade fluid in a rotating frame in a fractal porous medium. Fractals. 2023;31(08): 2340175. https://doi.org/10.1142/S0218348X23401758

6. Rashid S, Noorb MA, Noor K. I. Caputo fractional derivatives and inequalities via preinvex stochastic processes, Published by Faculty of Sciences and Mathematics. University of Nis. Serbia. Filomat. 2023;37(19):6569–6584.
https://doi.org/10.2298/FIL2319569R

7. Li W, Farooq U, Waqas H, Alharthi AM, Fatima N, Hassan AM, Muhammad T, Akgül A. Numerical simulations of Darcy-forchheimer flow of radiative hybrid nanofluid with Lobatto-IIIa scheme configured by a stretching surface. Case Studies in Thermal Engineering. 2023;49:103364:214-157X.
https://doi.org/10.1016/j.csite.2023.103364

8. Faridi WA, Abu Bakar M, Akgül A, Abd El-Rahman M, El Din SM. Exact fractional soliton solutions of thin-film ferroelectric matrial equation by analytical approaches. Alexandria Engineering Journal. 2023;78:483-497.
https://doi.org/10.1016/j.aej.2023.07.049

9. Ashraf R, Hussain S, Ashraf F, Akgül A, El Din SM. The extended Fan's sub-equation method and its application to nonlinear Schrödinger equation with saturable nonlinearity. Results in Physics. 2023;52:106755
https://doi.org/10.1016/j.rinp.2023.106755

10. Khan SA, Yasmin S, Waqas H, Az-Zo'bi EA, Alhushaybari A, Akgül A, Hassan A. M, Imran M. Entropy optimized Ferro-copper/blood based nanofluid flow between double stretchable disks: Application to brain dynamic. Alexandria Engineering Journal. 2023;79:296-307.
https://doi.org/10.1016/j.aej.2023.08.017

11. Faridi WA, Abu Bakar M, Myrzakulova Z, Myrzakulov R, Akgül A, El Din S. M. The formation of solitary wave solutions and their propagation for Kuralay equation. Results in Physics. 2023;52:106774.
https://doi.org/10.1016/j.rinp.2023.106774

12. Rashid S, Karim S, Akgül A, Bariq A, Elagan SK. Novel insights for a nonlinear deterministic-stochastic class of fractional-order Lassa fever model with varying kernels. Sci Rep 2023;13:15320.
https://doi.org/10.1038/s41598-023-42106-0

13. Zhou S-S, Rashid S, Set E, Garba Ahmad A, Hamed YS. On more general inequalities for weighted generalized proportional Hadamard fractional integral operator with applications. AIMS Mathematics. 2021;6(9):9154–9176.
https://doi.org/1010.3934/math.2021532

14. Rashid S, Abouelmagd E. I, Sultana S, Chu Y-M. New developments in weighted n-fold type inequalities via discrete generalized $\widehat{\hbar}$-proportional fractional operators. Fractals. 2022; 30(02):2240056.
https://doi.org/10.1142/S0218348X22400564

15. Rashid S, Abouelmagd E. I, Khalid A, Farooq FB, Chu Y-M. Some recent developments on dynamical $\widehat{\hbar}$-discrete fractional type inequalities in the frame of nonsingular and nonlocal kernels. Fractals. 2022; 30 (02):2240110.
https://doi.org/10.1142/S0218348X22401107

16. Rashid S, Sultana S, Hammouch Z, Jarad F, Hamed YS. Novel aspects of discrete dynamical type inequalities within fractional operators having generalized $\widehat{\hbar}$-discrete Mittag-Leffler kernels and application. Chaos. Solitons & Fractals. 2021;151:111204.
https://doi.org/10.1016/j.chaos.2021.111204

17. Atangana A, Baleanu D. New fractional derivatives with nonlocal and non-singular kernel: theory and application to heat transfer model. Therm. Sci. 2016;20(2):763--769.
http://dx.doi.org/10.2298/TSCI160111018A

18. Chu Y-M, Rashid S, Asif Q. U-A, Abdalbagi M. On configuring new choatic behaviours for a variable fractional-order memristor-based circuit in terms of Mittag-Leffler kernel. Results in Physics. 2023;53:106939. https://doi.org/10.1016/j.rinp.2023.106939

19. Rashid S, Khalid A, Bazighifan O, Oros G.I. New Modifications of Integral Inequalities via $\wp$-Convexity Pertaining to Fractional Calculus and Their Applications. Mathematics. 2021;9:1753.
https://doi.org/10.3390/math9151753

20. Érdlyi A. An integral equation involving Legendre functions, J. Soc. Indust. Appl. Math. 1964;12(1):15-30.
https://doi.org/10.1137/0112002

21. OSLR TJ. Leibniz rule for fractional derivatives and an application to infinite series. SIAM J. Appl. Math. 1970;18(3):658--674.
https://doi.org/10.1137/0118059

22. Almeida R. A caputo fractional derivative of a function with respect to another function. Communications in Nonlinear Science and Numerical Simulation. 2017;44:460--481.
https://doi.org/10.1016/j.cnsns.2016.09.006

23. Almeida R. Further properties of Osler's generalized fractional integrals and derivatives with respect to another function. Rocky Mountain J. Math. 2019;49(8):2459--2493.
https://doi.org/10.1216/RMJ-2019-49-8-2459

24. Sousa JV da C, Oliveira EC de. On the Ψ-Hilfer Fractional Derivative. Commun. Nonlinear Sci. Numer. Simul. 2018;60:72-91.
https://doi.org/10.1016/j.cnsns.2018.01.005

25. Yang X-J. General fractional derivatives: theory, methods and applications. CRC Press. New York 2019.
https://doi.org/10.1201/9780429284083

26. Jarad F, Abdeljawad T. Generalized fractional derivatives and Laplace transform. Discrete Contin. Dyn. Syst. 2020;13(3):709–722.
https://doi.org/10.3934/dcdss.2020039

27. Singh Y, Gill V, Kundu S, Kumar D. On the Elzaki transform and its application in fractional free electron laser equation. Acta Univ. Sapientiae Mathem. 2019;11(2):419--429.
https://doi.org/10.2478/ausm-2019-0030

28. Elzaki TM. The New Integral Transform (Elzaki Transform) fundamental properties investigations and applications. GJPAM. 2011;7(1):57—64.

29. Almeida R, Malinowska AB, Odzijewicz T. An extension of the fractional gronwall inequality, in Conference on Non-Integer Order Calculus and Its Applications. Springer. 2018:20-28.
https://doi.org/10.1007/978-3-030-17344-9_2

30. Ali A, Minamoto T. A new numerical technique for investigating boundary value problems with Ψ-Caputo fractional operator. Journal of Applied Analysis & Computation. 2023;13(1):275--297.
https://doi.org/10.11948/20220062

31. Sousa JV da C, Oliveira E C de. On the Ψ-fractional integral and applications. Comp. Appl. Math. 2019;38(4).
https://doi.org/10.1007/s40314-019-0774-z

32. Bulut H, Baskonus HM, Bin Muhammad Belgacem F. The Analytical Solutions of Some Fractional Ordinary Differential Equations By Sumudu Transform Method. Abs. Appl. Anal. 2013;2013(6):203875.
https://doi.org/10.1155/2013/203875

33. Jafari H. A new general integral transform for solving integral equations. J Adv Res. 2021;32:133--138.
https://doi.org/10.1016/j.jare.2020.08.016

34. Elzaki MT, Chamekh M. Solving nonlinear fractional differential equations using a new decomposition method. Universal Journal of Applied Mathematics & Computation. 2018;6:27-35.

35. Fahad HM, Ur Rehman M, Fernandez A. On Laplace transforms with respect to functions and their applications to fractional differential equations. Math. Methods Appl. Sci. 2021;1-20.
https://doi.org/10.1002/mma.7772

36. Prabhakar TR. A singular integral equation with a generalized Mittag-Leffler function in the kernel. Yokohama Math. J. 1971;19:7-15.

37. Pang D, Jiang W, Niazi AUK. Fractional derivatives of the generalized Mittag-Leffler functions. Adv. Differ. Equ. 2018;2018:415.
https://doi.org/10.1186/s13662-018-1855-9

38. Harikrishnan S, Shah K, Baleanu D, et al. Note on the solution of random differential equations via Ψ-Hilfer fractional derivative. Adv Differ Equ. 2018;2018:224.
https://doi.org/10.1186/s13662-018-1678-8

39. Li C, Zeng FH. Numerical methods for fractional calculus. Chapman and Hall/CRC 2015. https://doi.org/10.1201/b18503

Mourad Chamekh: https://orcid.org/0000-0002-6335-0804

Mohamed Ali Latrach: https://orcid.org/0000-0001-9274-1014

Tarig M. Elzaki: https://orcid.org/0000-0002-6946-9267