

PROGNOZOWANIE NA PODSTAWIE MODELI TRENDU

redaktor naukowy Joanicjusz Nazarko

Redaktor naukowy
Joanicjusz Nazarko

PROGNOZOWANIE
W ZARZĄDZANIU
PRZEDSIĘBIORSTWEM

Cz. IV
Prognozowanie
na podstawie
modeli trendu



Oficyna Wydawnicza Politechniki Białostockiej

Białystok 2018

Recenzenci:

dr hab. Stanisław Łobejko, prof. SGH
prof. dr hab. Dorota Witkowska

Redaktor naukowy:

Joanicjusz Nazarko

Autorzy:

Ewa Chodakowska
Katarzyna Halicka
Arkadiusz Jurczuk
Joanicjusz Nazarko

Redaktor wydawnictwa:

Elżbieta Dorota Alicka

© Copyright by Politechnika Białostocka, Białystok 2017

ISBN 978-83-65596-33-8 ISBN 978-83-65596-34-5 (eBook)



Publikacja jest udostępniona na licencji
Creative Commons Uznanie autorstwa-Użycie niekomercyjne-Bez utworów zależnych 4.0
(CC BY-NC-ND 4.0)

Pełna treść licencji dostępna na stronie
creativecommons.org/licenses/by-nc-nd/4.0/legalcode.pl

Publikacja jest dostępna w Internecie na stronie Oficyny Wydawniczej PB

Redakcja techniczna, skład:

Oficyna Wydawnicza Politechniki Białostockiej

Druk:

„UNI-DRUK” Wydawnictwo i Drukarnia Sp. J.

Nakład: 63 egz.

Oficyna Wydawnicza Politechniki Białostockiej
ul. Wiejska 45C, 15-351 Białystok
tel.: 85 746 91 37, fax: 85 746 90 12
e-mail: oficyna.wydawnicza@pb.edu.pl
www.pb.edu.pl

SPIS TREŚCI

WPROWADZENIE	5
1. MODELE ANALITYCZNE	7
1.1. Specyfikacja	11
1.2. Estymacja parametrów funkcji trendu	13
1.3. Predykcja	19
1.4. Nieliniowe funkcje trendu	22
1.4.1. Funkcje nieliniowe o przyspieszonym tempie wzrostu ..	23
1.4.2. Funkcje nieliniowe o malejącym tempie wzrostu	37
1.4.3. Funkcje nieliniowe o zmiennym tempie wzrostu	57
2. METODA WSKAŹNIKÓW SEZONOWOŚCI	65
3. ANALIZA HARMONICZNA	84
4. OCENA MODELU ORAZ DOPUSZCZALNOŚCI	
PROGNOZ	108
4.1. Ocena dopasowania modelu	108
4.2. Ocena dopuszczalności prognoz	112
5. PRZYKŁADY ZASTOSOWAŃ	136
5.1. Prognoza liczby gospodarstw domowych w Polsce posiadających dostęp do szerokopasmowego Internetu (wyrażona w procentach ogółu gospodarstw domowych)	136
5.2. Prognoza zainteresowania turystyką online na podstawie liczby zapytań w Google Trends	156
5.3. Prognoza wskaźnika ogólnego klimatu koniunktury w budownictwie	166
6. PROBLEMY DO SAMODZIELNEGO ROZWIĄZANIA	178
LITERATURA	180

WPROWADZENIE

Tematem czwartej części podręcznika *Prognozowane w zarządzaniu przedsiębiorstwem* jest prognozowanie z zastosowaniem klasycznych modeli składowych szeregów czasowych (tendencji rozwojowej, wahań periodycznych oraz wahań przypadkowych). Autorzy wybrali do zaprezentowania w podręczniku modele najczęściej wykorzystywane w prognozowaniu zjawisk gospodarczych: analityczne, wskaźników sezonowości oraz analizy harmonicznej. W opracowaniu omówiono także zagadnienie dopuszczalności prognoz.

Trend (tendencja rozwojowa) szeregu czasowego wyraża długookresową skłonność do monotonicznych (jednokierunkowych) zmian prognozowanej zmiennej. Trend jest wynikiem trwałego oddziaływania na dane zjawisko gospodarcze ustalonej kompozycji czynników, zarówno o charakterze obiektywnym, jak i subiektywnym. W modelach tendencji rozwojowej szereg czasowy jest przybliżany monotoniczną ze względu na czas funkcją o określonym wzorze analitycznym. Funkcję tę określa się mianem funkcji trendu. Modelu trendu stosuje się do prognozowania zjawisk gospodarczych opisywanych za pomocą szeregów czasowych, w których występują tendencja rozwojowa oraz wahania przypadkowe.

Wahania periodyczne szeregu czasowego przejawiają się w postaci okresowych wahań wartości prognozowanej zmiennej wokół jej tendencji rozwojowej. Wahania periodyczne odzwierciedlają na przykład zmiany nasilenia działalności gospodarczej związane bezpośrednio lub pośrednio z porami roku i kalendarzem (wahania sezonowe). Wahania periodyczne wynikać mogą też ze zmian aktywności gospodarczej, spowodowanej na przykład rytmem biologicznym człowieka czy też harmonogramem pracy (oscylacje krótkookresowe). Metodę wskaźników sezonowości stosuje się do prognozowania zjawisk gospodarczych, w których występują wahania sezonowe o znanym okresie, tendencja rozwojowa oraz wahania przypadkowe. Analiza harmoniczna przydatna jest szczególnie w sytuacjach, w których nie jest znany *a priori* okres wahań periodycznych, a także wówczas, gdy w szeregu czasowym mogą występować jednocześnie wahania periodyczne o różnych okresach.

Zaprezentowane metody charakteryzują się dużą przydatnością praktyczną i mogą być użyteczne w prognozowaniu wielu zjawisk go-

spodarczych o dosyć złożonej naturze, w których rolę zmiennej objaśniającej pełni zmienna czasowa, syntetyzująca wpływ wielu (często bliżej nieznanym) czynników oddziałujących na prognozowane zjawisko.

Oprócz wymienionych modeli składowych szeregów czasowych i ich zastosowania w prognozowaniu w podręczniku przedstawiono także zagadnienia związane z oceną dopuszczalności prognoz. Jest to jeden z zasadniczych aspektów procesu prognozowania. Dopuszczalność prognozy wyraża wystarczający poziom zaufania odbiorcy prognozy co do możliwości jej zastosowania w procesie decyzyjnym i związana jest przede wszystkim z oceną jej błędu *ex ante*. W podręczniku omówiono również inne kryteria dopuszczalności prognozy: prawdopodobieństwo realizacji, przedział ufności oraz ekspercką ocenę wiarygodności prognozy.

Intencją autorów było, aby zagadnienia podjęte w podręczniku opisane zostały w sposób możliwie przystępny i zrozumiały zarówno dla studentów podstawowego kursu prognozowania, jak i dla szerszego grona odbiorców pragnących poznać zasadnicze zagadnienia i metody prognozowania. Autorzy starali się zachować niezbędną ścisłość i przejrzystość wywodów, unikając jednocześnie zawilego dyskursu naukowego i subtelności teoretycznych. Wiele uwagi poświęcili sposobowi prezentacji materiału dydaktycznego. Zamieścili wiele przykładów obliczeniowych szczegółowo objaśniających sposób sporządzania prognoz z zastosowaniem omawianych metod i wskazujących na możliwości ich wykorzystania w procesach podejmowania decyzji menadżerskich. Akcentowali te zagadnienia, które – ich zdaniem – są szczególnie istotne w praktyce przedsiębiorstw. Intensywnie wykorzystali schematy graficzne ilustrujące przedstawiane kwestie. Każdy rozdział kończy wykaz kluczowych zagadnień oraz spis literatury podstawowej i uzupełniającej, przydatnej do samodzielnego studiowania i poszerzania wiedzy z zakresu omawianych zagadnień. Podręcznik zamyka rozdział zawierający problemy do samodzielnego rozwiązania, które mogą pomóc czytelnikowi ugruntować zdobytą wiedzę.

Autorzy podręcznika wyrażają nadzieję, że będzie on przydatną pomocą w studiowaniu zagadnień prognozowania, a jednocześnie czytelnicy znajdą w nim źródło wiedzy i inspirację do własnych badań prognostycznych. Będą jednocześnie wdzięczni za wszelkie sugestie i opinie przydatne do udoskonalenia ewentualnych następnych wydań i kolejnych części podręcznika.

1. MODELE ANALITYCZNE

Znajomość kształtowania się tendencji rozwojowej określonego zjawiska może być wykorzystywana zarówno do opisu jego przeszłości, jak i prognozowania¹. Modele tendencji rozwojowej stosuje się między innymi do prognozowania przyszłych wartości szeregów czasowych, w których występuje tendencja rozwojowa oraz wahania przypadkowe. Rolę zmiennej objaśniającej odgrywa zmienna czasowa.

Wykorzystanie modeli tendencji rozwojowej do prognozowania sprowadza się do znalezienia funkcji czasu $f(t)$, która w sposób analityczny opisuje dane zjawisko. W metodzie tej zakłada się, że kształtowanie się zmiennej prognozowanej y w czasie można opisać z dokładnością do składnika (czynnika) losowego ξ za pomocą funkcji²:

$$y_t = f(t) + \xi_t, \quad t = 1, \dots, n \text{ (model addytywny)} \quad (1.1)$$

lub

$$y_t = f(t) \cdot \xi_t, \text{ (model multiplikatywny)} \quad (1.2)$$

gdzie:

y_t – wartość zmiennej prognozowanej w momencie lub okresie t ,

$f(t)$ – funkcja czasu charakteryzująca tendencję rozwojową,

ξ – zmienna losowa charakteryzująca efekty oddziaływań wahań przypadkowych.

W modelu addytywnym (1.1) zakłada się, że zmienna losowa ma wartość oczekiwaną równą zero i skończoną wariancję:

$$E(\xi) = 0, \quad (1.3)$$

$$V(\xi) = \sigma^2. \quad (1.4)$$

¹ J. Nazarko (red.), *Prognozowanie w zarządzaniu przedsiębiorstwem. Cz. II. Prognozowanie na podstawie szeregów czasowych*, Wydawnictwo Politechniki Białostockiej, Białystok 2004.

² M. Cieślak (red.), *Prognozowanie gospodarcze. Metody i zastosowania*, Wydawnictwo Naukowe PWN, Warszawa 2011, s. 76.

Natomiast w wypadku modelu multiplikatywnego (1.2) zakłada się, że zmienna losowa ma wartość oczekiwaną równą jeden i skończoną wariancję:

$$E(\xi) = 1, \quad (1.5)$$

$$V(\xi) = \sigma^2. \quad (1.6)$$

Model (1.1) opisuje sytuację, w której efekty działania składnika losowego nakładają się addytywnie na tendencję rozwojową szeregu. Model taki określa się wówczas jako addytywny. Model opisany równaniem (1.2) odpowiada sytuacji, gdy efekty działania składnika losowego nakładają się multiplikatywnie na tendencję rozwojową szeregu, co jest charakterystyczne dla modelu multiplikatywnego.

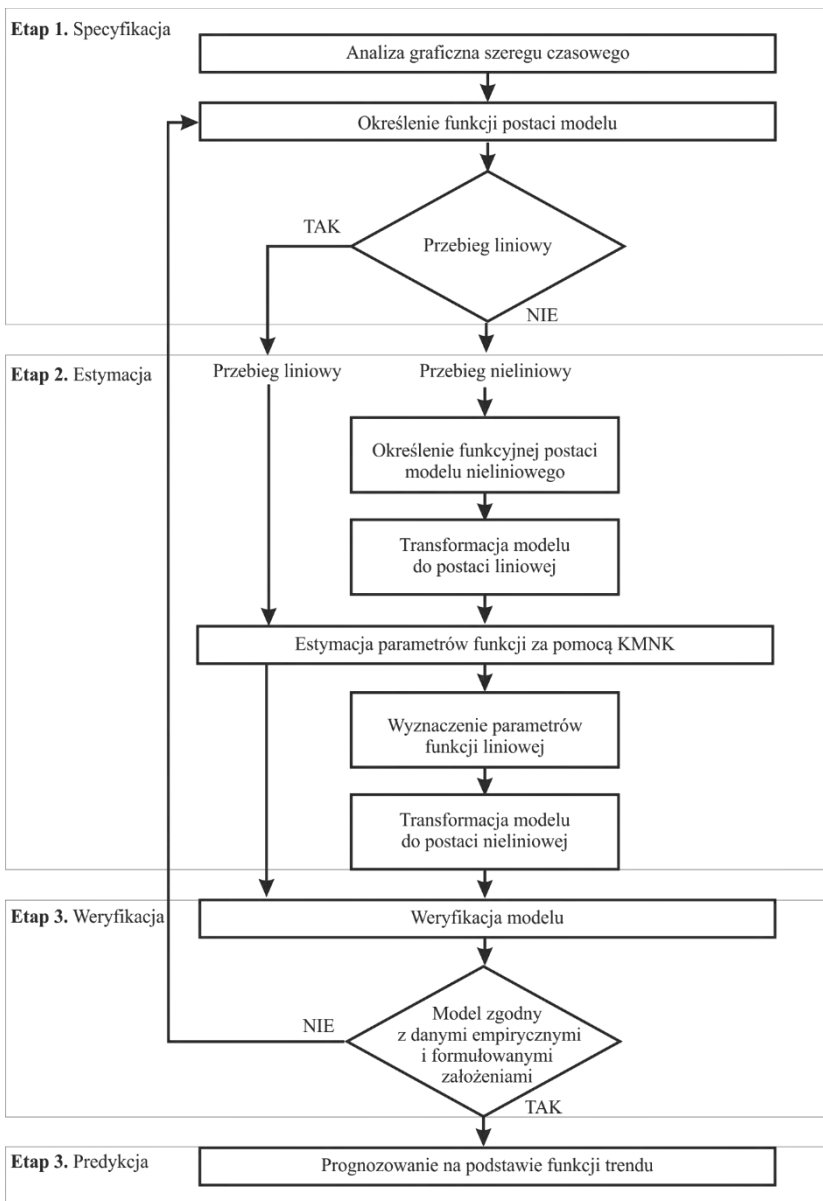
Funkcję $f(t)$ można wyznaczyć, wykorzystując modele analityczne lub modele adaptacyjne³. Modele adaptacyjne zostały szczegółowo omówione w trzeciej części niniejszego podręcznika⁴.

W wypadku modeli analitycznych przybliża się wartość szeregu czasowego pewną funkcją matematyczną. Modele te stosuje się głównie do prognozowania zjawisk, które charakteryzowały się w przeszłości regularnymi zmianami, dającymi się opisać za pomocą funkcji czasu. Zakłada się niezmiennosć kierunku trendu (rosnący, malejący) oraz stałość charakteru zmian zjawiska w rozpatrywanym czasie, wyrażoną poprzez niezmiennosć postaci analitycznej funkcji trendu. Przyjmuje się również, że wahania przypadkowe nie wpływają w istotny sposób na charakter badanego zjawiska⁵.

³ K. Halicka, J. Godlewska, *Prognozowanie średniego miesięcznego kursu kupna USD*, *Ekonomia i Zarządzanie*, 2013, t. 5, nr 1; t. 4, nr 2, Oficyna Wydawnicza Politechniki Białostockiej, Białystok; s. 19-29; K. Halicka, C. Wienkowski, *Wykorzystanie metod wykładniczego do prognozowania kursu sprzedaży EUR*, *Ekonomia i Zarządzanie* 2013, t. 5, nr 2, Oficyna Wydawnicza Politechniki Białostockiej, Białystok 2013 s. 70-80.

⁴ J. Nazarko (red.), *Prognozowanie w zarządzaniu przedsiębiorstwem. Cz. III. Prognozowanie na podstawie modeli adaptacyjnych*, Wydawnictwo Politechniki Białostockiej, Białystok 2005.

⁵ J. Nazarko (red.), *Prognozowanie w zarządzaniu przedsiębiorstwem. Cz. I. Wprowadzenie do metodyki prognozowania*, Wydawnictwo Politechniki Białostockiej, Białystok 2004.



Rys. 1.1. Schemat procesu budowy prognozy z wykorzystaniem modeli analitycznych

Źródło: opracowanie własne.

Modele tendencji rozwojowej, w tym modele analityczne, są powszechnie wykorzystywane do prognozowania wielu zjawisk ekonomicznych. Znajdują one zastosowanie głównie do ustalania prognoz krótko- i średniookresowych.

Prognozowanie na podstawie modelu analitycznych przebiega w czterech podstawowych etapach:

- 1) specyfikacja,
- 2) estymacja,
- 3) weryfikacja,
- 4) predykcja.

Schemat procesu budowy prognozy z wykorzystaniem funkcji analitycznych został przedstawiony na rysunku 1.1.

Etap pierwszy polega na analizie graficznej szeregu i określeniu charakteru prawidłowości zmian poziomu badanego zjawiska w czasie. Etap ten kończy się zaproponowaniem potencjalnej postaci funkcji trendu najlepiej dopasowanej do wartości, które badana wielkość przyjmowała w kolejnych okresach, i określeniem sposobu wprowadzenia składnika losowego do modelu (addytywnego lub multiplikatywnego).

Etap drugi polega na estymacji parametrów wybranej funkcji. W modelach liniowych oraz modelach, które można sprowadzić poprzez transformację do postaci liniowej, stosuje się najczęściej klasyczną metodę najmniejszych kwadratów (KMNK).

W etapie trzecim przeprowadzana jest weryfikacja modelu, polegająca na zbadaniu jego zgodności z danymi empirycznymi i z sformułowanymi założeniami dotyczącymi kształtowania się analizowanego zjawiska w czasie. Zagadnienie to zostanie szerzej omówione w rozdziale 4.

W etapie czwartym dokonuje się predykcji. Zweryfikowany model może być wykorzystywany do prognozowania, przy założeniu, że postać analityczna modelu oraz wartości ocen jego parametrów, a także rozkład składnika losowego nie ulegną zmianie w rozpatrywanym horyzoncie czasowym.

Omówione etapy powinny być traktowane w sposób integralny.

1.1. Specyfikacja

W badaniach ekonomicznych można posłużyć się różnymi postaciami analitycznymi funkcji trendu $f(t)$ w zależności od charakteru zmiennej prognozowanej, dotychczasowego przebiegu jej realizacji oraz założeń co do przewidywanego kształtowania się tej zmiennej w przyszłości.

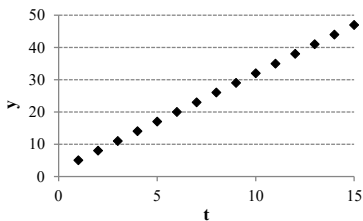
Wybór postaci modelu jest jednym z trudniejszych etapów budowy modelu. Nie powstała do tej pory uniwersalna i w pełni obiektywna metoda wyboru. Wybór właściwego kształtu funkcji jest kwestią intuicji badacza oraz rozeznania w rozwoju analizowanego zjawiska⁶. Przy wyborze określonej postaci analitycznej modelu mającego opisać zależność danego zjawiska ekonomicznego od czasu często korzysta się z wiedzy ekonomicznej o badanych prawidłowościach. Istnieją teorie ekonomiczne, które dotyczą zachowania się gospodarki narodowej, rynku czy przedsiębiorstw, które mogą być podstawą sformułowania założenia, że pewna zależność może być w przybliżeniu opisana za pomocą funkcji o określonej postaci analitycznej. Na przykład, jeśli wiedza o badanym zjawisku wskazuje, że elastyczność zmiennej objaśnianej y względem zmiennej objaśniającej t jest stała, to model powinien mieć postać potęgową⁷.

Często wykorzystywaną procedurą określania postaci analitycznej trendu jest również metoda empiryczna. Metoda ta polega na budowie kilku różnych modeli tendencji rozwojowej i wyborze tego, który jest najlepiej dopasowany do danych empirycznych (szerzej w rozdziale 4). W literaturze przedmiotu dotyczącej procedur określania postaci analitycznej trendu można spotkać się również z metodą krzywych wzrostu, analizy dynamicznych własności równania trendu lub metodą uśrednionych gradientów⁸.

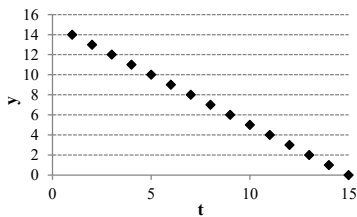
⁶ T. Stanisławski, *Funkcje jednej zmiennej w badaniach ekonomicznych*, Wydawnictwo Naukowe PWN, Warszawa 1993, s. 286.

⁷ E. Nowak, *Zarys metod ekonometrii. Zbiór zadań*, Wydawnictwo Naukowe PWN, Warszawa 2002, s. 65.

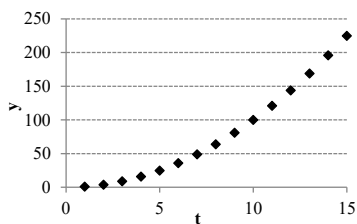
⁸ T. Stanisławski, *Funkcje jednej zmiennej...*, op. cit., s. 287.



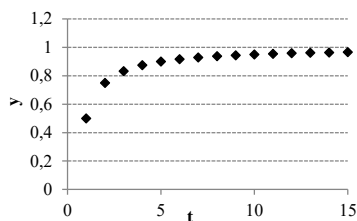
a) tendencja wzrostowa, związek liniowy



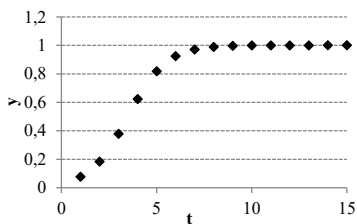
b) tendencja spadkowa, związek liniowy



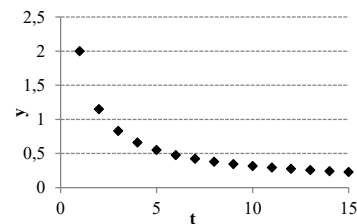
c) tendencja wzrostowa, przyspieszone tempo wzrostu, związek nieliniowy



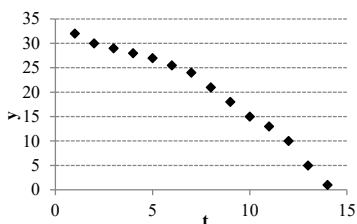
d) tendencja wzrostowa, malejące tempo wzrostu, związek nieliniowy



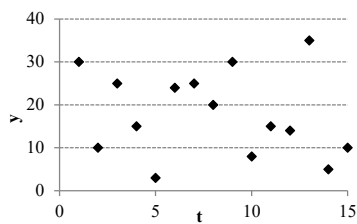
e) tendencja wzrostowa, zmienne tempo wzrostu (początkowo przyspieszone tempo wzrostu, później lekkie tempo wzrostu), związek nieliniowy



f) tendencja malejąca, malejące tempo spadku, związek nieliniowy



g) tendencja malejąca, przyspieszone tempo spadku, związek nieliniowy



h) brak związku funkcyjnego między zmiennymi

Rys. 1.2. Typowe postacie związków między zmienną obserwowaną y a czasem t

Źródło: opracowanie własne.

Często wyboru postaci analitycznej modelu dokonuje się, wykorzystując metodę graficzną. Początkowo szereg czasowy przedstawiany jest w układzie współrzędnych. Następnie dobór postaci modelu odbywa się na podstawie oceny wzrokowej rozstawienia punktów, które odpowiadają wynikom obserwacji zmiennej y . Analiza wykresu, wsparta znajomością przebiegu określonych funkcji, pozwala sformułować postać analityczną funkcji trendu. Najczęściej spotykane postacie funkcji trendu zostały przedstawione na rysunku 1.2.

Analizując rysunki 1.2 (a) i (b) można zauważyć związek liniowy pomiędzy zmienną obserwowaną y a czasem t . Z kolei na rysunkach 1.2 (c), (d), (e), (f) i (g) przedstawiono związki nieliniowe. Natomiast na rysunku 1.2 (h) widoczny jest brak związku funkcyjnego między zmiennymi. Funkcje analityczne o rosnącym tempie wzrostu zostały przedstawione na rysunkach 1.2 (a), (c), (d) i (e), a o malejącym tempie wzrostu na rysunkach 1.2 (b), (f) i (g).

Najprostszą funkcją trendu jest funkcja liniowa:

$$y_t = \beta_0 + \beta_1 t, \quad (1.7)$$

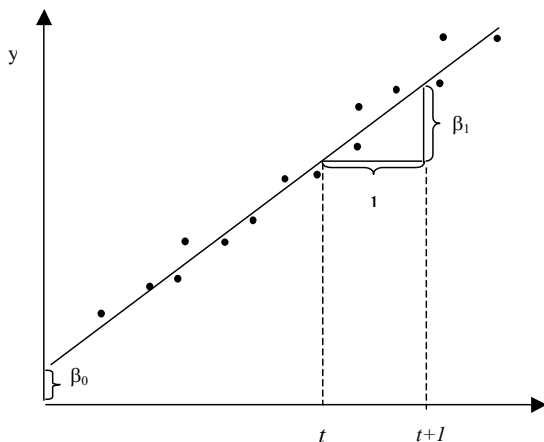
gdzie:

β_0 – wyraz wolny, parametr reprezentujący średni poziom zmiennej prognozowanej;

β_1 – współczynnik wyrażający przyrost wartości zmiennej prognozowanej w ciągu jednostki czasu (z okresu na okres).

1.2. Estymacja parametrów funkcji trendu

Funkcja liniowa (patrz rys. 1.2a) reprezentuje stały kierunek rozwoju danego zjawiska, określony przez współczynnik kierunkowy prostej β_1 . Graficzną postać funkcji liniowej wraz z interpretacją jej parametrów przedstawiono na rysunku 1.3.



Rys. 1.3. Graficzna prezentacja postaci funkcji liniowej

Źródło: opracowanie własne na podstawie: A.D. Aczel, *Statystyka w zarządzaniu*, Wydawnictwo Naukowe PWN, Warszawa 2000, s. 461.

Do szacowania parametrów modelu linowego można wykorzystać *klasyczną metodę najmniejszych kwadratów* (KMNK). Idea tej metody sprowadza się do takiego wyznaczenia wartości oszacowań $\hat{\beta}_0, \hat{\beta}_1, \dots, \hat{\beta}_k$ parametrów $\beta_0, \beta_1, \dots, \beta_k$, aby suma kwadratów odchylenia wartości teoretycznych \hat{y}_t od wartości empirycznych y_t była najmniejsza. Warunek ten można zapisać w następujący sposób:

$$\sum_{t=1}^n e_t^2 = \sum_{t=1}^n (y_t - \hat{y}_t)^2 \rightarrow \min, \quad (1.8)$$

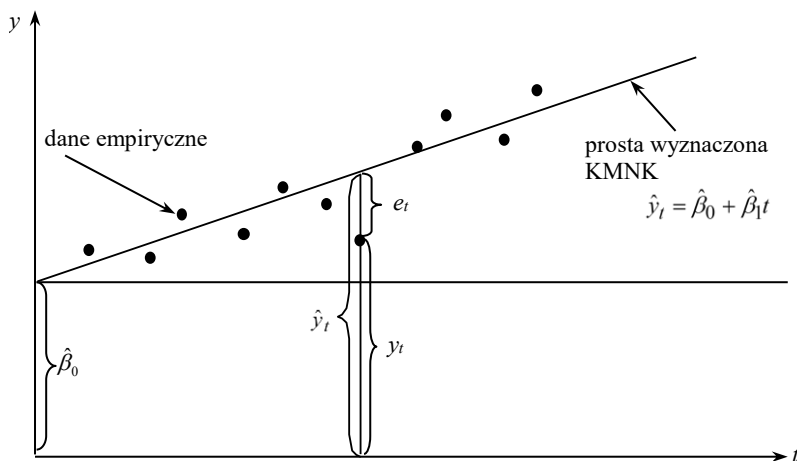
gdzie:

- e_t – bezwzględny błąd prognozy *ex post* w okresie t ,
- y_t – empiryczna wartość zmiennej prognozowanej w momencie lub okresie t ,
- \hat{y}_t – wartość teoretyczna zmiennej y wyznaczona na moment lub okres t .

Zastosowanie metody najmniejszych kwadratów wymaga przyjęcia następujących założeń⁹:

- 1) szacowany model jest modelem liniowym, bądź sprowadzalnym do postaci liniowej;
- 2) zmienne objaśniające są wielkościami nielosowymi o ustalonych elementach;
- 3) nie występuje zjawisko współliniowości zmiennych objaśniających;
- 4) składnik losowy ma wartość oczekiwaną równą zero i stałą skończoną wariancję;
- 5) nie występuje zjawisko autokorelacji składnika losowego, czyli zależność wartości składnika losowego w czasie.

Graficzną ilustrację idei wyznaczenia sumy kwadratów reszt na podstawie znajomości reszt modelu przedstawiono na rysunku 1.4.



Rys. 1.4. Graficzna prezentacja wyznaczenia sumy kwadratów reszt na podstawie znajomości reszt modelu

Źródło: opracowanie własne na podstawie: A.D Aczel, *Statystyka w zarządzaniu*, Wydawnictwo Naukowe PWN, Warszawa 2000, s. 466.

⁹ E. Nowak, *Zarys metod ...*, op. cit., s. 36.

Wyznaczając parametry β_0 oraz β_1 klasyczną metodą najmniejszych kwadratów, otrzymuje się oszacowania określone wzorami¹⁰:

$$\hat{\beta}_1 = \frac{\sum_{t=1}^n (t - \bar{t}) y_t}{\sum_{t=1}^n (t - \bar{t})^2}, \quad (1.9)$$

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{t}, \quad (1.10)$$

gdzie:

$\hat{\beta}_1$ – wartość oszacowania parametru β_1 ,

$\hat{\beta}_0$ – wartość oszacowania parametru β_0 ,

t – zmienna czasowa,

\bar{t} – średnia arytmetyczna wyznaczona dla zmiennej t ,

y_t – empiryczna wartość zmiennej prognozowanej w momencie lub okresie t ,

\bar{y} – średnia arytmetyczna wyznaczona dla empirycznych wartości zmiennej prognozowanej,

przy czym:

$$\bar{t} = \frac{1}{n} \sum_{t=1}^n t, \quad (1.11)$$

$$\bar{y} = \frac{1}{n} \sum_{t=1}^n y_t. \quad (1.12)$$

Po oszacowaniu wartości parametrów modelu β_j metodą najmniejszych kwadratów zaleca się sprawdzenie ich statystycznej istotności. Badanie to polega na wykonaniu testu t -Studenta weryfikującego istotność parametru β_j . Jeżeli parametr β_j nie jest istotnie różny od zera, to zmienna przy danym parametrze nie ma istotnego wpływu na zmienną objaśnianą y .

¹⁰ M. Cieślak (red.), *Prognozowanie gospodarcze...*, op. cit., s. 81.

Z uwagi na powszechność opisu i zastosowań testu t -Studenta¹¹ weryfikacja istotności parametrów modelu jest przedmiotem wielu komplementarnych podręczników (patrz: Zeliaś i in. 2003, s. 76-79) i dlatego też w tym rozdziale przedstawiono tylko podstawowe wyjaśnienia dotyczące tej kwestii. Zasady estymacji parametrów modelu metodą najmniejszych kwadratów zilustrowano przykładem 1.1.

Przykład 1.1. Estymacja parametrów modelu metodą najmniejszych kwadratów

Należy zbadać kształtowanie się wielkości produkcji energii elektrycznej w Polsce w latach 1996-2010. Zebrane dane przedstawiono w tabeli 1.1.

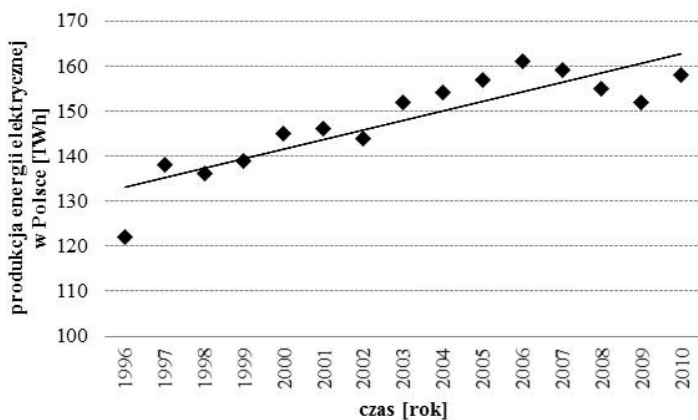
Tab. 1.1. Dane dotyczące produkcji energii elektrycznej w Polsce w latach 1996-2010

Czas	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
Produkcja energii elektrycznej [TWh]	122	138	136	139	145	146	144	152	154	157	161	159	155	152	158

Źródło: opracowanie własne na podstawie Rocznika Statystycznego Przemysłu 2012, dokument elektroniczny http://www.stat.gov.pl/gus/5840_3921_PLK_HTML.html, data wejścia 29.04.2016.

Wykres przebiegu danych oraz linię opisującą tendencję wzrostu produkcji energii elektrycznej w Polsce przedstawiono na rysunku 1.5.

¹¹ Test został opracowany przez absolwenta Oxfordu Williama Gosseta na potrzeby jednego z irlandzkich browarów z Dublina. Z uwagi na problemy publikacyjne Gosset przedstawił jego założenia w 1908 roku w czasopiśmie „Biometrika” (vol. VI, no. 1) pod pseudonimem „Student”. Test ten ma zatem „piwny”, bardzo praktyczny i „studentcki” rodowód.



Rys. 1.5. Produkcja energii elektrycznej w Polsce w latach 1996-2010

Źródło: opracowanie własne.

Tab. 1.2. Obliczenia parametrów modelu metodą najmniejszych kwadratów

t	y_t	$t - \bar{t}$	$(t - \bar{t})^2$	$(t - \bar{t})y_t$
1	122	-7,00	49,00	-854,00
2	138	-6,00	36,00	-828,00
3	136	-5,00	25,00	-680,00
4	139	-4,00	16,00	-556,00
5	145	-3,00	9,00	-435,00
6	146	-2,00	4,00	-292,00
7	144	-1,00	1,00	-144,00
8	152	0,00	0,00	0,00
9	154	1,00	1,00	154,00
10	157	2,00	4,00	314,00
11	161	3,00	9,00	483,00
12	159	4,00	16,00	636,00
13	155	5,00	25,00	775,00
14	152	6,00	36,00	912,00
15	158	7,00	49,00	1106,00
$\bar{t} = 8,000$	$\bar{y} = 147,867$			
Σ			280,000	591,000

Źródło: obliczenia własne.

Na podstawie dekompozycji szeregu można zauważyć, że w szeregu czasowym badanej zmiennej występuje rosnąca tendencja rozwojowa oraz wahania przypadkowe. Dokonując analizy graficznej, zauważono, że linia prosta będzie odpowiednio opisywała tendencję wzrostu produkcji energii elektrycznej w Polsce w miarę upływu czasu. Niezbędne obliczenia pośrednie przedstawione są w tabeli 1.2.

Korzystając z równań (1.9) i (1.10), otrzymano następujące wartości parametrów modelu:

$$\hat{\beta}_1 = \frac{591}{280} = 2,111;$$

$$\hat{\beta}_0 = 147,867 - 2,111 \cdot 8,000 = 130,979.$$

Oszacowane klasyczną metodą najmniejszych kwadratów parametry liniowej funkcji trendu produkcji energii elektrycznej w Polsce można zatem przedstawić ostatecznie w następujący sposób:

$$\hat{y}_t = 130,979 + 2,111 t.$$

Wartość oceny $\beta_1 = 2,111$ wskazuje, że w latach 1980-2010 średni roczny przyrost produkcji energii elektrycznej w Polsce wynosił 2,111 TWh.

1.3. Predykcja

Wyznaczenie prognozy, z wykorzystaniem modeli analitycznych, sprowadza się do obliczenia wartości funkcji dla przyszłego momentu lub okresu czasu¹². Należy jednak dokonać pewnych istotnych założeń, których przyjęcie określa sposób sporządzenia prognozy (ekstrapolacja modelu) i ocenę jej jakości *ex ante*. Założenia te dotyczą¹³:

- stabilności relacji strukturalnych w czasie, co oznacza, że postać analityczna modelu oraz wartości ocen jego parametrów nie ulegną zmianie w przedziale czasu, dla którego wyznacza się prognozę;
- stabilności rozkładu składnika losowego, umożliwiającej ocenę błędu *ex ante* prognozy.

¹² A. Maciąg, R. Pietroń, S. Kukła, *Prognozowanie i symulacje w przedsiębiorstwie*, Polskie Wydawnictwo Ekonomiczne, Warszawa 2013, s. 77.

¹³ M. Cieślak (red.), *Prognozowanie gospodarcze...*, op. cit., s. 81.

Przyjęcie powyższych założeń stanowi jednocześnie akceptację pasywnej postawy prognosty (szerzej w rozdziale 4) i jest zazwyczaj wykorzystywane przy budowie prognoz krótkoterminowych.

Prognozowanie na podstawie modelu tendencji rozwojowej polega na ekstrapolacji funkcji trendu. Oznacza to, że do modelu jako wartość zmiennej czasowej t podstawia się numer momentu lub okresu, na który wyznaczana jest prognoza¹⁴:

$$y_t^* = \hat{f}(t), \quad t > n, \quad (1.13)$$

gdzie:

$\hat{f}(t)$ – estymowana funkcja trendu.

Przewidywaną produkcję energii elektrycznej w Polsce na kolejne lata można więc, wykorzystując równania (1.7) i (1.13) oraz oszacowane parametry β_0 i β_1 , wyznaczyć w następujący sposób:

$$\begin{array}{l}
 t = 1 \quad \left. \begin{array}{l} y_1^* = 130,979 + 2,111 \cdot 1 \\ y_1^* = 133,09 \end{array} \right\} \text{wartości estymowane modelem,} \\
 t = 2 \quad \left. \begin{array}{l} y_2^* = 130,979 + 2,111 \cdot 2 \\ y_2^* = 135,20 \end{array} \right\} \\
 t = 3 \quad \left. \begin{array}{l} y_3^* = 130,979 + 2,111 \cdot 3 \\ y_3^* = 137,31 \end{array} \right\} \\
 \vdots \\
 t = 16 \quad \left. \begin{array}{l} y_{16}^* = 130,979 + 2,111 \cdot 16 = 164,76 \\ y_{17}^* = 130,979 + 2,111 \cdot 17 = 166,87 \\ y_{18}^* = 130,979 + 2,111 \cdot 18 = 168,98 \end{array} \right\} \text{prognozowana wielkość} \\
 t = 17 \quad \left. \begin{array}{l} \\ \\ \end{array} \right\} \text{produkcji energii} \\
 t = 18 \quad \left. \begin{array}{l} \\ \\ \end{array} \right\} \text{elektrycznej w Polsce.}
 \end{array}$$

¹⁴ Ibidem.

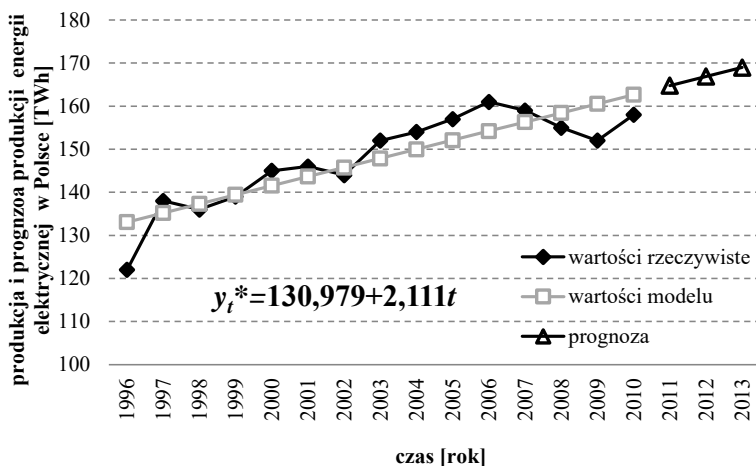
Zatem prognoza produkcji energii elektrycznej w Polsce, sporządzona na koniec 2010 roku według zaproponowanego modelu, na rok 2011 wynosi 164,76 TWh, na rok 2012 – 166,87 TWh, natomiast na rok 2013 – 168,98 TWh. Wyniki obliczeń umieszczono w tabeli 1.3.

Tab. 1.3. Wartości teoretyczne i prognozy produkcji energii elektrycznej w Polsce

<i>t</i>	Lata	Produkcja energii elektrycznej [TWh]	Model i prognoza produkcji energii elektrycznej [TWh]
		y_t	y^*_t
1	1996	122	133,09
2	1997	138	135,20
3	1998	136	137,31
4	1999	139	139,42
5	2000	145	141,53
6	2001	146	143,65
7	2002	144	145,76
8	2003	152	147,87
9	2004	154	149,98
10	2005	157	152,09
11	2006	161	154,20
12	2007	159	156,31
13	2008	155	158,42
14	2009	152	160,53
15	2010	158	162,64
16	2011		164,76
17	2012		166,87
18	2013		168,98

Źródło: obliczenia własne.

Wykres wartości rzeczywistych, wartości modelu oraz prognozę produkcji energii elektrycznej w Polsce przedstawiono na rysunku 1.6.



Rys. 1.6. Wartości rzeczywiste oraz wartości modelu i prognoza produkcji energii elektrycznej w Polsce w latach 1996-2013

Źródło: opracowanie własne.

Na podstawie analizy wizualnej przebiegów przedstawionych na wykresie 1.6 można stwierdzić, że funkcja liniowa o równaniu $y_t^* = 130,979 + 2,111t$ dosyć dobrze odzwierciedla dane rzeczywiste.

1.4. Nieliniowe funkcje trendu

W opisie zjawisk ekonomicznych, przy których stosowanie liniowych funkcji trendu słabo odzwierciedla rzeczywisty przebiegu tych zjawisk, można skorzystać z innych analitycznych postaci funkcji trendu. W tym rozdziale zostaną przedstawione funkcje nieliniowe, które po pewnych przekształceniach można sprowadzić do postaci liniowej względem parametrów modelu. Omówione zostaną funkcje o przyspieszonym tempie wzrostu, funkcje o malejącym tempie wzrostu oraz funkcje o zmiennym tempie wzrostu (logistyczne).

Nieliniowy związek między zmiennymi i parametrami funkcji uniemożliwia bezpośrednie stosowanie klasycznej metody najmniejszych kwadratów do estymacji parametrów tych funkcji. Dlatego też w celu uzyskania liniowej postaci modelu i szacowania parametrów metodą najmniejszych kwadratów należy dokonać pewnych transformacji¹⁵.

Oprócz metody najmniejszych kwadratów do szacowania parametrów regresji można wykorzystywać inne metody, takie jak na przykład minimalizowanie sumy bezwzględnych wartości odchyłeń. Jednak KMNK jest najczęściej wykorzystywaną metodą dopasowania linii prostej do wyników obserwacji¹⁶.

1.4.1. Funkcje nieliniowe o przyspieszonym tempie wzrostu

Do najczęściej stosowanych w opisie zjawisk ekonomicznych postaci funkcji nieliniowych o przyspieszonym tempie wzrostu należą:

- funkcja potęgowa,
- funkcja wykładnicza,
- wielomian stopnia drugiego.

Funkcja potęgowa o przyspieszonym tempie wzrostu wyrażona jest następującym równaniem:

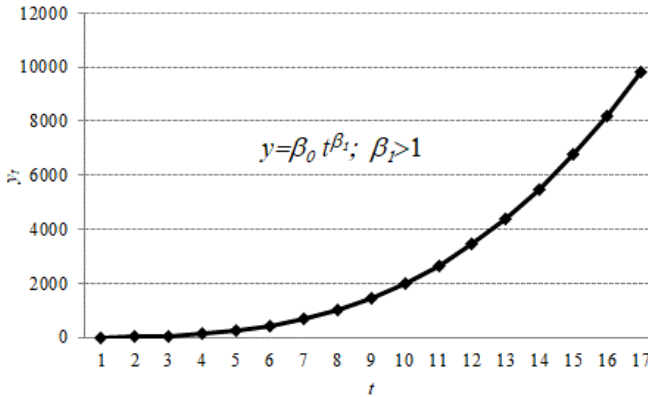
$$y_t = \beta_0 t^{\beta_1}, \quad \beta_1 > 1. \quad (1.14)$$

Podstawową własnością tej funkcji jest stała elastyczność zmiennej y względem zmiennej t . Powinna być ona wykorzystywana do opisu tendencji rozwojowych, które w układzie współrzędnych logarytmicznych wykazują przebieg linowy¹⁷. Przebieg funkcji potęgowej dla $\beta_1 > 1$ przedstawiono na rysunku 1.7.

¹⁵ M. Cieślak (red.), *Prognozowanie gospodarcze...*, op. cit., s. 77.

¹⁶ A.D. Aczel, *Statystyka w zarządzaniu*, Wydawnictwo Naukowe PWN, Warszawa 2000, s. 466.

¹⁷ P. Dittman, *Prognozowanie w przedsiębiorstwie, Metody i ich zastosowanie*, Wolters Kluwer, Kraków 2009, s. 69.



Rys. 1.7. Graficzna reprezentacja funkcji potęgowej ($\beta_0 = 2$; $\beta_1 = 3$)

Źródło: opracowanie własne.

Funkcja potęgowa znajduje bardzo szerokie zastosowanie w analizie rynku przy badaniu popytu na dobra nowe, które znajdują się w fazie rozpowszechniania.

Sprowadzając funkcję potęgową do postaci liniowej, należy dokonać przekształceń, które w przypadku funkcji potęgowej polegają na zlogarytmowaniu obu stron równania (1.14):

$$\ln y_t = \ln \beta_0 + \beta_1 \ln t. \quad (1.15)$$

Następnie należy dokonać następujących podstawień:

$$y'_t = \ln y_t, \quad (1.16)$$

$$\beta'_0 = \ln \beta_0, \quad (1.17)$$

$$t' = \ln t. \quad (1.18)$$

W wyniku powyższych operacji powstaje model liniowy:

$$y'_t = \beta'_0 + \beta'_1 t'. \quad (1.19)$$

Ocenę parametru β_0 otrzymuje się poprzez delogarytmowanie¹⁸:

$$\beta_0 = \exp \beta'_0, \quad (1.20)$$

przy czym:

$$\beta'_1 = \frac{\sum_{t=1}^n \left(t' - \bar{t}' \right) y_t}{\sum_{t=1}^n \left(t' - \bar{t}' \right)^2}, \quad (1.21)$$

$$\beta'_0 = \bar{y}' - \beta'_1 \bar{t}'. \quad (1.22)$$

Kolejnymi omawianymi funkcjami nieliniowymi o przyspieszonym tempie wzrostu są funkcje wykładnicze mające postać:

$$y_t = \beta_0 \beta_1^t, \quad \beta_1 > 1 \quad (1.23)$$

lub

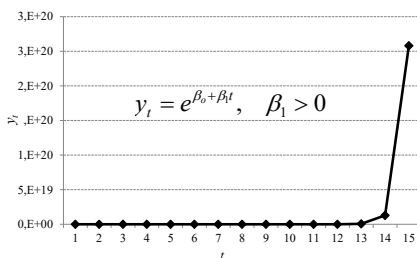
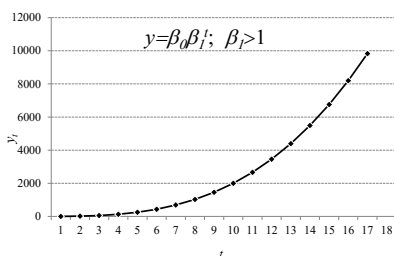
$$y_t = e^{\beta_0 + \beta_1 t}, \quad \beta_1 > 0. \quad (1.24)$$

Funkcje te wykorzystywane są w sytuacjach, w których tempo wzrostu danej wielkości jest stałe, na przykład w badaniu dynamiki dochodu narodowego, w analizie rynku (przy badaniu popytu na dobra nowe, $\beta_1 > 1$).

Właściwością tych funkcji jest stopa wzrostu wynosząca $\ln \beta_1$ i β_1 ¹⁹. Przebieg funkcji wykładniczej o równaniach (1.23) i (1.24) dla $\beta_0 > 1$ i $\beta_1 > 0$ przedstawiono na rysunku 1.8.

¹⁸ M. Cieślak (red.), *Prognozowanie gospodarcze...*, op. cit., s. 80.

¹⁹ Ibidem, s. 77.



Rys. 1.8. Graficzna reprezentacja funkcji wykładniczej ($\beta_0 = 2; \beta_1 = 3$)

Źródło: opracowanie własne.

Dokonując transformacji funkcji wykładniczej w celu uzyskania liniowej postaci modelu, należy dokonać następujących przekształceń:

$$\ln y_t = \ln \beta_0 + t \ln \beta_1, \quad (1.25)$$

$$y'_t = \ln y_t, \quad (1.26)$$

$$\beta'_0 = \ln \beta_0, \quad (1.27)$$

$$\beta'_1 = \ln \beta_1. \quad (1.28)$$

W wyniku powyższych operacji powstaje model zlinearyzowany:

$$y'_t = \beta'_0 + \beta'_1 t. \quad (1.29)$$

Przykładem funkcji o rosnącym tempie wzrostu jest również wielomianu stopnia drugiego o równaniu:

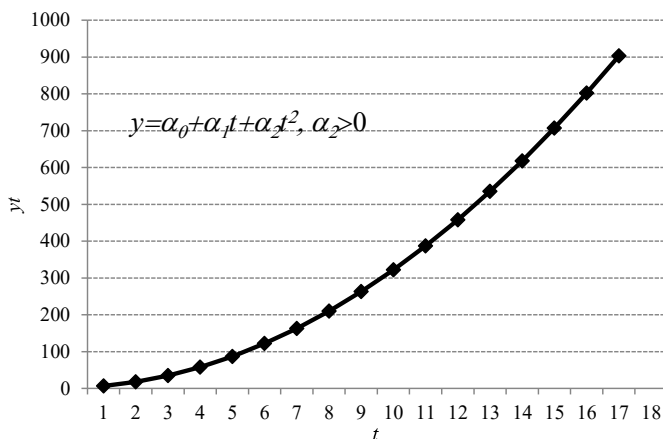
$$y_t = \alpha_0 + \alpha_1 t + \alpha_2 t^2, \quad \alpha_2 > 0. \quad (1.30)$$

Funkcja ta bardzo często jest stosowana w badaniach ekonomiczno-rolniczych do wyjaśniania przyrostów wielkości produkcji w zależności od poziomu nakładów materiałowych oraz do znajdowania optymalnego poziomu tych nakładów²⁰. Zaletą tej funkcji jest duża elastyczność, która wynika z uwzględniania w równaniu trzech parametrów²¹. Dzięki temu

²⁰ P. Dittman, *Prognozowanie w przedsiębiorstwie...*, op. cit., s. 242.

²¹ Ibidem, s. 240.

funkcja ta może lepiej odzwierciedlać nieliniowe tendencje rozwojowe. Przebieg wielomianu drugiego stopnia przedstawiono na rysunku 1.9.



Rys. 1.9. Graficzna reprezentacja wielomianu drugiego stopnia ($\alpha_0 = 2; \alpha_1 = 2; \alpha_2 = 3$)

Źródło: opracowanie własne.

Parametry α_0 , α_1 , α_2 można oszacować przez rozwiązanie następującego układu równań²²:

$$\begin{cases} \sum_{t=1}^n y_t = n\alpha_0 + \alpha_1 \sum_{t=1}^n t + \alpha_2 \sum_{t=1}^n t^2, \\ \sum_{t=1}^n ty_t = \alpha_0 \sum_{t=1}^n t + \alpha_1 \sum_{t=1}^n t^2 + \alpha_2 \sum_{t=1}^n t^3, \\ \sum_{t=1}^n t^2 y_t = \alpha_0 \sum_{t=1}^n t^2 + \alpha_1 \sum_{t=1}^n t^3 + \alpha_2 \sum_{t=1}^n t^4. \end{cases} \quad (1.31)$$

²² M. Sobczyk, *Prognozowanie*, Placet, Warszawa 2008, s. 57.

Przy czym²³:

$$\sum_{t=1}^n t = \frac{n(n+1)}{2}, \quad (1.32)$$

$$\sum_{t=1}^n t^2 = \frac{n(n+1)(2n+1)}{6}, \quad (1.33)$$

$$\sum_{t=1}^n t^3 = \left[\frac{n(n+1)}{2} \right]^2, \quad (1.34)$$

$$\sum_{t=1}^n t^4 = \frac{n(n+1)(2n+1)}{6} \cdot \frac{3n^2 + 3n - 1}{5}, \quad (1.35)$$

gdzie:

n – długość szeregu czasowego.

Parametry α_0 , α_1 , α_2 można też oszacować, rozwiązując następujące równania²⁴:

$$\begin{cases} \sum_{t=1}^n y_t = n\alpha_0 + \alpha_2 \sum_{t=1}^n (t-\bar{t})^2, \\ \sum_{t=1}^n (t-\bar{t})y_t = \alpha_1 \sum_{t=1}^n (t-\bar{t})^2, \\ \sum_{t=1}^n (t-\bar{t})^2 y_t = \alpha_0 \sum_{t=1}^n (t-\bar{t})^2 + \alpha_2 \sum_{t=1}^n (t-\bar{t})^4. \end{cases} \quad (1.36)$$

gdzie:

\bar{t} – średnia arytmetyczna obliczona z numerów jednostek $\bar{t} = \frac{1}{n} \sum_{t=1}^n t$.

Wymienione wyżej funkcje analityczne (potęgowa, wykładnicza i wielomian stopnia drugiego) mogą być odpowiednie do konstrukcji prognoz krótkoterminowych. Ekstrapolacja tych funkcji na zbyt długie

²³ Ibidem, s. 58.

²⁴ Ibidem, s. 59.

okresy przyczynia się do zwiększenia ryzyka budowy prognoz obarczonych dużymi błędami. Związane to jest z faktem, że założenia o przyspieszonym tempie wzrostu nie mogą utrzymać się w dłuższym okresie²⁵.

W dalszej części niniejszego rozdziału zostały przedstawione przykłady oszacowania parametrów wybranych funkcji analitycznych o rosnącym tempie wzrostu. Na podstawie danych empirycznych przedstawionych w postaci jednowymiarowego szeregu czasowego (tab. 1.4) wykonana została estymacja parametrów funkcji potęgowej, wykładniczej i wielomianu stopnia drugiego.

Tab. 1.4. Dane dotyczące produkcji odbiorników telewizyjnych w Polsce w latach 1996-2010

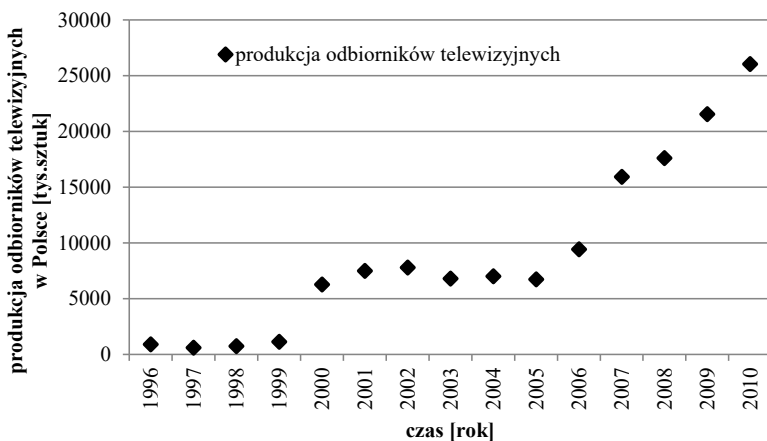
Czas	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
Produkcja odbiorników telewizyjnych [tys. szt.]	900	610	748	1138	6287	7502	7795	6818	7012	6733	9427	15936	17618	21559	26055

Źródło: opracowanie własne na podstawie Rocznika Statystycznego Przemysłu 2012, dokument elektroniczny http://www.stat.gov.pl/gus/5840_3921_PLK_HTML.htm, data wejścia 29.04.2016.

Wykres rozproszenia danych dotyczących produkcji odbiorników telewizyjnych w Polsce w latach 1996-2010 przedstawiono na rysunku 1.10.

Na podstawie rozrzutu punktów przedstawionych graficznie na rysunku 1.10 można zauważyć, że w miarę upływu czasu wzrost wartości zmiennej jest coraz szybszy. W takim wypadku do przybliżenia szeregu należy zastosować funkcję o rosnącym tempie wzrostu. Początkowo jako model tendencji rozwojowej wykorzystano funkcję potęgową (przykł. 1.2), w kolejnych przykładach zastosowano funkcję wykładniczą (przykł. 1.3) oraz wielomian stopnia drugiego (przykł. 1.4).

²⁵ M. Cieślak (red.), *Prognozowanie gospodarcze...*, op. cit., s. 77.



Rys. 1.10. Produkcja odbiorników telewizyjnych w Polsce w latach 1996-2010

Źródło: opracowanie własne.

Przykład 1.2. Oszacowanie parametrów funkcji o rosnącym tempie wzrostu – funkcja potęgowa

Z rozważań z rozdziału 1.4.1 wynika, że funkcja potęgowa o równaniu $y_t = \beta_0 t^{\beta_1}$; $\beta_1 > 1$ powinna dobrze opisywać tendencję do przyspieszonego wzrostu produkcji odbiorników telewizyjnych w Polsce. W celu oszacowania parametrów funkcji potęgowej dokonano jej linearyzacji.

Korzystając z równań (1.20), (1.21) i (1.22) na estymatory KMNK parametrów modelu oraz obliczeń pośrednich umieszczonych w tabeli 1.5, otrzymano następujące wartości:

$$\beta_1 = \frac{12,544}{8,561} = 1,465;$$

$$\beta_0' = 8,583 - 1,465 \cdot 1,860 = 5,858;$$

$$\beta_0 = \exp \beta_0' = 350,02.$$

Niezbędne pośrednie obliczenia do rozwiązania zadania zostały umieszczone w tabeli 1.5.

Tab. 1.5. Obliczenia parametrów modelu klasyczną metodą najmniejszych kwadratów

t	y_t	t'	y'_t	$t' - \bar{t}'$	$(t' - \bar{t}')^2$	$(t' - \bar{t}')y'_t$
1	900	0,00	6,80	-1,86	3,46	-12,65
2	610	0,69	6,41	-1,17	1,36	-7,48
3	748	1,10	6,62	-0,76	0,58	-5,04
4	1138	1,39	7,04	-0,47	0,22	-3,33
5	6287	1,61	8,75	-0,25	0,06	-2,19
6	7502	1,79	8,92	-0,07	0,00	-0,61
7	7795	1,95	8,96	0,09	0,01	0,77
8	6818	2,08	8,83	0,22	0,05	1,94
9	7012	2,20	8,86	0,34	0,11	2,99
10	6733	2,30	8,81	0,44	0,20	3,90
11	9427	2,40	9,15	0,54	0,29	4,92
12	15936	2,48	9,68	0,62	0,39	6,05
13	17618	2,56	9,78	0,70	0,50	6,89
14	21559	2,64	9,98	0,78	0,61	7,77
15	26055	2,71	10,17	0,85	0,72	8,62
		$\bar{t}' = 1,860$	$\bar{y}' = 8,583$			
Σ					8,561	12,544

Źródło: obliczenia własne.

Oszacowana metodą najmniejszych kwadratów funkcja potęgowa opisująca produkcję odbiorników telewizyjnych Polsce ma następującą postać:

$$\hat{y}_t = \beta_0 t^{\beta_1} = 350,02 t^{1,465}.$$

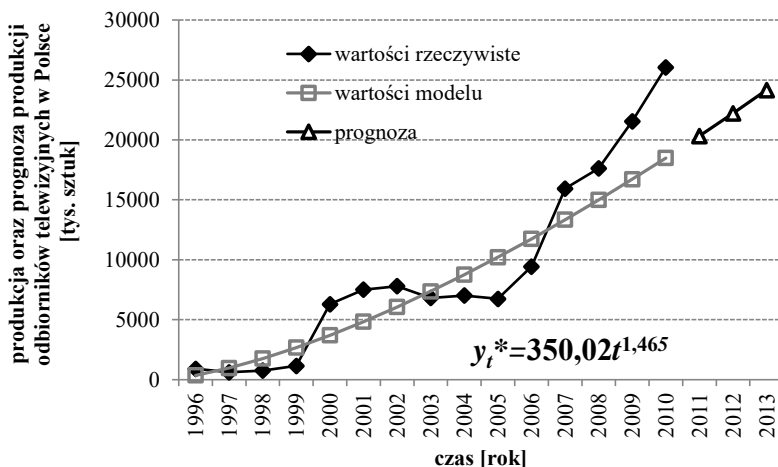
Prognozę produkcji odbiorników telewizyjnych w Polsce na 2011, 2012 i 2013 rok ($t=16, 17, 18$) można wyznaczyć w następujący sposób:

$$y_{16}^* = 350,02 \cdot 16^{1,465} = 2032979;$$

$$y_{17}^* = 350,02 \cdot 17^{1,465} = 2221799;$$

$$y_{18}^* = 350,02 \cdot 18^{1,465} = 2415857.$$

Zatem prognoza produkcji odbiorników telewizyjnych w Polsce w 2011 roku wyniesie 20 329,79 tys. sztuk, w 2012 roku około 22 218 tys. sztuk, natomiast w 2013 roku prognozowana produkcja będzie się kształtowała na poziomie 24 158,57 tys. sztuk.



Rys. 1.11. Wartości rzeczywiste oraz wartości modelu i prognoza produkcji odbiorników telewizyjnych w Polsce wyznaczona z wykorzystaniem funkcji nieliniowej potęgowej

Źródło: opracowanie własne.

Wykres wartości rzeczywistych, wartości modelu oraz prognozę produkcji odbiorników telewizyjnych w Polsce wyznaczoną z wykorzystaniem funkcji potęgowej przedstawiono na rysunku 1.11.

Przykład 1.3. Oszacowanie parametrów funkcji o rosnącym tempie wzrostu – funkcja wykładnicza

Na podstawie analizy wartości szeregu czasowego (tab. 1.4), jak i wzrokowej oceny przebiegu zmiennej przedstawionej na rysunku 1.10 w przykładzie 1.2 zauważano, że do budowy modelu analitycznego można również wykorzystać funkcję wykładniczą o równaniu $y_t = \beta_0 \beta_1^t$, $\beta_0 > 1$.

W celu oszacowania parametrów funkcji wykładniczej sprowadzono ją do postaci liniowej względem parametrów zgodnie z przedstawionymi wcześniej zasadami transformacji. Korzystając z równań (1.25-1.29), oszacowano następujące wartości parametrów zlinearyzowanej postaci modelu β'_1 , β'_0 oraz postaci pierwotnej β_1 , β_0 :

$$\beta'_1 = \frac{72,196}{280,000} = 0,258;$$

$$\beta'_0 = 8,583 - 0,258 \cdot 8,000 = 6,519;$$

$$\beta_1 = \exp \beta'_1 = 1,294;$$

$$\beta_0 = \exp \beta'_0 = 677,900.$$

Obliczenia pośrednie niezbędne do oszacowania wartości ww. parametrów modelu wykładniczego zawarto w tabeli 1.6.

Tab. 1.6. Obliczenia parametrów modelu metodą najmniejszych kwadratów

t	y_t	y_t'	$t - \bar{t}$	$(t - \bar{t})^2$	$(t - \bar{t})y_t'$
1	900	6,80	-7,00	49,00	-47,62
2	610	6,41	-6,00	36,00	-38,48
3	748	6,62	-5,00	25,00	-33,09
4	1138	7,04	-4,00	16,00	-28,15
5	6287	8,75	-3,00	9,00	-26,24
6	7502	8,92	-2,00	4,00	-17,85
7	7795	8,96	-1,00	1,00	-8,96
8	6818	8,83	0,00	0,00	0,00
9	7012	8,86	1,00	1,00	8,86
10	6733	8,81	2,00	4,00	17,63
11	9427	9,15	3,00	9,00	27,45
12	15936	9,68	4,00	16,00	38,71
13	17618	9,78	5,00	25,00	48,88
14	21559	9,98	6,00	36,00	59,87
15	26055	10,17	7,00	49,00	71,18
$\bar{t} = 8,000$		$\bar{y}' = 8,583$			
Σ				280,000	72,196

Źródło: obliczenia własne.

Po oszacowaniu metodą najmniejszych kwadratów parametrów modelu bazującego na wykładniczej postaci funkcji trendu można jego postać przedstawić w następujący sposób:

$$\hat{y}_t = \beta_0 \beta_1^t = 677,900 \cdot 1,294^t.$$

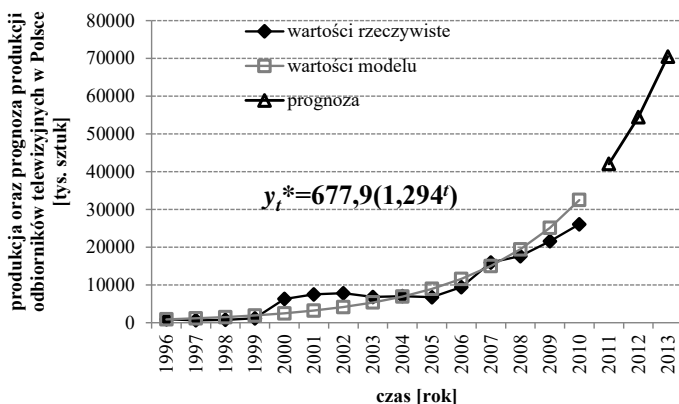
Funkcja ta opisuje wielkość produkcji odbiorników telewizyjnych w Polsce w latach 1996-2010. Prognozę produkcji odbiorników telewizyjnych w Polsce na 2011, 2012 i 2013 rok ($t = 16, 17, 18$) można wyznaczyć w następujący sposób:

$$y_{16}^* = 677,900 \cdot 1,294^{16} = 42066,21;$$

$$y_{17}^* = 677,900 \cdot 1,294^{17} = 54447,92;$$

$$y_{18}^* = 677,900 \cdot 1,294^{18} = 70474,06.$$

Zatem prognoza produkcji odbiorników telewizyjnych w Polsce w 2011 roku wyniesie 42 066,21 tys. sztuk, w 2012 roku około 54 447 tys. sztuk, natomiast w 2013 roku prognozowana produkcja będzie się kształtowała na poziomie 70 474,06 tys. sztuk.



Rys. 1.12. Wartości rzeczywiste oraz wartości modelu i prognoza produkcji odbiorników telewizyjnych w Polsce wyznaczona z wykorzystaniem funkcji wykładniczej

Źródło: opracowanie własne.

Wykres wartości rzeczywistych, wartości modelu oraz prognozę produkcji odbiorników telewizyjnych w Polsce wyznaczoną z wykorzystaniem funkcji wykładniczej przedstawiono na rysunku 1.12.

Analizując wizualnie wykresy zmiennych przedstawione na rysunku 1.12 można zauważyć, że wykorzystywana funkcja wykładnicza o równaniu $y_t^* = 677,900 \cdot 1,294^t$ jest dosyć dobrze dopasowana do wartości rzeczywistych badanej zmiennej.

Przykład 1.4. Oszacowanie parametrów funkcji o rosnącym tempie wzrostu – wielomian stopnia drugiego

Przedstawiona na rysunku 1.10 tendencja do przyśpieszonego wzrostu produkcji odbiorników telewizyjnych w Polsce może być opisana również przez wielomian stopnia drugiego o równaniu (1.30).

Parametry α_0 , α_1 , α_2 zostały oszacowane przez rozwiązanie omówionego wcześniej układu równań (1.31) oraz (1.32-1.35).

Szacując:

$$\sum_{t=1}^n t = \frac{n(n+1)}{2} = \frac{15(15+1)}{2} = 120,$$

$$\sum_{t=1}^n t^2 = \frac{n(n+1)(2n+1)}{6} = \frac{15(15+1)(30+1)}{6} = 1\,240,$$

$$\sum_{t=1}^n t^3 = \left[\frac{n(n+1)}{2} \right]^2 = \left[\frac{15(15+1)}{2} \right]^2 = 14\,400,$$

$$\begin{aligned} \sum_{t=1}^n t^4 &= \frac{n(n+1)(2n+1)}{6} \cdot \frac{3n^2 + 3n - 1}{5} = \\ &= \frac{15(15+1)(30+1)}{6} \cdot \frac{3(15)^2 + 3(15) - 1}{5} = 178\,312. \end{aligned}$$

Po rozwiązaniu układu równań (1.31) otrzymano następujące oszacowania parametrów modelu analitycznego:

$$\alpha_0 = 1898409;$$

$$\alpha_1 = -413,212;$$

$$\alpha_2 = 126,812.$$

Oszacowany model można zatem przedstawić następująco:

$$\hat{y}_t = 1898409 - 413212t + 126812t^2.$$

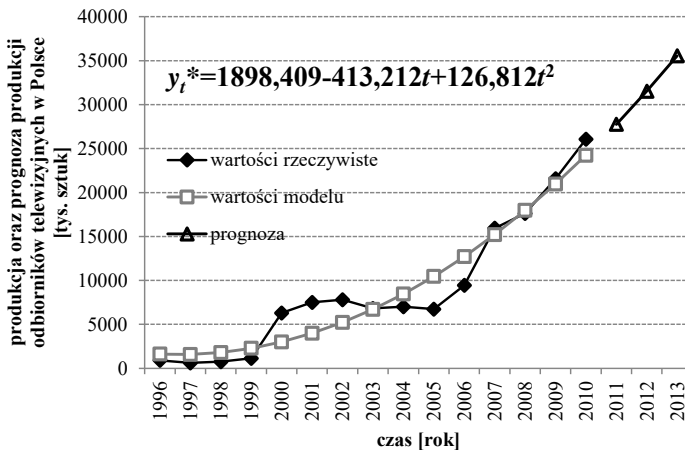
Prognozę produkcji odbiorników telewizyjnych w Polsce na 2011, 2012 i 2013 rok ($t = 16, 17, 18$), wykorzystując powyższy model, można wyznaczyć w następujący sposób:

$$y_{16}^* = 1898409 - 413212 \cdot 16 + 126812 \cdot 16^2 = 2775089;$$

$$y_{17}^* = 1898409 - 413212 \cdot 17 + 126812 \cdot 17^2 = 3152247;$$

$$y_{18}^* = 1898409 - 413212 \cdot 18 + 126812 \cdot 18^2 = 3554768.$$

Zatem prognoza produkcji odbiorników telewizyjnych w Polsce w 2011 roku wyniesie 27 750,89 tys. sztuk, w 2012 roku około 31 522,47 tys. sztuk, natomiast w 2013 roku prognozowana produkcja będzie się kształtowała na poziomie 35 547,68 tys. sztuk. Wykres wartości rzeczywistych i wartości modelu oraz prognozę produkcji odbiorników telewizyjnych w Polsce wyznaczoną z wykorzystaniem wielomianu stopnia drugiego przedstawiono na rysunku 1.13.



Rys. 1.13. Wartości rzeczywiste oraz wartości modelu i prognoza produkcji odbiorników telewizyjnych w Polsce wyznaczona z wykorzystaniem wielomianu stopnia drugiego

Źródło: opracowanie własne.

Analizując wizualnie wykres zmiennej przedstawiony na rysunku 1.13, można zauważyć, że wykorzystany do budowy modelu analitycznego wielomian drugiego stopnia o równaniu $\hat{y}_t = 1898409 - 413,212t + 126,812t^2$ poprawnie odzwierciedla dane rzeczywiste, tj. podąża za zmianami zmiennej prognozowanej y .

1.4.2. Funkcje nieliniowe o malejącym tempie wzrostu

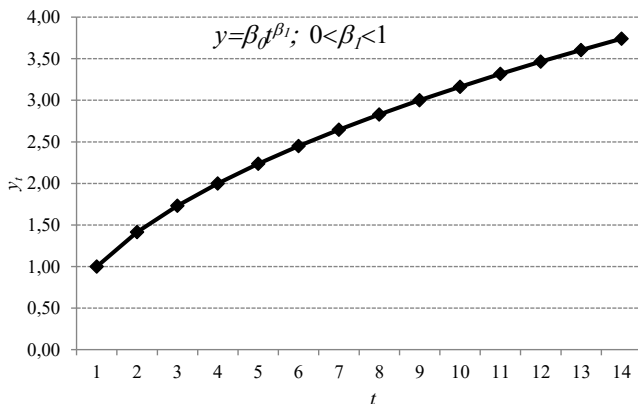
Funkcje o malejącym tempie wzrostu znajdują zastosowanie w przypadku, gdy wzrost wartości zmiennej prognozowanej przebiega coraz wolniej i zdąża do pewnego poziomu. Dotyczy to na przykład sytuacji względnego nasycenia rynku. Wśród możliwych do zastosowania funkcji opisujących coraz wolniejszy wzrost wartości zmiennej prognozowanej można wymienić:

- funkcję potęgową,
- funkcję logarytmiczną,
- funkcję liniowo-odwrotnościową,
- funkcję ilorazową,
- wielomian odwrotnościowy,
- wielomian stopnia drugiego.

Przebieg przykładowej funkcji potęgowej o malejącym tempie wzrostu opisanej równaniem:

$$y_t = \beta_0 t^{\beta_1}, \quad 0 < \beta_1 < 1 \quad (1.37)$$

przedstawiono na rysunku 1.14.



Rys. 1.14. Graficzna reprezentacja funkcji potęgowej ($\beta_0 = 1; \beta_1 = 0,5$)

Źródło: opracowanie własne.

Dokonując transformacji funkcji potęgowej w celu uzyskania liniowej postaci modelu, należy dokonać przekształceń, które w wypadku tej funkcji polegają na zlogarytmowaniu obu stron równania (1.37) i zastosowaniu następujących podstawień:

$$\ln y_t = \ln \beta_0 + \beta_1 \ln t, \quad (1.38)$$

$$y'_t = \ln y_t, \quad (1.39)$$

$$\beta'_0 = \ln \beta_0, \quad (1.40)$$

$$t' = \ln t. \quad (1.41)$$

W wyniku powyższych operacji powstaje model liniowy:

$$y'_t = \beta'_0 + \beta_1 t'. \quad (1.42)$$

Ocenę parametru β_0 otrzymuje się poprzez delogarytmowanie²⁶:

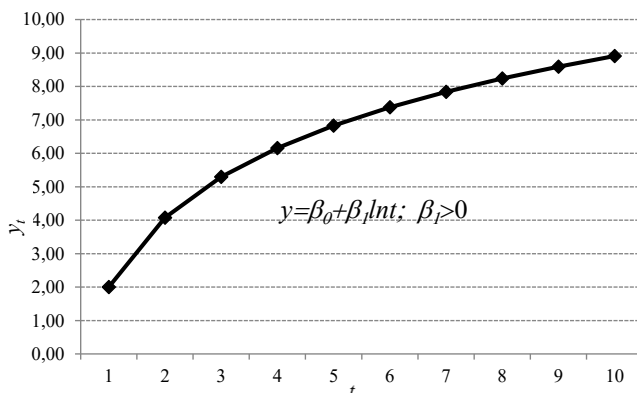
$$\beta_0 = \exp \beta'_0. \quad (1.43)$$

²⁶ Ibidem, s. 80.

Z kolei funkcję logarytmiczną można określić następującym równaniem:

$$y_t = \beta_0 + \beta_1 \ln t, \quad \beta_1 > 0. \quad (1.44)$$

Przebieg przykładowej funkcji logarytmicznej przedstawiono na rysunku 1.15.



Rys. 1.15. Graficzna reprezentacja funkcji logarytmicznej
($\beta_0 = 2; \beta_1 = 3$)

Źródło: opracowanie własne.

Funkcja logarytmiczna należy do funkcji sprowadzalnych do postaci liniowej względem parametrów. W celu uzyskania liniowej postaci modelu analitycznego bazującego na logarytmicznej postaci funkcji trendu należy dokonać następującego przekształcenia:

$$t' = \ln t. \quad (1.45)$$

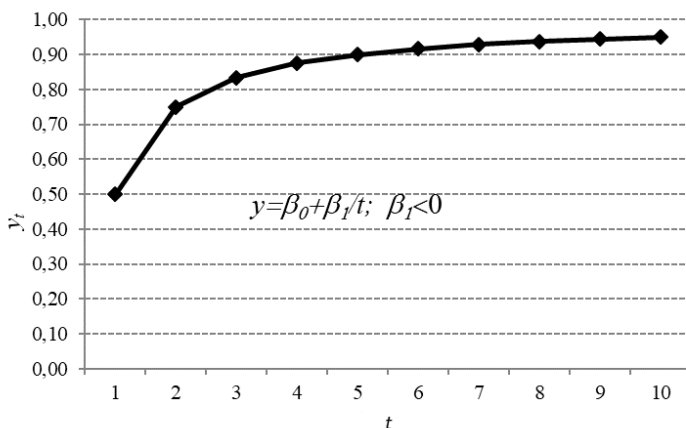
W wyniku powyższej operacji powstaje model liniowy:

$$y_t = \beta_0 + \beta_1 t', \quad \beta_1 > 0. \quad (1.46)$$

Kolejną funkcją nieliniową o malejącym tempie wzrostu jest funkcja liniowo-odwrotnościowa:

$$y_t = \beta_0 + \frac{\beta_1}{t}, \quad \beta_1 < 0. \quad (1.47)$$

Przebieg przykładowej funkcji liniowo-odwrotnościowej przedstawiono na rysunku 1.16.



Rys. 1.16. Graficzna reprezentacja funkcji liniowo-odwrotnościowej ($\beta_0 = 1; \beta_1 = -0,5$)

Źródło: opracowanie własne.

Dokonując transformacji funkcji liniowo-odwrotnościowej w celu uzyskania liniowej postaci modelu, należy dokonać następujących przekształceń:

$$t' = \frac{1}{t}. \quad (1.48)$$

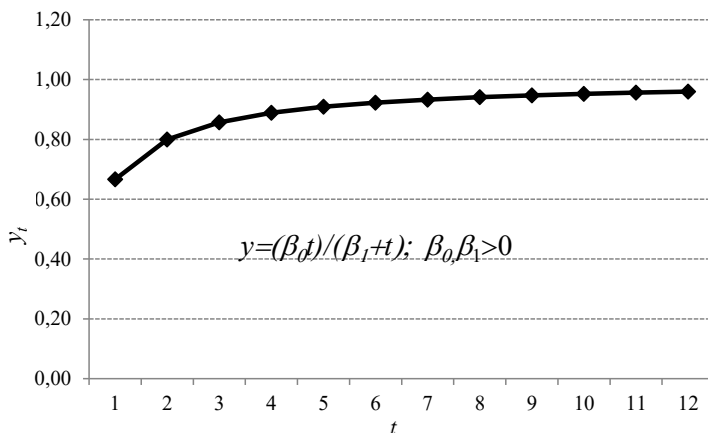
W wyniku powyższej operacji powstaje model liniowy:

$$y_t = \beta_0 + \beta_1 t', \quad \beta_1 < 0. \quad (1.49)$$

Do funkcji nieliniowych o malejącym tempie wzrostu można również zaliczyć funkcję ilorazową przedstawioną równaniem:

$$y_t = \frac{\beta_0 t}{\beta_1 + t}, \quad \beta_0, \beta_1 > 0. \quad (1.50)$$

Przebieg przykładowej funkcji ilorazowej przedstawiono na rysunku 1.17.



Rys. 1.17. Graficzna reprezentacja funkcji ilorazowej ($\beta_0 = 1; \beta_1 = 0,5$)

Źródło: opracowanie własne.

Dokonując transformacji funkcji ilorazowej w celu uzyskania liniowej postaci modelu należy dokonać następujących przekształceń:

$$y'_t = \frac{1}{y_t}, \quad (1.51)$$

$$\beta'_0 = \frac{1}{\beta_0}, \quad (1.52)$$

$$\beta'_1 = \frac{\beta_1}{\beta_0}, \quad (1.53)$$

$$t' = \frac{1}{t}. \quad (1.54)$$

W wyniku powyższych operacji powstaje model liniowy:

$$y'_t = \beta'_0 + \beta'_1 t'. \quad (1.55)$$

Ocenę parametrów β_0 i β_1 otrzymuje się poprzez następujące przekształcenie odwrotnościowe²⁷:

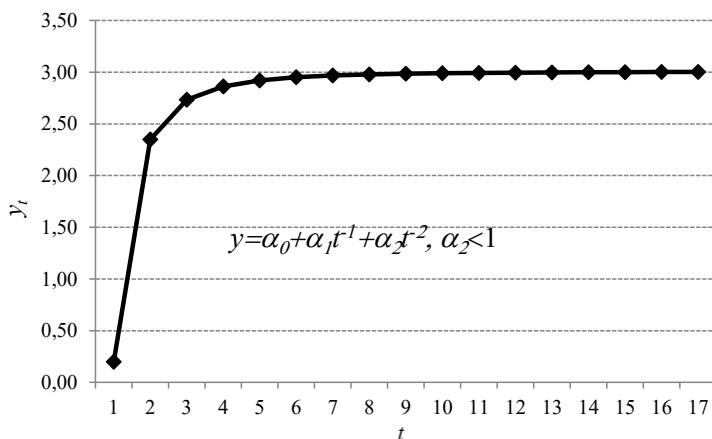
$$\beta_0 = \frac{1}{\beta'_0}, \quad (1.56)$$

$$\beta_1 = \beta_0 \beta'_1. \quad (1.57)$$

Wśród funkcji nieliniowych o malejącym tempie wzrostu można także wyróżnić wielomian odwrotnościowy, który da się przedstawić w postaci podanego poniżej równania:

$$y_t = \alpha_0 + \alpha_1 t^{-1} + \alpha_2 t^{-2}, \quad \alpha_2 < 1. \quad (1.58)$$

Natomiast przebieg przykładowej funkcji wielomianu odwrotnościowego przedstawiono na rysunku 1.18.



Rys. 1.18. Graficzna reprezentacja wielomianu odwrotnościowego ($\alpha_0 = 3; \alpha_1 = 0,2; \alpha_2 = -3$)

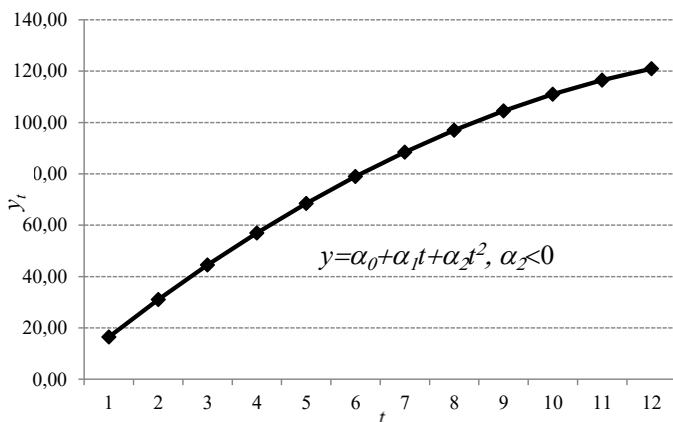
Źródło: opracowanie własne.

²⁷ M. Cieślak (red.), *Prognozowanie gospodarcze...*, op. cit., s. 80.

Ostatnią omawianą funkcją analityczną o malejącym tempie wzrostu jest wielomian stopnia drugiego (parabola). Wyrażony jest on równaniem:

$$y_t = \alpha_0 + \alpha_1 t + \alpha_2 t^2, \quad \alpha_2 < 0. \quad (1.59)$$

Przebieg przykładowej funkcji wielomianu stopnia drugiego przedstawiono na rysunku 1.19.



Rys. 1.19. Graficzna reprezentacja wielomianu stopnia drugiego ($\alpha_0 = 1; \alpha_1 = 16; \alpha_2 = -0,5$)

Źródło: opracowanie własne.

Parametry α_0 , α_1 , α_2 można oszacować przez rozwiązanie układu równań (1.31-1.35) lub (1.36).

Przedstawiając ten rodzaj funkcji trendu, warto wspomnieć, że prognozowanie na podstawie funkcji o malejącym tempie wzrostu jest obarczone na ogół mniejszym ryzykiem wystąpienia błędów niż w przypadku funkcji o przyśpieszonym tempie wzrostu.

W dalszej części niniejszego rozdziału, dla tych samych danych dotyczących liczby gospodarstw domowych w Polsce wyposażonych w komputery, przedstawiono przykłady oszacowania parametrów wybranych funkcji analitycznych o malejącym tempie wzrostu.

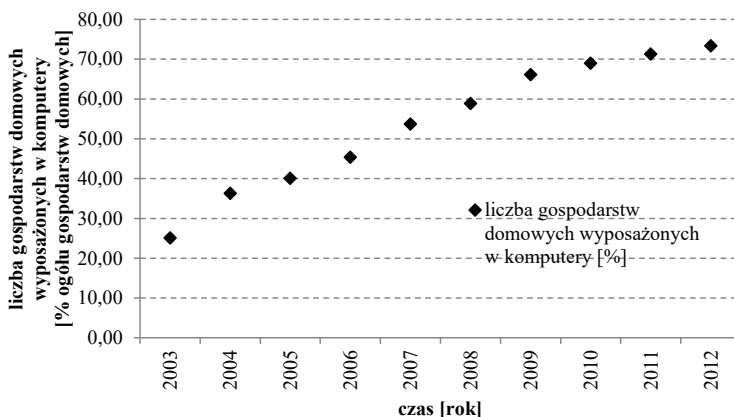
Na podstawie danych empirycznych przedstawionych w tabeli 1.7 w postaci szeregu czasowego zidentyfikowano adekwatną postać funkcji oraz oszacowano wartości jej parametrów metodą najmniejszych kwadratów.

Tab. 1.7. Dane dotyczące liczby gospodarstw domowych wyposażonych w komputery [% ogółu gospodarstw domowych] w latach 2003-2012

Czas	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012
Liczby gospodarstw domowych wyposażonych w komputery	25,10	36,30	40,10	45,40	53,70	58,90	66,10	69,00	71,30	73,40

Źródło: opracowanie własne na podstawie rocznych wskaźników makroekonomicznych, dokument elektroniczny http://www.stat.gov.pl/gus/wskazniki_makroekon_PLK_HTML.htm, data wejścia 29.04.2016.

Wykres danych dotyczących liczby gospodarstw domowych w Polsce wyposażonych w komputery [% ogółu gospodarstw domowych] w latach 2003-2012 przedstawiono na rysunku 1.20.



Rys. 1.20. Liczba gospodarstw domowych wyposażonych w komputery w Polsce w latach 2003-2012 [% ogółu gospodarstw domowych]

Źródło: opracowanie własne.

Na podstawie oceny wzrokowej szeregu przedstawionego na rysunku 1.20 można zauważyć, że początkowo szybki wzrost wartości zmiennej w miarę upływu czasu jest coraz wolniejszy. W takim wypadku do oszacowania parametrów należy zastosować jedną z funkcji o malejącym tempie wzrostu. W celu wybrania najbardziej adekwatnej postaci modelu i funkcji na wstępnym etapie prac prognostycznych wykorzystano funkcję potęgową (przykł. 1.5), w kolejnych krokach do szacowania parametrów funkcji zastosowano bardziej złożone funkcje, tj. logarytmiczną (przykł. 1.6), liniowo-odwrotnościową (przykł. 1.7), ilorazową (przykł. 1.8) i wielomian stopnia drugiego (przykł. 1.9).

Przykład 1.5. Oszacowanie parametrów funkcji o malejącym tempie wzrostu – funkcja potęgowa

Funkcja potęgowa o równaniu $y_t = \beta_0 t^{\beta_1}$ $0 < \beta_1 < 1$ powinna dobrze opisywać tendencję do malejącego wzrostu liczby gospodarstw domowych w Polsce wyposażonych w komputery.

W celu oszacowania parametrów funkcji potęgowej dokonano linearyzacji funkcji potęgowej. Korzystając z równań (1.38-1.43), otrzymano następujące wartości parametrów modelu:

$$\beta_1 = \frac{2,329}{4,836} = 0,482;$$

$$\beta_0' = 3,937 - 0,482 \cdot 1,510 = 3,209;$$

$$\beta_0 = \exp \beta_0' = 24,754.$$

Niezbędne pośrednie obliczenia wykorzystane do oszacowania wartości parametrów modelu umieszczono w tabeli 1.8.

Tab. 1.8. Obliczenia parametrów modelu metodą najmniejszych kwadratów

Rok	t	y_t	t'	y'_t	$t'-\bar{t}$	$(t'-\bar{t})^2$	$(t'-\bar{t})y'_t$
2003	1	25,10	0,00	3,22	-1,51	2,28	-4,87
2004	2	36,30	0,69	3,59	-0,82	0,67	-2,93
2005	3	40,10	1,10	3,69	-0,41	0,17	-1,52
2006	4	45,40	1,39	3,82	-0,12	0,02	-0,47
2007	5	53,70	1,61	3,98	0,10	0,01	0,40
2008	6	58,90	1,79	4,08	0,28	0,08	1,15
2009	7	66,10	1,95	4,19	0,44	0,19	1,83
2010	8	69,00	2,08	4,23	0,57	0,32	2,41
2011	9	71,30	2,20	4,27	0,69	0,47	2,93
2012	10	73,40	2,30	4,30	0,79	0,63	3,40
			$\bar{t} = 1,510$	$\bar{y}' = 3,937$			
Σ						4,836	2,329

Źródło: obliczenia własne.

Po oszacowaniu klasyczną metodą najmniejszych kwadratów parametrów funkcji można ją przedstawić w postaci funkcji trendu o malejącym tempie wzrostu za pomocą podanej poniżej zależności:

$$\hat{y}_t = \beta_0 t^{\beta_1} = 24,754 t^{0,482}.$$

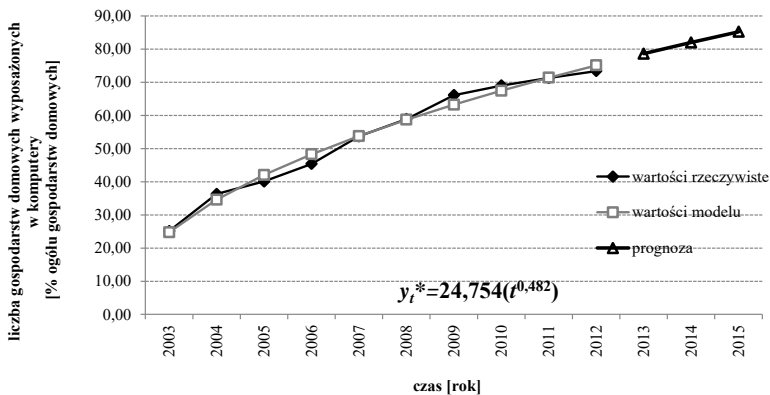
Prognozę liczby gospodarstw domowych w Polsce wyposażonych w komputery na 2013 ($t=11$), 2014 ($t=12$) i 2015 rok ($t=13$) można wyznaczyć w następujący sposób:

$$t = 11, \quad y_{11}^* = 24,754 \cdot 11^{0,482} = 78,63;$$

$$t = 12, \quad y_{12}^* = 24,754 \cdot 12^{0,482} = 82,00;$$

$$t = 13, \quad y_{13}^* = 24,754 \cdot 13^{0,482} = 85,23.$$

Zatem prognoza liczby gospodarstw domowych w Polsce wyposażonych w komputery, wyrażona w % ogółu gospodarstw domowych, dla 2013 roku wynosi 78,63%, 2014 roku 82,00% a dla 2015 roku 85,23%. Wykres wartości rzeczywistych, wartości modelu oraz prognozę liczby gospodarstw domowych wyposażonych w komputery w Polsce (wyrażoną w % ogółu gospodarstw domowych) wyznaczoną z wykorzystaniem funkcji potęgowej przedstawiono na rysunku 1.21.



Rys. 1.21. Wartości rzeczywiste oraz wartości modelu i prognoza liczby gospodarstw domowych wyposażonych w komputery w Polsce (wyrażona w % ogółu gospodarstw domowych) wyznaczona z wykorzystaniem funkcji potęgowej

Źródło: opracowanie własne.

Analizując wizualnie przebiegi przedstawione na rysunku 1.21, można dostrzec, że wartości modelu wyznaczone z wykorzystaniem funkcji potęgowej są zbliżone w dużym stopniu z wartościami zaobserwowanymi. Zatem model zbudowany z wykorzystaniem funkcji potęgowej o równaniu $\hat{y}_t = 24,754t^{0,482}$ dobrze odzwierciedla zmiany wartości rzeczywistych badanej zmiennej.

Przykład 1.6. Oszacowanie parametrów funkcji o malejącym tempie wzrostu – funkcja logarytmiczna

Analizując zmiany wartości i sposób kształtowania się zmiennej zaprezentowane w tabeli 1.7 oraz na rysunku 1.20, stwierdzono, że do budowy modelu prognostycznego można wykorzystać również funkcję logarytmiczną.

W celu oszacowania parametrów funkcji logarytmicznej dokonano jej linearyzacji zgodnie z podanymi wcześniej zasadami transformacji. W tabeli 1.9 umieszczono obliczenia pośrednie, niezbędne do oszacowania wartości parametrów modelu.

Tab. 1.9. Obliczenia parametrów modelu metodą najmniejszych kwadratów

Rok	t	y_t	t'	$t' - \bar{t}'$	$(t' - \bar{t}')^2$	$(t' - \bar{t}')y_t$
2003	1	25,10	0,00	-1,51	2,28	-37,90
2004	2	36,30	0,69	-0,82	0,67	-29,65
2005	3	40,10	1,10	-0,41	0,17	-16,50
2006	4	45,40	1,39	-0,12	0,02	-5,62
2007	5	53,70	1,61	0,10	0,01	5,34
2008	6	58,90	1,79	0,28	0,08	16,60
2009	7	66,10	1,95	0,44	0,19	28,81
2010	8	69,00	2,08	0,57	0,32	39,29
2011	9	71,30	2,20	0,69	0,47	49,00
2012	10	73,40	2,30	0,79	0,63	58,18
		$\bar{y} = 53,930$	$\bar{t}' = 1,510$			
Σ					4,836	107,550

Źródło: obliczenia własne.

Korzystając z równań (1.45) i (1.46) oraz wyników obliczeń pomocniczych umieszczonych w tabeli 1.9, otrzymano następujące wartości parametrów modelu:

$$\beta_1 = \frac{107,550}{4,836} = 22,240;$$

$$\beta_0 = 53,930 - 22,240 \cdot 1,510 = 20,348.$$

Oszacowane klasyczną metodą najmniejszych kwadratów parametry funkcji trendu w wypadku analizowanej zmiennej można przedstawić za pomocą następującej zależności:

$$\hat{y}_t = \beta_0 + \beta_1 \ln t = 20,348 + 22,240 \ln t.$$

Prognozę liczby gospodarstw domowych w Polsce na 2013, 2014 i 2015 rok ($t=11, 12, 13$) można wyznaczyć w następujący sposób:

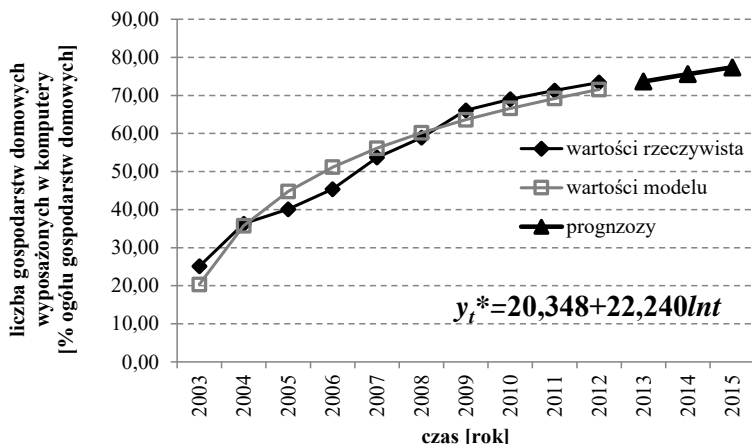
$$y_{11}^* = 20,348 + 22,240 \ln 11 = 73,68;$$

$$y_{12}^* = 20,348 + 22,240 \ln 12 = 75,61;$$

$$y_{13}^* = 20,348 + 22,240 \ln 13 = 77,39.$$

Zatem prognoza liczby gospodarstw domowych w Polsce wyposażonych w komputery, wyrażona w % ogółu gospodarstw domowych,

dla 2013 roku wynosi około 73,68%, dla 2014 roku kształtuje się na poziomie 75,614%, a dla 2015 roku na poziomie 77,39%. Wykres wartości rzeczywistych, wartości modelu oraz prognoza liczby gospodarstw domowych wyposażonych w komputery w Polsce (wyrażoną w % ogółu gospodarstw domowych) wyznaczona z wykorzystaniem funkcji nieliniowej logarytmicznej została przedstawiona na rysunku 1.22.



Rys. 1.22. Wartości rzeczywiste oraz wartości modelu i prognoza liczby gospodarstw domowych wyposażonych w komputery w Polsce (wyrażona w % ogółu gospodarstw domowych) wyznaczona z wykorzystaniem funkcji logarytmicznej

Źródło: opracowanie własne.

Analizując wizualnie przebiegi teoretyczne i empiryczne zmiennej przedstawione na rysunku 1.22 można zauważyć, że analityczny model zbudowany z wykorzystaniem funkcji logarytmicznej o równaniu $y_t^* = 20,348 + 22,240 \ln t$ dobrze odzwierciedla zmiany wartości rzeczywistych.

Przykład 1.7. Oszacowanie parametrów funkcji o malejącym tempie wzrostu – funkcja liniowo-odwrotnościowa

Przedstawiona na wykresie 1.20 tendencja do malejącego tempa wzrostu liczby gospodarstw domowych w Polsce wyposażonych w komputery może być opisana także przez funkcję liniowo-odwrotnościową o równaniu (1.47). W celu uzyskania liniowej postaci funkcji liniowo-odwrotnościowej dokonano przekształcenia (1.48). Korzystając z równań (1.47), (1.48) i (1.49), otrzymano następujące wartości parametrów modelu:

$$\beta_1 = \frac{-36,161}{0,692} = -52,256;$$

$$\beta_0 = 53,930 + 52,256 \cdot 0,293 = 69,241.$$

Obliczenia pośrednie niezbędne do oszacowania wartości parametrów ww. modelu zamieszczono w tabeli 1.10.

Tab. 1.10. Obliczenia parametrów modelu klasyczną metodą najmniejszych kwadratów

Rok	t	y_t	t'	$t' - \bar{t}'$	$(t' - \bar{t}')^2$	$(t' - \bar{t}')y_t$
2003	1	25,10	1,00	0,71	0,50	17,75
2004	2	36,30	0,50	0,21	0,04	7,51
2005	3	40,10	0,33	0,04	0,00	1,62
2006	4	45,40	0,25	-0,04	0,00	-1,95
2007	5	53,70	0,20	-0,09	0,01	-4,99
2008	6	58,90	0,17	-0,13	0,02	-7,44
2009	7	66,10	0,14	-0,15	0,02	-9,92
2010	8	69,00	0,13	-0,17	0,03	-11,59
2011	9	71,30	0,11	-0,18	0,03	-12,97
2012	10	73,40	0,10	-0,19	0,04	-14,17
		$\bar{y} = 53,930$	$\bar{t}' = 0,293$			
Σ					0,692	-36,161

Źródło: obliczenia własne.

Po oszacowaniu klasyczną metodą najmniejszych kwadratów parametrów funkcji postać modelu analitycznego można przedstawić w następujący sposób:

$$\hat{y}_t = \beta_0 + \frac{\beta_1}{t} = 69,241 + \frac{-52,256}{t}.$$

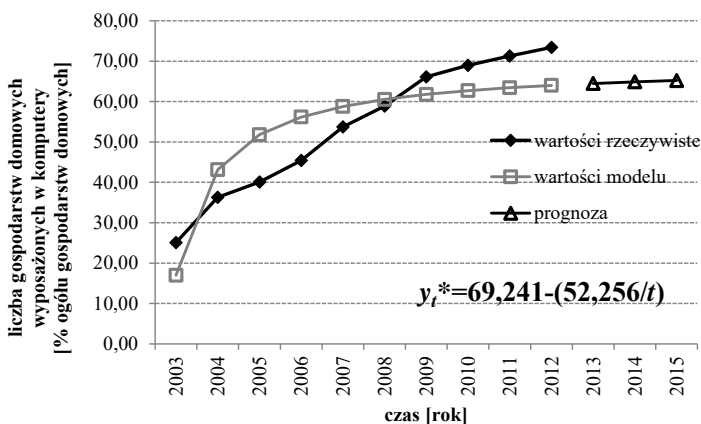
Prognozę liczby gospodarstw domowych w Polsce wyposażonych w komputery wyrażoną w % ogółu gospodarstw domowych na 2013, 2014 i 2015 rok ($t = 11, 12, 13$) można wyznaczyć w następujący sposób:

$$y_{11}^* = 69,241 + \frac{-52,256}{11} = 64,49;$$

$$y_{12}^* = 69,241 + \frac{-52,256}{12} = 64,89;$$

$$y_{13}^* = 69,241 + \frac{-52,256}{13} = 65,22.$$

Zatem prognoza liczby gospodarstw domowych w Polsce wyposażonych w komputery, wyrażona w % ogółu gospodarstw domowych, dla 2013 roku wynosi około 64,49%, dla 2014 roku kształtuje się na poziomie 64,89%, a dla roku 2015 na poziomie 65,22%.



Rys. 1.23. Wartości rzeczywiste, wartość modelu oraz prognoza liczby gospodarstw domowych wyposażonych w komputery w Polsce wyznaczona z wykorzystaniem funkcji liniowo-odwrotnościowej (wyrażona w % ogółu gospodarstw domowych)

Źródło: opracowanie własne.

Wykres wartości rzeczywistych, wartości modelu oraz prognozę liczby gospodarstw domowych w Polsce wyposażonych w komputery

(wyrażoną w % ogółu gospodarstw domowych) wyznaczoną z wykorzystaniem funkcji nieliniowej liniowo-odwrotnościowej przedstawiono na rysunku 1.23. Analizując rysunek, można zauważyć, że wykorzystana do budowy modelu analitycznego funkcja liniowo-odwrotnościowa nie odzwierciedla poprawnie danych rzeczywistych, tj. nie podąża za zmianami zmiennej prognozowanej y .

Przykład 1.8. Oszacowanie parametrów funkcji o malejącym tempie wzrostu – funkcja ilorazowa

Na podstawie podobnych jak poprzednio przesłanek założono, że tendencja rozwojowa o malejącym tempie wzrostu, przedstawiona na wykresie 1.20, może być również odzwierciedlona za pomocą funkcji ilorazowej.

W celu oszacowania parametrów funkcji ilorazowej dokonano liczych przekształceń. Korzystając z równań (1.51-1.54), (1.55) i (1.56-1.57), otrzymano następujące wartości parametrów analitycznego modelu:

$$\beta'_1 = \frac{0,020}{0,692} = 0,029;$$

$$\beta'_0 = 0,021 - 0,029 \cdot 0,293 = 0,013;$$

$$\beta_0 = \frac{1}{\beta'_0} = \frac{1}{0,013} = 76,923;$$

$$\beta_1 = \beta_0 \cdot \beta'_1 = 76,923 \cdot 0,029 = 2,231.$$

Podobnie jak w poprzednich przykładach, obliczenia pośrednie niezbędne do oszacowania wartości parametrów modelu umieszczono w tabeli 1.11.

Tab. 1.11. Obliczenia parametrów modelu metodą najmniejszych kwadratów

Rok	t	y_t	t'	y'	$t'-\bar{t}$	$(t'-\bar{t})^2$	$(t'-\bar{t})y_t'$
2003	1	25,10	1,00	0,04	0,71	0,50	0,028
2004	2	36,30	0,50	0,03	0,21	0,04	0,006
2005	3	40,10	0,33	0,02	0,04	0,00	0,001
2006	4	45,40	0,25	0,02	-0,04	0,00	-0,001
2007	5	53,70	0,20	0,02	-0,09	0,01	-0,002
2008	6	58,90	0,17	0,02	-0,13	0,02	-0,002
2009	7	66,10	0,14	0,02	-0,15	0,02	-0,002
2010	8	69,00	0,13	0,01	-0,17	0,03	-0,002
2011	9	71,30	0,11	0,01	-0,18	0,03	-0,003
2012	10	73,40	0,10	0,01	-0,19	0,04	-0,003
			$\bar{t} = 0,293$	$\bar{y}' = 0,021$			
Σ						0,692	0,020

Źródło: obliczenia własne.

Po oszacowaniu parametrów funkcji trendu o malejącym tempie wzrostu, wyrażonego za pomocą funkcji ilorazowej, postać modelu analitycznego można przedstawić w następujący sposób:

$$\hat{y}_t = \frac{\beta_0 t}{\beta_1 + t} = \frac{76,923t}{2,231+t}.$$

Prognozę liczby gospodarstw domowych w Polsce wyposażonych w komputery wyrażoną w % ogółu gospodarstw domowych na rok 2013, 2014 i 2015 ($t = 11, 12, 13$) można wyznaczyć w następujący sposób:

$$y_{11}^* = \frac{76,923 \cdot 11}{2,231+11} = 63,95;$$

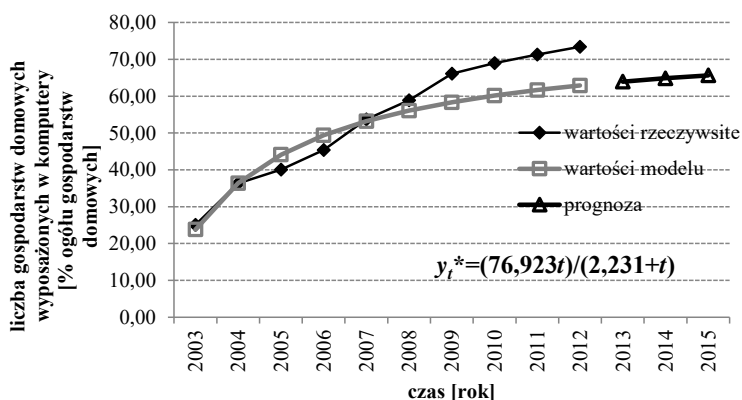
$$y_{12}^* = \frac{76,923 \cdot 12}{2,231+12} = 64,86;$$

$$y_{13}^* = \frac{76,923 \cdot 13}{2,231+13} = 65,66.$$

Zatem prognoza liczby gospodarstw domowych w Polsce wyposażonych w komputery w 2013 roku wynosi około 63,95%, natomiast

dla 2014 roku kształtuje się na poziomie 64,86%, a dla roku 2015 na poziomie 65,66%.

Wykres wartości rzeczywistych, wartości modelu oraz prognozę liczby gospodarstw domowych w Polsce wyposażonych w komputery (wyrażoną w % ogółu gospodarstw domowych) wyznaczoną z wykorzystaniem funkcji nieliniowej ilorazowej przedstawiono na rysunku 1.24. Analizując wykres, można stwierdzić, że wykorzystana do budowy modelu analitycznego funkcja ilorazowa o równaniu $\hat{y}_t = \frac{76,923t}{2,231+t}$ nie odzwierciedla poprawnie danych rzeczywistych, tj. nie podąża za zmianami zmiennej prognozowanej y .



Rys. 1.24. Wartości rzeczywiste, wartości modelu oraz prognoza liczby gospodarstw domowych wyposażonych w komputery w Polsce wyznaczona z wykorzystaniem funkcji ilorazowej (wyrażona w % ogółu gospodarstw domowych)

Źródło: opracowanie własne.

Przykład 1.9. Oszacowanie parametrów funkcji o malejącym tempie wzrostu – wielomian stopnia drugiego

Analizując dane zebrane w tabeli 1.7 i zobrazowane na rysunku 1.20, stwierdzono, że specyfikę tendencji rozwojowej o malejącym tempie wzrostu można odzwierciedlić także za pomocą funkcji o postaci wielomianu stopnia drugiego. Wyniki oceny wzrokowej przebiegu zmiennej

uzasadniają dalszy etap prac nad budowa modelu, tj. estymację parametrów wybranej postaci funkcji.

Parametry α_0 , α_1 , α_2 zostały wyznaczone przez rozwiązane równań (1.31-1.35).

Przy czym:

$$\begin{aligned}\sum_{t=1}^n t &= \frac{n(n+1)}{2} = \frac{10(10+1)}{2} = 55, \\ \sum_{t=1}^n t^2 &= \frac{n(n+1)(2n+1)}{6} = \frac{10(10+1)(20+1)}{6} = 385, \\ \sum_{t=1}^n t^3 &= \left[\frac{n(n+1)}{2} \right]^2 = \left[\frac{15(15+1)}{2} \right]^2 = 3025, \\ \sum_{t=1}^n t^4 &= \frac{n(n+1)(2n+1)}{6} \cdot \frac{3n^2 + 3n - 1}{5} = \\ &= \frac{10(10+1)(20+1)}{6} \cdot \frac{3(10)^2 + 3(10) - 1}{5} = 25333.\end{aligned}$$

Wartości pomocnicze, niezbędne do oszacowania parametrów α_0 , α_1 , α_2 , umieszczono w tabeli 1.12.

Tab. 1.12. Obliczenia pomocnicze

Rok	t	y_t	ty_t	$t^2 y_t$
2003	1	25,10	25,10	25,10
2004	2	36,30	72,60	145,20
2005	3	40,10	120,30	360,90
2006	4	45,40	181,60	726,40
2007	5	53,70	268,50	1342,50
2008	6	58,90	353,40	2120,40
2009	7	66,10	462,70	3238,90
2010	8	69,00	552,00	4416,00
2011	9	71,30	641,70	5775,30
2012	10	73,40	734,00	7340,00
Σ		539,30	3411,90	25490,70

Źródło: obliczenia własne.

Po rozwiązaniu układu równań (1.31-1.35) i uwzględnieniu wartości pomocniczych otrzymano:

$$\alpha_0 = 16,897;$$

$$\alpha_1 = 9,061;$$

$$\alpha_2 = -0,333.$$

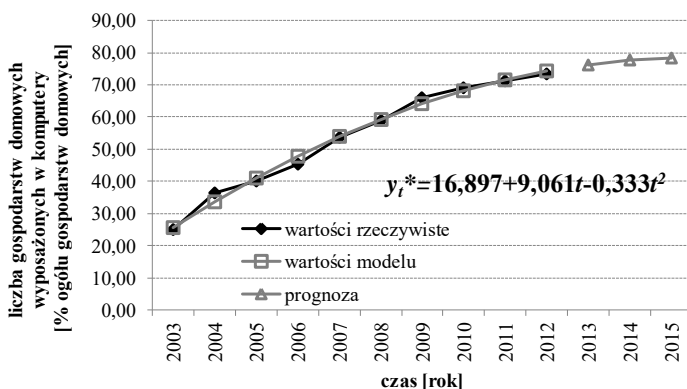
Model analityczny bazujący na ww. postaci funkcji można zatem przedstawić w następujący sposób: $\hat{y}_t = 16,897 + 9,061t - 0,333t^2$.

Prognozę liczby gospodarstw domowych w Polsce wyposażonych w komputery wyrażoną w % ogółu gospodarstw domowych na 2013 ($t = 1$), 2014 ($t = 12$) i 2015 rok ($t = 13$) można wyznaczyć w następujący sposób:

$$y_{11}^* = 16,897 + 9,061 \cdot 11 - 0,333 \cdot 11^2 = 76,28;$$

$$y_{12}^* = 16,897 + 9,061 \cdot 12 - 0,333 \cdot 12^2 = 77,68;$$

$$y_{13}^* = 16,897 + 9,061 \cdot 13 - 0,333 \cdot 13^2 = 78,41.$$



Rys. 1.25. Wartości rzeczywiste, wartości modelu oraz prognoza liczby gospodarstw domowych wyposażonych w komputery w Polsce wyznaczone z wielomianu stopnia drugiego (wyrażona w % ogółu gospodarstw domowych)

Źródło: opracowanie własne.

Prognoza liczby gospodarstw domowych w Polsce wyposażonych w komputery, wyrażona w % ogółu gospodarstw domowych, na rok 2013 wynosi 76,28%, na rok 2014 około 77,68%, natomiast na rok 2015 kształtuje się na poziomie 78,41%. Wykres wartości rzeczywistych, wartości modelu oraz prognozę liczby gospodarstw domowych w Polsce wyposażonych w komputery (wyrażoną w % ogółu gospodarstw domowych) wyznaczoną z wykorzystaniem wielomianu stopnia drugiego przedstawiono na rysunku 1.25.

Analizując wykres przedstawiony na rysunku 1.25, można zauważyć, że wykorzystana funkcja wielomianu stopnia drugiego o równaniu $\hat{y}_t = 16,897 + 9,061t - 0,333t^2$ dobrze odzwierciedla zmiany wartości rzeczywistych.

1.4.3. Funkcje nieliniowe o zmiennym tempie wzrostu

Obserwując przebieg zjawisk gospodarczych i przyrodniczych w dłuższych okresach, można dostrzec cały przebieg cyklu rozwojowego. Przykładowo, analizując krzywą życia produktu, można wyodrębnić poszczególne fazy rozwoju produktu, tj. fazę wprowadzania produktu na rynek, fazę przyspieszonego i malejącego tempa wzrostu popytu na produkt, stan nasycenia rynku i spadek popytu. Wówczas do opisu tendencji rozwojowych pierwszych trzech faz tych zjawisk można się posłużyć zamiast kilkoma funkcjami, odpowiednimi dla poszczególnych faz cyklu rozwojowego, tylko jedną. Przykładem takiej funkcji jest funkcja logistyczna postaci²⁸:

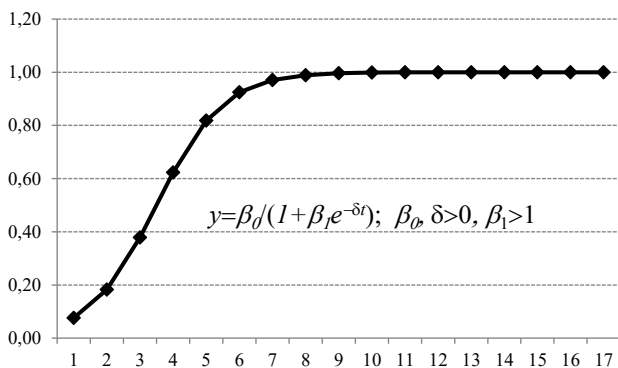
$$y_t = \frac{\beta_0}{1 + \beta_1 e^{-\delta t}}, \quad \beta_0, \delta > 0, \quad \beta_1 > 1. \quad (1.60)$$

Do punktu przegięcia (o współrzędnych $(1/\delta) \ln \beta_1$ dla zmiennej czasowej oraz β_0 dla zmiennej y) funkcja rośnie w tempie przyspieszonym, po czym rośnie w tempie malejącym do asymptoty poziomej (punkt nasycenia) $y = \beta_0$.

Funkcja logistyczna jest bardzo przydatna przy opisie i prognozowaniu wielu zjawisk ekonomicznych. Dobrze wyraża długotrwały wzrost pewnych podstawowych wielkości charakteryzujących rozwój

²⁸ P. Dittman, *Metody prognozowania sprzedaży w przedsiębiorstwie*, Wydawnictwo Akademii Ekonomicznej im. Oskara Langego we Wrocławiu, Wrocław 2000, s. 65.

gospodarczy, takich jak zatrudnienie, dochód narodowy itp. Wykorzystywana jest do opisu popytu na dobra trwałego użytku, na przykład samochody w krajach wysoko rozwiniętych, na telewizory, radioodbiorniki, dobrze obrazuje również wzrost liczby abonentów telefonicznych w Polsce²⁹. Graficzną postać funkcji logistycznej przedstawiono na rysunku 1.26.



Rys. 1.26. Graficzna reprezentacja funkcji logistycznej
($\beta_0 = 1$; $\beta_1 = 33$; $\delta = 1$)

Źródło: opracowanie własne.

Przebiegiem logistycznym charakteryzuje się również wiele zjawisk przyrodniczych, zwłaszcza tych, których rozwój jest ograniczony do pewnej przestrzeni. Przykładowo demografowie i biolodzy posługują się krzywą logistyczną do modelowania wzrostu liczby ludności. W Stanach Zjednoczonych funkcję logistyczną stosowano do zagadnień demograficznych, a w szczególności do opracowania prognozy rozwoju ludności na lata 1940-2100. Podejmowano też próby przedstawienia za pomocą krzywej logistycznej rozwoju poszczególnych gałęzi przemysłu, okręgów przemysłowych, często całego przemysłu i handlu³⁰.

Niektóre procesy gospodarcze, które podlegają tradycyjnie logistycznemu prawu wzrostu, można dobrze opisywać funkcją loglogistyczną. Sy-

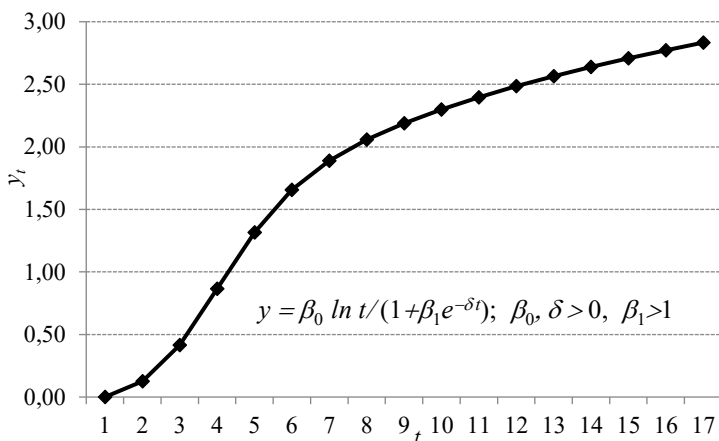
²⁹ T. Stanisławski, *Funkcje jednej zmiennej w badaniach ekonomicznych...*, op. cit., s. 133.

³⁰ Ibidem, s. 133-135.

tuacja taka ma miejsce w wypadku, gdy zjawisko charakteryzuje się stałym, nieograniczonym wzrostem, z malejącym do zera tempem wzrostu to do opisu i prognozowania jego przebiegu. Funkcja loglogistyczna ma postać³¹:

$$y_t = \frac{\beta_0 \ln t}{1 + \beta_1 e^{-\delta t}}, \quad \beta_0, \delta > 0, \quad \beta_1 > 1. \quad (1.61)$$

Graficzną postać funkcji loglogistycznej przedstawiono na rysunku 1.27.



Rys. 1.27. Graficzna reprezentacja funkcji loglogistycznej ($\beta_0 = 1; \beta_1 = 33; \delta = 1$)

Źródło: opracowanie własne.

Nieliniowy związek między zmiennymi i parametrami funkcji logistycznej i loglogistycznej uniemożliwia stosowanie KMNK. Większość znanych metod szacowania funkcji logistycznej (m.in. Hotellinga, Hellwiga) i loglogistycznych ma charakter jedynie przybliżony i opiera się na uproszczeniach³².

Szacowanie parametrów funkcji logistycznej i loglogistycznej jest skomplikowane i czasochłonne. Dlatego też do szacowania parametrów tych funkcji wykorzystuje się programy komputerowe. W dalszej części

³¹ M. Cieślak (red.), *Prognozowanie gospodarcze...*, op. cit., s. 79.

³² Ibidem, s. 81.

rozdziału zostaną przedstawione przykłady, w których oszacowano parametry funkcji o zmiennym tempie wzrostu z wykorzystaniem programu komputerowego Statistica 10.

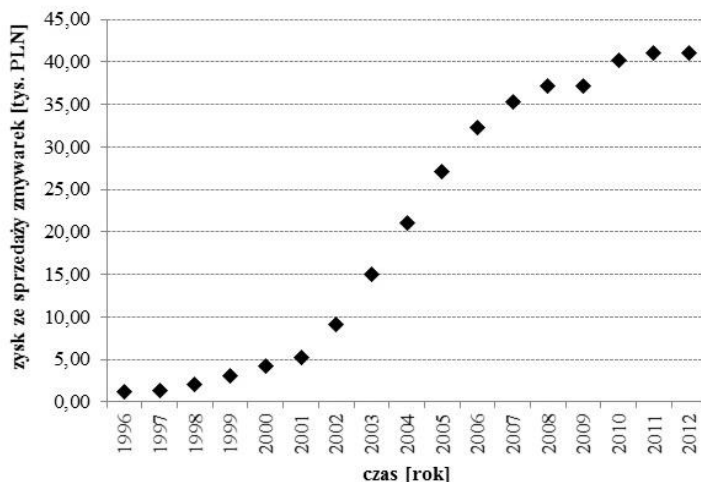
Przykład 1.10. Oszacowanie parametrów funkcji o zmiennym tempie wzrostu – funkcja logistyczna i loglogistyczna

Na podstawie zebranych danych, przedstawionych w tabeli 1.13 w postaci jednowymiarowego szeregu czasowego, można przeprowadzić estymację parametrów funkcji logistycznej i loglogistycznej.

Tab. 1.13. Dane dotyczące zysku ze sprzedaży zmywarek

Czas	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012
Zysk [tys. PLN]	1,21	1,32	2,11	3,11	4,23	5,21	9,11	15,02	21,05	27,08	32,23	35,23	37,11	37,14	40,23	41,11	41,11

Źródło: opracowanie własne na podstawie danych wewnętrznych sklepu z artykułami ADG.



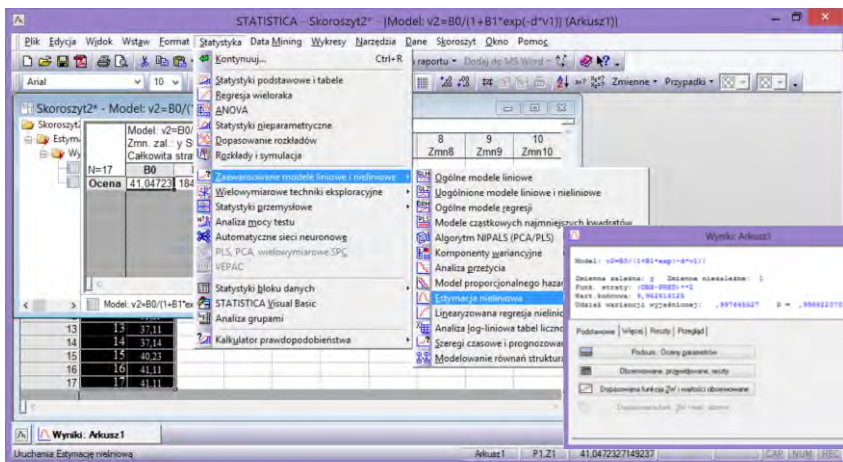
Rys. 1.28. Zysk ze sprzedaży zmywarek w jednym z podlaskich sklepów z artykułami ADG w latach 1996-2012 [tys. PLN]

Źródło: opracowanie własne.

Wykres danych dotyczących zysku ze sprzedaży zmywarek w jednym ze sklepów z artykułami AGD w latach 1996-2012 przedstawiono na rysunku 1.28.

Na podstawie analizy wizualnej wykresu przedstawionego na rysunku 1.28 można zauważyć, że początkowo stopniowy wzrost wartości zmiennej w miarę upływu czasu jest coraz szybszy, jednak po pewnym czasie tempo zmian jest coraz wolniejsze. W takim wypadku do budowy analitycznego modelu prognostycznego można zastosować jedną z funkcji o zmiennym tempie wzrostu, tj. funkcję logistyczną lub loglogistyczną.

Do oszacowania parametrów ww. funkcji wykorzystano pakiet komputerowy Statistica 10. Sposób wykorzystania tej aplikacji do rozwiązania tego zadania ilustruje rysunek 1.29.



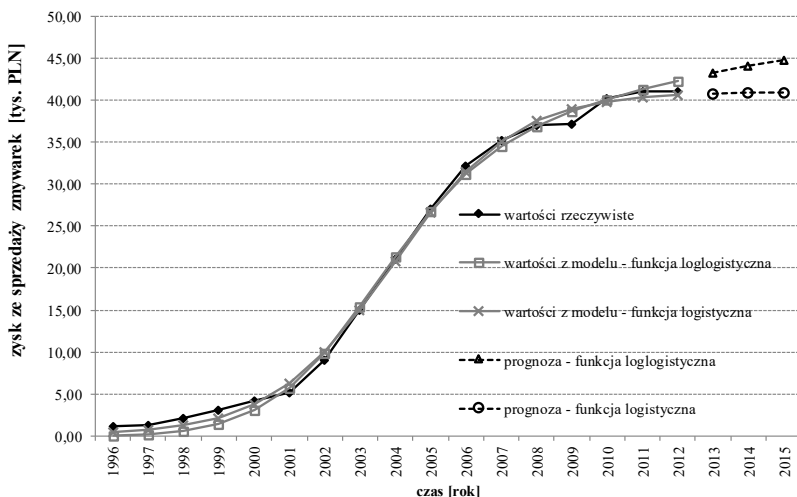
Rys. 1.29. Zastosowanie pakietu Statistica 10 do szacowania parametrów funkcji logistycznej i loglogistycznej

Źródło: opracowanie własne.

W wypadku funkcji logistycznej otrzymano następujące parametry: $\beta_0 = 41,047$; $\beta_1 = 184,711$; $\delta = 0,584$.

Z kolei dla funkcji loglogistycznej wyznaczono następujące parametry: $\beta_0 = 14,981$; $\beta_1 = 170,601$; $\delta = 0,639$.

Wykres wartości rzeczywistych, wartości teoretycznych wynikających z modeli oraz prognoz zysku ze sprzedaży w latach 2013-2015, wyznaczonych z wykorzystaniem funkcji logistycznej i loglogistycznej, przedstawiono na rysunku 1.30.



Rys. 1.30. Wartości rzeczywiste, wartości modelu oraz prognozy zysku ze sprzedaży zmywarek w jednym z podlaskich sklepów z artykułami ADG latach 1996-2015 [tys. PLN]

Źródło: opracowanie własne.

Analizując wykres przedstawiony na rysunku 1.30, można zauważyć, że zarówno funkcja logistyczna o równaniu $\hat{y}_t = \frac{41,047}{1 + 180,711e^{-0,584t}}$, jak i funkcja loglogistyczna o równaniu $\hat{y}_t = \frac{14,98 \ln t}{1 + 170,601e^{-0,639t}}$ dość do-
brze odzwierciedlają zmiany wartości rzeczywistych zmiennej prognozowanej.

Kluczowe zagadnienia:

trend, klasyczna metoda najmniejszych kwadratów, postacie funkcji trendu, wybór analitycznej postaci modelu, linearyzacja, funkcje o przyspieszonym tempie wzrostu, funkcje o malejącym tempie wzrostu, funkcje o zmiennym tempie wzrostu, funkcja potęgowa, funkcja wykładnicza, wielomian stopnia drugiego, funkcja logarytmiczna, funkcja liniowo-odwrotnościowa, funkcja ilorazowa, wielomian odwrotnościowy, funkcja logistyczna, funkcja loglogistyczna

Literatura podstawowa:

1. Nazarko J. (red.), *Prognozowanie w zarządzaniu przedsiębiorstwem. Cz. I. Wprowadzenie do metodyki prognozowania*, Wydawnictwo Politechniki Białostockiej, Białystok 2004.
2. Nazarko J. (red.), *Prognozowanie w zarządzaniu przedsiębiorstwem. Cz. II. Prognozowanie na podstawie szeregów czasowych*, Wydawnictwo Politechniki Białostockiej, Białystok 2004.
3. Nazarko J. (red.), *Prognozowanie w zarządzaniu przedsiębiorstwem. Cz. III. Prognozowanie na podstawie modeli adaptacyjnych*, Wydawnictwo Politechniki Białostockiej, Białystok 2005.
4. Sobczyk M., *Prognozowanie*, Placet, Warszawa 2008.

Literatura uzupełniająca:

1. Aczel A.D., *Statystyka w zarządzaniu*, Wydawnictwo Naukowe PWN, Warszawa 2000.
2. Cieślak M. (red.), *Prognozowanie gospodarcze. Metody i zastosowanie*, Wydawnictwo Naukowe PWN, Warszawa 2011.
3. Dittman P., *Prognozowanie w przedsiębiorstwie. Metody i ich zastosowanie*, Wolters Kluwer, Kraków 2009.

4. Dittman P., *Metody prognozowania sprzedaży w przedsiębiorstwie*, Wydawnictwo Akademii Ekonomicznej im. Oskara Langego we Wrocławiu, Wrocław 2000.
5. Halicka K., Godlewska J., *Prognozowanie średniego miesięcznego kursu kupna USD*, *Ekonomia i Zarządzanie*, t. 5, nr 1, Oficyna Wydawnicza Politechniki Białostockiej, Białystok 2013, s. 19-29.
6. Halicka K., Więnkowski C., *Wykorzystanie metod wygładzania wykładniczego do prognozowania kursu sprzedaży EUR*, *Ekonomia i Zarządzanie*, t. 5, nr 2, Oficyna Wydawnicza Politechniki Białostockiej, Białystok 2013, s. 70-80.
7. Maciąg A., Pietroń R., Kukła S., *Prognozowanie i symulacje w przedsiębiorstwie*, Polskie Wydawnictwo Ekonomiczne, Warszawa 2013, s. 77.
8. Nowak E., *Zarys metod ekonometrii. Zbiór zadań*, Wydawnictwo Naukowe PWN, Warszawa 2002.
9. Stanisław T., *Funkcje jednej zmiennej w badaniach ekonomicznych*, Wydawnictwo Naukowe PWN, Warszawa 1993.

2. METODA WSKAŹNIKÓW SEZONOWOŚCI

Metodę wskaźników sezonowości wykorzystuje się do prognozowania wartości zmiennych charakteryzujących się wahaniami sezonowymi lub cyklicznymi występującymi wraz z tendencją rozwojową lub stałym poziomem (patrz: *Prognozowanie w zarządzaniu przedsiębiorstwem*. Cz. II, rozdz. 4, s. 47-51³³).

U podstaw stosowania metody leży założenie, że składową systematyczną szeregu można rozłożyć na składową trendu oraz składową sezonową³⁴:

$$Y_t = T_t + S_t + C_t + E_t \quad (2.1)$$

lub

$$Y_t = T_t \cdot S_t \cdot C_t \cdot E_t, \quad (2.2)$$

gdzie:

Y_t – zmienna prognozowana,

T_t – składowa trendu,

C_t – składowa cykliczna,

S_t – składowa sezonowa,

E_t – składowa losowa.

Składową trendu i cykliczną modeluje się, korzystając z analitycznych funkcji trendu przedstawionych w rozdziale 1, zaś składnik okresowy metodą wywodzącą się z koncepcji indeksów (wskaźników), czyli liczb względnych, charakteryzujących, zazwyczaj w procentach, zmianę poziomu zjawiska zachodzącą w czasie.

Indeksy są stosunkowo często wykorzystywane do przedstawienia danych ekonomicznych, gdyż określając zmianę w stosunku do wartości poprzedniej (bazowej), w wielu wypadkach ułatwiają interpretację. Można się z nimi spotkać w raportach dotyczących bezrobocia, inflacji, produkcji, wynagrodzeń i wielu innych wielkości charakteryzujących

³³ J. Nazarko (red.), *Prognozowanie w zarządzaniu przedsiębiorstwem*. Cz. II., op. cit., s. 47-51.

³⁴ C.S. Hilas, S.K. Goudos, J.N. Sahalos, Seasonal decomposition and forecasting of telecommunication data: A comparative case study, *Technological Forecasting & Social Change* 73, 2006, s. 495–509.

tendencje rozwoju społeczno-gospodarczego. Przykładowe zastosowanie indeksów prostych, sposób ich obliczenia oraz interpretację dla stopy bezrobocia rejestrowanego zamieszczono w tabeli 2.1.

Tab. 2.1. Wykorzystanie indeksów na przykładzie stopy bezrobocia rejestrowanego

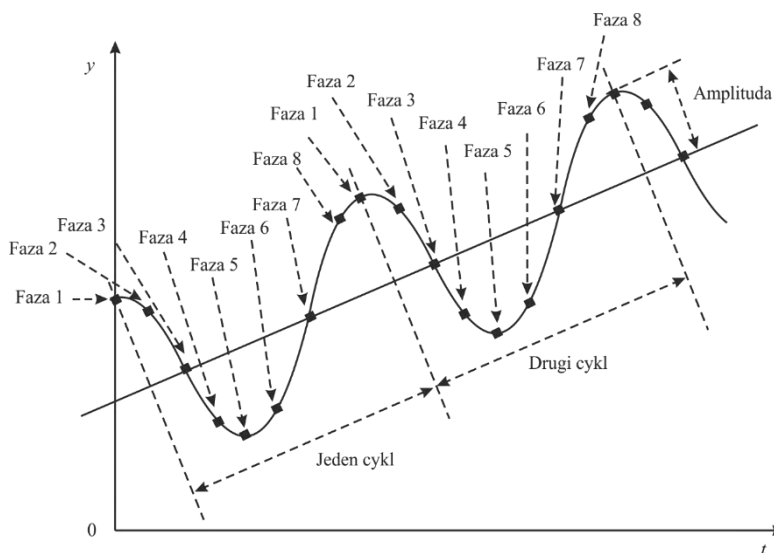
Mie- siąc	Stopa bezrobocia w 2012 [%]	Interpretacja stopy bezrobocia	Indeksy proste (zmiana w stosunku do poprzedniego miesiąca)	Interpretacja indeksu
I	13,2		–	–
II	13,4	$13,4 - 13,2 =$ $= 0,2$ Wzrost o 0,2 punktu procentowego (p.p.)	$13,4 / 13,2 =$ $= 1,015$	Wartość w miesiącu II stanowi 101,5% wartości z miesiąca I. Wzrost o 1,5% wartości z miesiąca I
III	13,3	$13,3 - 13,4 =$ $= - 0,1$ Spadek o 0,1 p.p.	$13,3 / 13,4 =$ $= 0,993$	Wartość w miesiącu III stanowi 99,3% wartości z miesiąca II. Spadek o 100% – 99,3% = 0,7% wartości z miesiąca II
			...	
X	12,5			
XI	12,9	Wzrost o 0,4 p.p.	$12,9 / 12,5 =$ $= 1,032$	Wzrost o 3,2%
XII	13,4	Wzrost o 0,5 p.p.	$13,4 / 12,9 =$ $= 1,039$	Wzrost o 3,9%

Źródło: opracowanie własne na podstawie danych opublikowanych na stronie internetowej Głównego Urzędu Statystycznego, <http://www.stat.gov.pl>, data wejścia 29.04.2016.

Wyjaśnienie koncepcji metody wskaźników sezonowości wymaga przypomnienia kilku podstawowych pojęć związanych z szeregami czasowym, w których występują wahania sezonowe, takich jak: faza, cykl, amplituda.

Fazą określa się momenty lub okresy odległe od siebie o stały okres, w których dane zjawisko tak samo się kształtuje, np. szybki wzrost, lekki wzrost, spadek. Cykl to okres obejmujący wszystkie fazy wahań. Amplituda mierzy zakres wahań, czyli największe odchylenie wartości zmiennej od stałego poziomu lub trendu. Powyższe definicje zilustrowano

na rysunku 2.1. Przedstawiono na nim dwa pełne, ośmifazowe cykle szeregu niestacjonarnego.



Rys. 2.1. Wahania sezonowe – objaśnienie wybranych pojęć

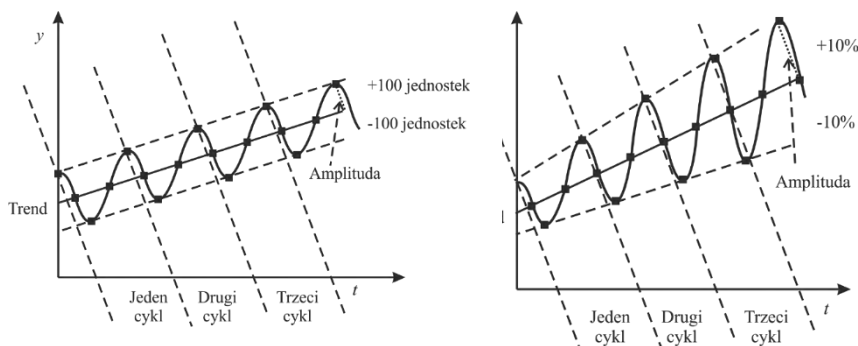
Źródło: opracowanie własne na podstawie: A. Zeliaś, B. Pawełek, S. Wanat, *Prognozowanie ekonomiczne: teoria, przykłady, zadania*, Wydawnictwo Naukowe PWN, Warszawa 2004, s. 87.

W metodzie wskaźników sezonowości wylicza się zmiany, wyrażone w procentach, wartości badanej zmiennej w każdej z faz cyklu sezonowego, czyli wskaźniki sezonowości. Obliczone wskaźniki są następnie wykorzystywane do korygowania wartości trendu lub stałego poziomu.

Wyróżnia się dwa modele prognostyczne: addytywny (2.1) i multiplikatywny (2.2), które stosuje się w zależności od charakteru wzajemnych relacji występujących w szeregu wahań sezonowych.

Model addytywny wykorzystuje się w wypadku, gdy wahania sezonowe są nieskorelowane ze stałym poziomem lub trendem (wahania bezwzględnie stałe), czyli amplitudy wahań w analogicznych fazach cyklu są w przybliżeniu takie same. Model multiplikatywny stosuje się, jeżeli

amplitudy wahań zmieniają się w tym samym stosunku, tzn. wartość szeregu w danej fazie zależy od poziomu trendu (wahania względnie stałe)³⁵. Wybór właściwego modelu zazwyczaj dokonywany jest na podstawie wizualnej analizy zmienności szeregu przy uwzględnieniu natury zjawiska. Na rysunku 2.2 przedstawiono schematycznie szeregi, do których prognozowania można użyć modelu addytywnego – amplituda wahań sezonowych wynosi 100 jednostek oraz modelu multiplikatywnego, w którym wahania sezonowe stanowią 10% poziomu trendu.



Rys. 2.2. Sezonowość addytywna i multiplikatywna

Źródło: opracowanie własne na podstawie: S.A. DeLurgio, *Forecasting Principles and Applications*, Irwin/McGraw-Hill, Boston 1998, s. 177.

W modelu addytywnym wartość zmiennej prognozowanej jest sumą oszacowanego stałego poziomu lub trendu, wskaźnika sezonowości oraz składnika losowego. Model ten można zapisać za pomocą równania³⁶:

$$y_{t_{lj}} = f(t_{lj}) + s_j + \xi_{t_{lj}}, \quad t_{lj} = 1, \dots, n, \quad l = 1, \dots, N, \quad j = 1, \dots, m, \quad (2.3)$$

gdzie:

$y_{t_{lj}}$ – wartość zmiennej prognozowanej w momencie lub okresie t_{lj} ,

czyli w j -tej fazie l -tego cyklu: $t_{lj} = m(l-1) + j$;

³⁵ S.A. DeLurgio, *Forecasting Principles and Applications*, Irwin/McGraw-Hill, Boston 1998, s. 177.

³⁶ A. Zeliaś, B. Pawełek, S. Wanat, *Prognozowanie ekonomiczne: teoria, przykłady, zadania*, Wydawnictwo Naukowe PWN, Warszawa 2004, s. 91.

$f(t_{lj})$ – wartość zmiennej w momencie lub okresie t_{lj} oszacowana za pomocą właściwego modelu analitycznego lub stały poziom;

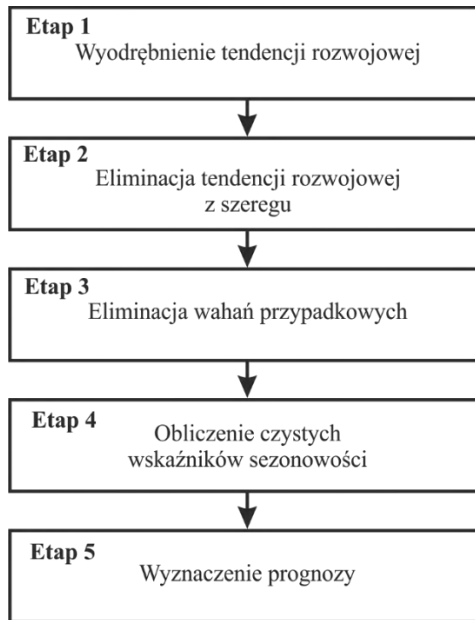
s_j – wskaźnik sezonowości dla j -tej fazy cyklu;

$\xi_{t_{lj}}$ – składnik losowy w momencie lub okresie t (w j -tej fazie l -tego cyklu);

n – liczba obserwacji w szeregu;

m – liczba faz w cyklu;

N – liczba cykli.



Rys. 2.3. Etapy opracowywania prognozy na podstawie metody wskaźników sezonowości

Źródło: opracowanie własne na podstawie: A. Zeliaś., B. Pawełek, S. Wanat, *Prognozowanie ekonomiczne: teoria, przykłady, zadania*, Wydawnictwo Naukowe PWN, Warszawa 2004, s. 91.

W wypadku modelu multiplikatywnego wartość zmiennej prognozowanej jest iloczynem stałego poziomu lub trendu, wskaźnika sezonowości oraz składnika losowego):

$$y_{t_{ij}} = f(t_{ij})s_j \xi_{t_{ij}}, \quad t_{ij} = 1, \dots, n, \quad l = 1, \dots, N, \quad j = 1, \dots, m, \quad (2.4)$$

gdzie: oznaczenia jak we wzorze 2.3.

W procesie opracowywania prognozy za pomocą metody wskaźników sezonowości można wyróżnić pięć zasadniczych etapów³⁷. Zilustrowano je na rysunku 2.3.

Etap 1. Wyodrębnienie tendencji rozwojowej $f(t_{ij})$

W pierwszym kroku należy określić analityczną postać funkcji trendu za pomocą rozwiązań przedstawionych w rozdziale 1. Alternatywnym, w stosunku do klasycznej metody najmniejszych kwadratów, podejściem do określenia współczynników liniowej funkcji trendu $f(t_{ij})$ dla zmiennej Y może być skorzystanie z następującego uproszczonego algorytmu³⁸:

1. Różnicuje się szereg z opóźnieniem równym długości cyklu w celu usunięcia trendu i oblicza średnią zróżnicowanego szeregu (np. dla danych miesięcznych długość cyklu to wynosi 12, dla kwartalnych 4).
2. Otrzymaną wartość średniej zróżnicowanego szeregu dzieli się przez długość jednego cyklu, aby otrzymać wartość trendu w jednym momencie lub okresie, czyli współczynnik kierunkowy funkcji trendu β_1 .
3. W celu oszacowania stałej β_0 liczy się średnią oryginalnego szeregu \bar{Y} , a także średnią czasu \bar{t} . Następnie korzysta się z faktu, że linia trendu przechodzi przez punkt (\bar{t}, \bar{Y}) , tzn.: $\bar{Y} = \beta_0 + \beta_0 \bar{t}$.
4. Funkcja trendu ma postać: $f^*(t_{ij}) = \beta_0 + \beta_1 t_{ij}$.

³⁷ Ibidem, s. 91.

³⁸ S.A. DeLurgio, *Forecasting Principles...*, op. cit., s. 249.

Etap 2. Eliminacja tendencji rozwojowej z szeregu

Eliminacja tendencji rozwojowej z szeregu w wypadku budowy modelu addytywnego polega na obliczeniu różnic między rzeczywistymi wartościami prognozowanej zmiennej a wartościami otrzymanymi z określonego w etapie pierwszym modelu tendencji rozwojowej:

$$z_{t_{ij}} = y_{t_{ij}} - f^*(t_{ij}), \quad (2.5)$$

gdzie:

$y_{t_{ij}}$ – wartość zmiennej prognozowanej w momencie lub okresie t_{ij} ,

$f^*(t_{ij})$ – wartość zmiennej w momencie lub okresie t_{ij} oszacowana za pomocą właściwego modelu analitycznego lub stały poziom.

W wypadku modelu multiplikatywnego należy obliczyć iloraz odpowiednich wartości szeregu czasowego oraz wartości obliczonych na podstawie modelu trendu:

$$z_{t_{ij}} = \frac{y_{t_{ij}}}{f^*(t_{ij})}, \quad (2.6)$$

gdzie: oznaczenia jak we wzorze 2.3.

Otrzymany w ten sposób szereg $z_{t_{ij}}$ zawiera tylko wahania przypadkowe i sezonowe.

Etap 3. Eliminacja wahań przypadkowych

W kolejnym kroku należy obliczyć tzw. surowe wskaźniki sezonowości dla każdej j -tej fazy cyklu według wzoru:

$$s'_j = \frac{1}{N} \sum_{t=1}^N z_{t_{ij}} = \frac{1}{N} (z_{t_{1j}} + z_{t_{2j}} + \dots + z_{t_{Nj}}), \quad j=1, 2, \dots, m, \quad (2.7)$$

gdzie:

$z_{t_{ij}}$ – wartość zmiennej szeregu bez tendencji rozwojowej w momencie lub okresie t_{ij} ,

N – liczba cykli.

Surowy wskaźnik sezonowości dla j -tej fazy jest to średnia arytmetyczna wartości szeregu bez trendu w tej fazie w każdym cyklu.

Etap 4. Obliczenie czystych wskaźników sezonowości

W wypadku, gdy dla modelu addytywnego suma surowych wskaźników sezonowości $\sum_{j=1}^m s'_j$ jest różna od zera, powinno się skorygować wskaźniki sezonowości poprzez odjęcie od każdego z nich ich średniej arytmetycznej $\overline{s'_j}$:

$$s_j = s'_j - \overline{s'_j}, \quad (2.8)$$

gdzie: $\overline{s'_j} = \frac{1}{m} \sum_{j=1}^m s'_j$.

W przeciwnym wypadku $s_j = s'_j$. Podobnie dla modelu multiplikatywnego, gdy suma wskaźników sezonowości $\sum_{j=1}^m s'_j$ jest różna od m , tzn. liczby faz tworzących cykl, dokonuje się korekcji według następującego wzoru:

$$s_j = \frac{s'_j}{s'}. \quad (2.9)$$

W ten sposób otrzymuje się tzw. czyste wskaźniki sezonowości. Wskaźniki te informują o zaobserwowanych w analizowanym szeregu średnich bezwzględnych (w wypadku modelu addytywnego) lub względnych (dla modelu multiplikatywnego) odchyleniach wartości prognozowanej od funkcji trendu w poszczególnych fazach cyklu sezonowego.

Etap 5. Wyznaczenie prognozy

Prognozę za pomocą metody wskaźników sezonowości oblicza się poprzez ekstrapolację linii trendu i jej korekcję obliczonymi wskaźnikami sezonowości dla danej fazy cyklu. Wymaga to przyjęcia postawy pasywnej i założenia, że w okresie prognozowanym utrzyma się zaobserwowana tendencja rozwojowa, a rodzaj i siła wahań nie ulegną zmianie³⁹.

³⁹ B. Radzikowska (red.), *Metody prognozowania. Zbiór zadań*, Wydawnictwo Akademii Ekonomicznej im. Oskara Langego we Wrocławiu, Wrocław 2000, s. 43.

Prognostyczną wartość zmiennej na moment lub okres t_{lj} dla modelu addytywnego można wyznaczyć jako sumę tendencji rozwojowej/stałego poziomu i czystego wskaźnika sezonowości:

$$y_{t_{lj}}^* = f^*(t_{lj}) + s_j, \quad t_{lj} > n, \quad l = N + 1, N + 2, \dots, \quad j = 1, \dots, m, \quad (2.10)$$

gdzie:

- $y_{t_{lj}}^*$ – prognoza wartości zmiennej w momencie lub okresie t_{lj}
(j -tej fazie l -tego cyklu),
- $f^*(t_{lj})$ – wartość oszacowana za pomocą funkcji trendu w okresie prognozowanym t_{lj} ,
- s_j – czysty wskaźnik sezonowości dla j -tej fazy cyklu,
- n – liczba obserwacji w szeregu,
- N – liczba cykli w szeregu,
- l – numer cyklu,
- j – numer fazy w cyklu,
- m – liczba faz w cyklu.

W wypadku przyjęcia modelu multiplikatywnego oblicza się iloczyn tendencji rozwojowej/stałego poziomu i czystego wskaźnika sezonowości:

$$y_{t_{lj}}^* = f^*(t_{lj})s_j, \quad t_{lj} > n, \quad l = N + 1, N + 2, \dots, \quad j = 1, \dots, m, \quad (2.11)$$

gdzie: oznaczenia jak we wzorze 2.10.

Przykład 2.1. Prognoza wielkości produktu krajowego brutto w cenach bieżących metodą wskaźników sezonowości

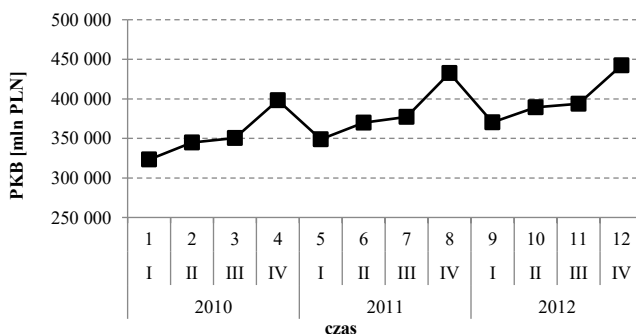
Produkt krajowy brutto (PKB) w cenach bieżących w mln PLN w Polsce w latach 2010-2012 przedstawiono w tabeli 2.2. Należy oszacować wartość produkcji w kolejnych kwartałach roku 2013.

Tab. 2.2. Produkt krajowy brutto w cenach bieżących w mln PLN

<i>t</i>	Rok	Kwartał	PKB [mln PLN]
			y_t
1	2010	I	323206,6
2		II	344700,0
3		III	350607,2
4		IV	398071,5
5	2011	I	348858,9
6		II	369749,9
7		III	377079,8
8		IV	432438,6
9	2012	I	370117,7
10		II	389122,4
11		III	393792,4
12		IV	442231,4

Źródło: opracowanie własne na podstawie danych opublikowanych na stronie internetowej Głównego Urzędu Statystycznego, <http://www.stat.gov.pl>, data wejścia 29.04.2016.

Analiza wizualna wykresu szeregu czasowego przedstawionego na rysunku 2.4 pozwala stwierdzić, że w kolejnych latach wartość produktu krajowego brutto systematycznie rosła.



Rys. 2.4. Produkt krajowy brutto w cenach bieżących w mln PLN

Źródło: opracowanie własne na podstawie danych opublikowanych na stronie internetowej Głównego Urzędu Statystycznego, <http://www.stat.gov.pl>, data wejścia 29.04.2016.

Szereg ten charakteryzuje się też wyraźną sezonowością: w każdym z badanych lat najniższa wartość PKB występuje w I kwartale, w kolejnych jest coraz większa, a wartość maksymalną badana zmienna osiąga w kwartale IV. Oznacza to, że przedstawione dane obejmują 3 roczne cykle, każdy składający się z 4 faz. Można też przyjąć, że amplitudy wahań w analogicznych fazach cyklu są w przybliżeniu takie same.

Ponieważ wahania w analizowanym szeregu są raczej bezwzględnie stałe, właściwym modelem prognostycznym będzie model addytywny metody wskaźników sezonowości. Opracowanie prognozy wartości PKB w kolejnym cyklu za pomocą metody wskaźników sezonowości wymaga wyodrębnienia i eliminacji tendencji rozwojowej z szeregu (etap 1 i 2), eliminacji wahań przypadkowych (etap 3) oraz obliczenia wskaźników sezonowości (etap 4). Prognozowana wartość zmiennej będzie wartością wyodrębnionego trendu skorygowanego obliczonymi wskaźnikami sezonowości (etap 5).

Etap 1. Wyodrębnienie tendencji rozwojowej

W celu określenia analitycznej postaci tendencji rozwojowej oszacowano parametry liniowej funkcji trendu za pomocą klasycznej metody najmniejszych kwadratów. Obliczono także podstawowe charakterystyki prostej regresji. Zestawienie wyników zawiera tabela 2.3.

Tab. 2.3. Parametry liniowej funkcji trendu oraz ich ocena

	Współczynnik kierunkowy	Stała
	7479,1	329717,5
Standardowe wartości błędu	1997,7	14702,6
Statystyka t-Studenta	3,744	22,426
Charakterystyki modelu		
Współczynnik determinacji R^2	0,584	
Statystyka F-Snedecora	14,016	
Stopnie swobody	10	
Wartość krytyczna rozkładu F-Snedecora	4,965	
Wartość krytyczna rozkładu t-Studenta	2,228	

Źródło: obliczenia własne za pomocą programu STATISTICA 10.

Oszacowana funkcja trendu ma postać: $f^*(t_{lj}) = 32971,75 + 7479,1t_{lj}$.

Dodatkowe statystyki, przedstawione w tabeli 2.3, służą sprawdzeniu istotności parametrów wybranej funkcji trendu oraz weryfikacji jej dopasowania do danych empirycznych. W wypadku wybranego modelu trendu uogólniony test Walda wykorzystujący statystykę *F-Snedecora* pozwala odrzucić hipotezę o braku zależności między wartością PKB a zmienną objaśniającą t . Wartość tej statystyki wynosi 14,016 i na poziomie prawdopodobieństwa $\alpha = 0,05$ przekracza wartość krytyczną wynoszącą 4,965. Podobnie test istotności parametrów modelu wykorzystujący statystykę *t-Studenta* wskazuje, że zmienna objaśniająca jest użyteczna w szacowaniu przewidywanej wartości PKB. Wartość bezwzględna statystyki t równa 3,744 jest większa od wartości krytycznej wynoszącej 2,228 dla $\alpha = 0,05$ i 10 stopni swobody. Współczynnik determinacji R^2 informuje o tym, jaka część zmienności zmiennej objaśnianej została wyjaśniona przez model. Przyjmuje wartości z przedziału $\langle 0;1 \rangle$, a dopasowanie modelu jest tym lepsze, im wartość R^2 jest bliższa jedności. W wypadku oszacowanego modelu R^2 , wynoszące 0,583, wskazuje na raczej słabe dopasowanie, lecz wynika to z istnienia składowej okresowej w szeregu. Podsumowując, oszacowaną funkcję $f^*(t_{lj}) = 32971,75 + 7479,1t_{lj}$ można uznać za właściwie modelującą składową trendu w analizowanym szeregu.

Etap 2. Eliminacja tendencji rozwojowej z szeregu

Obliczenie wskaźników sezonowości wymaga usunięcia trendu z szeregu. W modelu addytywnym eliminacja trendu polega na odjęciu od wyrazów oryginalnego szeregu odpowiadającym im wyrazów z szeregu otrzymanego na podstawie modelu tendencji rozwojowej. Otrzymuje się w ten sposób szereg $z_{t_{lj}}$, zawierający jedynie wahania przypadkowe i sezonowe. Dla okresu pierwszego (pierwszej fazy pierwszego cyklu, czyli $t = 1, l = 1, j = 1$) wartość szacowana modelem trendu wynosi:

$$f^*(1_{11}) = 32971,75 + 7479,1 \cdot 1 = 3371966.$$

Wartość szeregu bez trendu w okresie $t = 1$ wynosi zatem:

$$z_{1_{11}} = y_{1_{11}} - f^*(1_{11}) = 3232066 - 3371966 = -139900.$$

Podobnie dla okresu drugiego i kolejnych:

$$f^*(2_{12}) = 329717,5 + 74791 \cdot 2 = 3446757;$$

$$z_{2_{12}} = y_{2_{12}} - f^*(2_{12}) = 3447000 - 3446757 = 24,3;$$

⋮

$$f^*(12_{34}) = 329717,5 + 74791 \cdot 12 = 4194667;$$

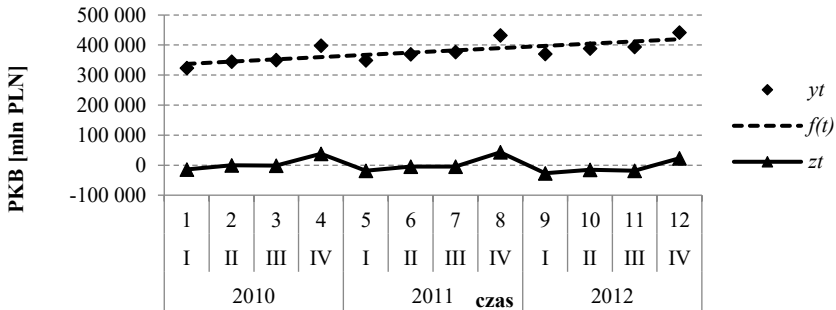
$$z_{12_{34}} = y_{12_{34}} - f^*(12_{34}) = 4422314 - 4194667 = 227647.$$

Zestawienie rezultatów obliczeń dla wszystkich okresów zawarto w tabeli 2.4 i zilustrowano na rysunku 2.5.

Tab. 2.4. Eliminacja trendu z szeregu

t_{ij}	Rok	Kwartal	Numer cyklu	Numer fazy	PKB	$f(t_{ij})$	$z_{t_{ij}} = y_{t_{ij}} - f(t_{ij})$
			l	j	$y_{t_{ij}}$		
1	2010	I	1	1	323206,6	337196,6	-13990,0
2		II		2	344700,0	344675,7	24,3
3		III		3	350607,2	352154,8	-1547,6
4		IV		4	398071,5	359633,9	38437,6
5	2011	I	2	1	348858,9	367113,0	-18254,1
6		II		2	369749,9	374592,1	-4842,2
7		III		3	377079,8	382071,2	-4991,4
8		IV		4	432438,6	389550,3	42888,3
9	2012	I	3	1	370117,7	397029,4	-26911,7
10		II		2	389122,4	404508,5	-15386,1
11		III		3	393792,4	411987,6	-18195,2
12		IV		4	442231,4	419466,7	22764,7

Źródło: obliczenia własne.



Rys. 2.5. Wartości produktu krajowego brutto w cenach bieżących w mln PLN – y_t , funkcji trendu – $f^*(t)$ oraz szeregu bez trendu – z_t

Źródło: opracowanie własne.

Etap 3. Eliminacja wahań przypadkowych – wyznaczenie surowych wskaźników sezonowości

Surowe wskaźniki sezonowości oblicza się jako średnią arytmetyczną wartości z danej fazy z każdego cyklu zgodnie ze wzorem (2.7). Ponieważ, jak już stwierdzono wcześniej, w szeregu liczącym $n = 12$ obserwacji występują wahania kwartalne, czyli liczba cykli wahań wynosi $N = 3$, a każdy cykl składa się z $m = 4$ faz, należy obliczyć 4 wskaźniki, każdy na podstawie wartości 3 obserwacji. Pierwszy wskaźnik sezonowości wyznaczono, wybierając z tabeli 2.4 wartości z pierwszego kwartału poszczególnych lat, tj. $t = 1_{11}, 5_{21}, 9_{31}$:

$$\begin{aligned}
 s'_1 &= \frac{1}{3} \sum_{l=1}^3 z_{t_{1l}} = \frac{1}{3} (-139900 - 182541 - 269117) = \\
 &= -\frac{1}{3} \cdot 591558 = -197186.
 \end{aligned}$$

Podobnie dla kolejnych faz cyklu:

$$\begin{aligned}
 s'_2 &= \frac{1}{3} \sum_{l=1}^3 z_{t_{12}} = \frac{1}{3} (24,3 - 4842,2 - 15386,1) = \\
 &= -\frac{1}{3} \cdot 202040 = -67346,7;
 \end{aligned}$$

$$s'_3 = \frac{1}{3} \sum_{l=1}^3 z_{t_{l3}} = \frac{1}{3} (-1547,6 + -4991,4 - 18195,2) =$$

$$= -\frac{1}{3} \cdot 24734,2 = -8244,73;$$

$$s'_4 = \frac{1}{3} \sum_{l=1}^3 z_{t_{l4}} = \frac{1}{3} (38437,6 + 42888,3 + 22764,7) =$$

$$= \frac{1}{3} \cdot 104090,6 = 34696,87.$$

Etap 4. Obliczenie czystych wskaźników sezonowości

Suma obliczonych w poprzednim etapie surowych wskaźników sezonowości jest równa:

$$\sum_{j=1}^4 s'_j = -19718,60 - 6734,67 - 8244,73 + 34696,87 = -1,13.$$

Wartość sumy surowych wskaźników sezonowości nie różni się znacząco od 0, biorąc pod uwagę bezwzględne wartości wskaźników, ale można dodatkowo skorygować surowe wskaźniki poprzez odjęcie od nich ich średniej arytmetycznej:

$$\bar{s}' = \frac{1}{4} \sum_{j=1}^4 s'_j = -\frac{1}{4} \cdot 1,13 = -0,28.$$

Wartość skorygowana, czyli czysty wskaźnik sezonowości fazy 1, wynosi: $s_1 = s'_1 - \bar{s}' = -19718,60 - (-0,28) = -19718,32$.

Analogicznie dla fazy drugiej:

$$s_2 = s'_2 - \bar{s}' = -6734,67 + 0,28 = -6734,39,$$

trzeciej: $s_3 = s'_3 - \bar{s}' = -8244,73 + 0,28 = -8244,45$

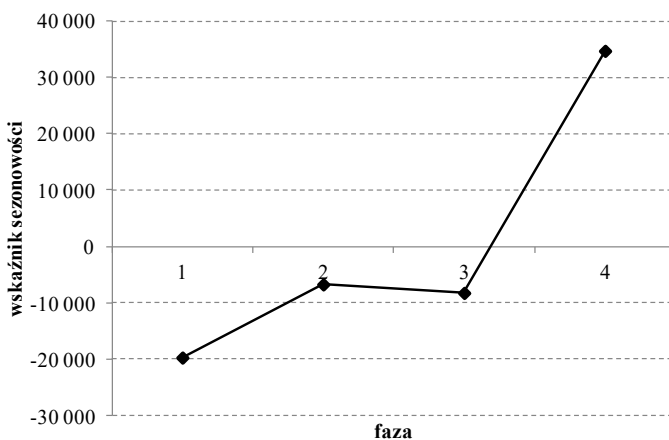
i czwartej: $s_4 = s'_4 - \bar{s}' = 34696,87 + 0,28 = 34697,15$.

Zestawienie wartości skorygowanych wartości wskaźników sezonowości dla wszystkich faz umieszczono w tabeli 2.5, zaś ich graficzną prezentację na rysunku 2.6.

Tab. 2.5. Kwartalne wskaźniki sezonowości

Numer fazy	Wskaźnik sezonowości
j	s_j
1	-19718,32
2	-6734,39
3	-8244,45
4	34697,15

Źródło: obliczenia własne.



Rys. 2.6. Kwartalne wskaźniki sezonowości

Źródło: opracowanie własne.

Obliczone wskaźniki sezonowości oznaczają, że w kwartale I wartość produkcji jest o -19718,32 mln PLN niższa od wartości wynikającej z oszacowanej linii trendu, w kwartale II jest o -6734,39 mln PLN niższa, w III o -8244,45 mln PLN, a w IV o 34697,15 mln PLN wyższa.

Korzystając z oszacowanych wskaźników sezonowości oraz funkcji trendu określonego w etapie 1, można obliczyć wartości modelu. W tym celu należy zsumować wartości funkcji trendu w każdym okresie oraz odpowiadający danej fazie wskaźnik sezonowości:

$$y_{1,1}^* = f^*(1_{1,1}) + s_1 = 3371966 - 1971832 = 3174783.$$

Podobnie dla okresu 2 (cykl $l=1$ faza $j=2$) i kolejnych:

$$y_{2_{12}}^* = f^*(2_{12}) + s_2 = 344675,7 - 6734,39 = 3379413;$$

$$y_{3_{13}}^* = f^*(3_{13}) + s_3 = 352154,8 - 8244,45 = 3439104;$$

$$y_{4_{14}}^* = f^*(4_{14}) + s_4 = 359633,9 + 34697,15 = 3943311;$$

$$y_{5_{21}}^* = f^*(5_{12}) + s_1 = 367113,0 - 1971832 = 3473947;$$

⋮

$$y_{12_{34}}^* = f^*(12_{34}) + s_4 = 419466,7 + 34697,15 = 454163,9.$$

Etap 5. Wyznaczenie prognozy

Zakładając, że utrzyma się dotychczasowa tendencja rozwojowa, a siła wahań sezonowych nie zmieni się, można obliczyć prognozę na kolejne cztery kwartały. Dokonując ekstrapolacji linii trendu na momenty 13, 14, 15 i 16 i korygując otrzymaną wartość odpowiednim wskaźnikiem sezonowości, uzyskano następujące wyniki:

$$y_{13_{41}}^* = f^*(13_{41}) + s_1 = 426945,8 - 19718,32 = 407227,5;$$

$$y_{14_{42}}^* = f^*(14_{42}) + s_2 = 434424,9 - 6734,39 = 427690,5;$$

$$y_{15_{43}}^* = f^*(15_{43}) + s_3 = 441904,0 - 8244,45 = 433659,6;$$

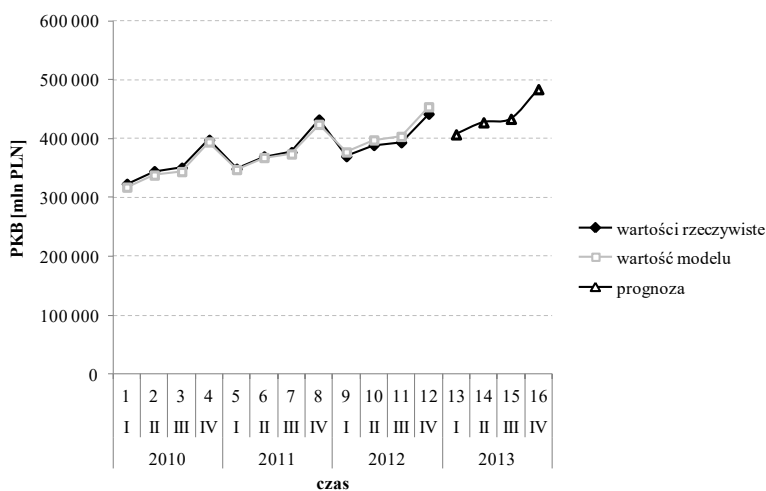
$$y_{16_{44}}^* = f^*(16_{44}) + s_4 = 449383,1 + 34697,15 = 484080,3.$$

Zestawienie rezultatów obliczeń dla wszystkich okresów t oraz prognozę zawiera tabela 2.6.

Tab. 2.6. Wartości modelu i prognozy PKB w mln PLN

t	Rok	Kwartał	PKB	Funkcja trendu	Wskaźnik sezonowości	Model i prognoza
			y_t	$f^*(t_{ij})$	s_j	$y_{tij}^* = f^*(t_{ij}) + s_j$
1	2010	I	323206,6	337196,6	-19718,32	317478,3
2		II	344700,0	344675,7	-6734,39	337941,3
3		III	350607,2	352154,8	-8244,45	343910,4
4		IV	398071,5	359633,9	34697,15	394331,1
5	2011	I	348858,9	367113,0	-19718,32	347394,7
6		II	369749,9	374592,1	-6734,39	367857,7
7		III	377079,8	382071,2	-8244,45	373826,8
8		IV	432438,6	389550,3	34697,15	424247,5
9	2012	I	370117,7	397029,4	-19718,32	377311,1
10		II	389122,4	404508,5	-6734,39	397774,1
11		III	393792,4	411987,6	-8244,45	403743,2
12		IV	442231,4	419466,7	34697,15	454163,9
13	2013	I		426945,8	-19718,32	407227,5
14		II		434424,9	-6734,39	427690,5
15		III		441904,0	-8244,45	433659,6
16		IV		449383,1	34697,15	484080,3

Źródło: obliczenia własne.



Rys. 2.7. Wartości rzeczywiste, modelu i prognozy kwartalnego produktu krajowego brutto

Źródło: opracowanie własne.

Wykres danych rzeczywistych, wartości modelu oraz prognozy PKB na 2013 rok przedstawiono na rysunku 2.7.

Jeżeli nie wystąpią nieoczekiwane dodatkowe przesłanki mogące zakłócić systematyczny wzrost PKB w Polsce, można przyjąć, że PKB w kolejnych kwartałach 2013 roku wynosić będzie odpowiednio: 407227,5 mln PLN, 427690,5 mln PLN, 433659,6 mln PLN i 484080,3 mln PLN.

Kluczowe zagadnienia:

metoda wskaźników sezonowości, cykl, faza, amplituda,
model addytywny wskaźników sezonowości,
model multiplikatywny wskaźników sezonowości

Literatura podstawowa:

1. Nazarko J. (red.), *Prognozowanie w zarządzaniu przedsiębiorstwem. Cz. II. Prognozowanie na podstawie szeregów czasowych*, Wydawnictwo Politechniki Białostockiej, Białystok 2004.
2. Zeliaś A., Pawełek B., Wanat S., *Prognozowanie ekonomiczne: teoria, przykłady, zadania*, Wydawnictwo Naukowe PWN, Warszawa 2004.

Literatura uzupełniająca:

1. DeLurgio S.A., *Forecasting principles and applications*, Irwin/McGraw-Hill, Boston 1998.
2. Hilas C.S., Goudos S.K., Sahalos J.N., *Seasonal decomposition and forecasting of telecommunication data: A comparative case study*, *Technological Forecasting & Social Change* 73, 2006, s. 495–509.
3. Radzikowska B. (red.), *Metody prognozowania: zbiór zadań*, Wydawnictwo Akademii Ekonomicznej im. Oskara Langego we Wrocławiu, Wrocław 2000.

3. ANALIZA HARMONICZNA

Analiza harmoniczna zwana jest także analizą spektralną lub analizą Fouriera na cześć francuskiego matematyka i fizyka Jeana Baptiste'a Josepha Fouriera. Podstawą metody jest opublikowane przez Josepha Fouriera w 1822 roku w *Théorie analytique de la chaleur*⁴⁰, twierdzenie, że każdy szereg okresowy można przedstawić za pomocą sumy funkcji trygonometrycznych sinusów i cosinusów⁴¹. Historycznie podstawy tego twierdzenia sięgają jednak czasów starożytnych. Uproszczonej analizie harmonicznej używali już Babilończycy do prognozowania zjawisk astronomicznych⁴². Współcześnie także wykorzystywana jest, na przykład do modelowania promieniowania słonecznego⁴³. Obecnie szeregi Fouriera mają jednak dużo szersze zastosowanie. Wykorzystywane są praktycznie w każdej dziedzinie, w której bada się częstotliwości wahań, począwszy od analizy drgań do przetwarzania obrazów. Szeregi Fouriera pomagają naukowcom określić skład chemiczny gwiazd lub zrozumieć, jak w układzie oddechowym powstaje mowa⁴⁴.

Podstawową zaletą prognozowania za pomocą analizy harmonicznej jest algorytmizacja modelowania szeregów ze składową sezonową lub cykliczną, stałym poziomem bądź trendem i wahaniami przypadkowymi. W odróżnieniu na przykład od metody Holta-Wintersa czy metody wskaźników analiza harmoniczna nie wymaga znajomości *a priori* składnika okresowego. Podczas opracowywania modelu identyfikuje się szczególnie silne wahania sezonowe. Dodatkowo w wypadku, gdy na podstawie obliczonych wskaźników okaże się, że opracowany model można uprościć poprzez rezygnację z odzwierciedlania w modelu nieistotnych wahań sezonowych, nie trzeba ponownie estymować jego parametrów.

⁴⁰ J.B.J. Fourier, *Théorie analytique de la chaleur* (*The Analytical Theory of Heat*), Firmin Didot, Paris 1822.

⁴¹ H. Dym, H.P. McKean, *Fourier Series and Integrals (Probability and Mathematical Statistics)*, ACADEMIC Press INC 1985, s. 2.

⁴² Ibidem.

⁴³ Y. Zong-Chang, Modeling and forecasting monthly movement of annual average solar insolation based on the least-squares Fourier-model, *Energy Conversion and Management*, 2014, vol. 81, s. 201–210.

⁴⁴ C.A. Pickover, *The Math Book: From Pythagoras to the 57th Dimension, 250 Milestones in the History of Mathematics*, Sterling Publishing Company, Inc. 2009, s. 210.

W modelach otrzymanych za pomocą innych metod usunięcie nawet jednego parametru wymaga zazwyczaj ponownego dopasowywania modelu⁴⁵.

W analizie harmoniczej model jest sumą harmonik. Harmonika to odpowiednio przesunięta i przekształcona sinusoida⁴⁶:

$$y_t = A \sin\left(\frac{2\pi}{n} ft + P\right), \quad (3.1)$$

gdzie:

t – czas ($t = 1, 2, 3, \dots, n$);

y_t – wartość harmoniki w momencie lub okresie t ;

A – amplituda – największe odchylenie wartości harmoniki od średniej (moduł różnicy między najniższym lub najwyższym punktem a poziomem średnim);

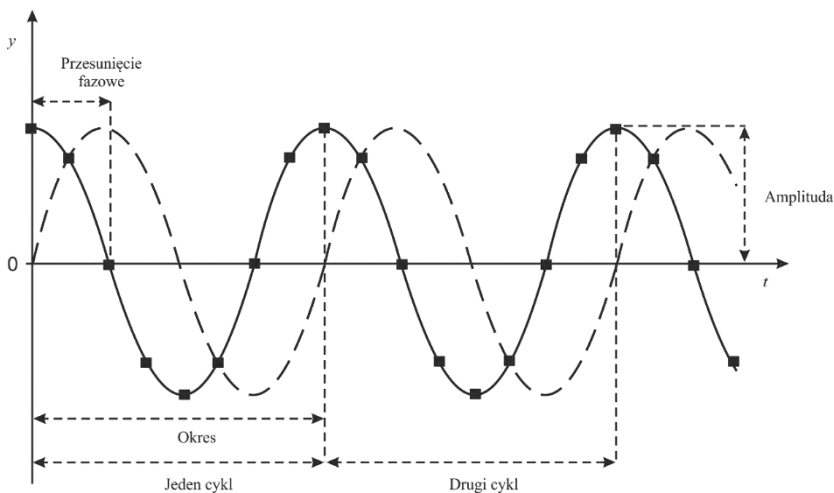
f – częstotliwość, liczba cykli na jednostkę czasu; odwrotność okresu, tj. długości jednego cyklu w jednostkach czasu;

P – przesunięcie fazowe – przesunięcie harmoniki względem początku układu współrzędnych.

Wykres harmoniki wraz z charakteryzującymi ją elementami zilustrowano na rysunku 3.1.

⁴⁵ S.A. DeLurgio, *Forecasting Principles...*, op. cit., s. 241.

⁴⁶ I.N. Bronsztejn, K.A. Siemiendiajew, *Matematyka. Poradnik encyklopedyczny*, Wydawnictwo Naukowe PWN, Warszawa 1995, s. 237.



Rys. 3.1. Harmonika

Źródło: opracowanie własne na podstawie: A. Zeliaś, *Prognozowanie ekonomiczne. Teoria, przykłady, zadania*, Wydawnictwo Naukowe PWN, Warszawa 2004, s. 87.

Korzystając z własności funkcji trygonometrycznych, równanie harmoniki (3.1) można przedstawić w następującej postaci⁴⁷:

$$y_t = A \sin\left(\frac{2\pi}{n} ft + P\right) = a \sin\left(\frac{2\pi}{n} ft\right) + b \cos\left(\frac{2\pi}{n} ft\right), \quad (3.2)$$

gdzie:

$$A = \sqrt{a^2 + b^2},$$

$$\operatorname{tg}(P) = b/a.$$

Dla szeregu składającego się z n obserwacji maksymalna liczba harmonik w modelu analizy harmonicznnej jest równa $q = n/2$, gdy n jest parzyste i $q = (n-1)/2$, gdy n jest nieparzyste. Związane jest to z faktem, iż do zidentyfikowania funkcji sinusoidalnej potrzeba co najmniej dwóch punktów danych: górnej i dolnej wartości ekstremalnej. Pierwsza harmonika w modelu ma okres równy n (częstotliwość $f = 1/n$), druga $n/2$ ($f = 2/n$), trzecia $n/3$ ($f = 3/n$) itd.

⁴⁷ I.N. Bronsztejn, K.A. Siemiendajew, *Matematyka. Poradnik...*, op. cit., s. 237.

Ogólny model szeregu, w którym występują stały poziom, wahania okresowe i przypadkowe, zapisuje się następującym wzorem⁴⁸:

$$\hat{y}_t = a_0 + \sum_{i=1}^q \left[a_i \sin\left(\frac{2\pi}{n} it\right) + b_i \cos\left(\frac{2\pi}{n} it\right) \right] \quad (3.3)$$

$$\text{dla } q = \begin{cases} (n-1)/2, & \text{gdy } n \text{ jest nieparzyste} \\ n/2, & \text{gdy } n \text{ jest parzyste,} \end{cases}$$

gdzie:

a_i, b_i – parametry modelu ($i = 0, 2, 3, \dots, q$).

Za wartość parametru a_0 można przyjąć średnią szeregu:

$$a_0 = \frac{1}{n} \sum_{t=1}^n y_t.$$

Jeżeli w szeregu istnieje trend, a nie stały poziom, wahania okresowe i przypadkowe, można zastosować zmodyfikowany model analizy harmonicznej, w którym parametr a_0 , reprezentujący stały poziom zmiennej prognozowanej, zastąpiony jest funkcją trendu⁴⁹:

$$\hat{y}_t = f(t) + \sum_{i=1}^q \left[a_i \sin\left(\frac{2\pi}{n} it\right) + b_i \cos\left(\frac{2\pi}{n} it\right) \right], \quad (3.4)$$

gdzie:

$f(t)$ – funkcja trendu.

Tab. 3.1. Oszacowania parametrów modelu analizy harmonicznej

n nieparzyste	n parzyste		
$i = 1, \dots, q$	$i = 1, \dots, q-1$	$i = q$	
$a_i = \frac{2}{n} \sum_{t=1}^n y_t \sin\left(\frac{2\pi}{n} it\right)$	(3.5)	$a_q = 0$	(3.6)
$b_i = \frac{2}{n} \sum_{t=1}^n y_t \cos\left(\frac{2\pi}{n} it\right)$	(3.7)	$b_q = \frac{1}{n} \sum_{t=1}^n y_t \cos(\pi t)$	(3.8)

Źródło: opracowanie własne na podstawie: P. Dittmann, *Prognozowanie w przedsiębiorstwie. Metody i ich zastosowanie*, Wolters Kluwers, Kraków 2009, s. 91.

⁴⁸ P. Dittmann, *Prognozowanie w przedsiębiorstwie...*, op. cit., s. 90.

⁴⁹ Ibidem, s. 91.

Wybór i estymacja funkcji trendu zostały opisane w rozdziale 1. Współczynniki poszczególnych harmonik a_i oraz b_i można oszacować na podstawie wzorów przedstawionych w tabeli 3.1. Z tym, że do obliczeń należy wziąć szereg stacjonarny, tzn. z odjętą średnią lub wyeliminowanym trendem*.

Wykorzystanie analizy harmoniczej do prognozowania polega na ekstrapolacji opracowanego modelu na następne momenty czasu:

$$y_{t+1}^* = f(t+1) + \sum_{i=1}^q \left[a_i \sin\left(\frac{2\pi}{n} i(t+1)\right) + b_i \cos\left(\frac{2\pi}{n} i(t+1)\right) \right]. \quad (3.9)$$

Liczba możliwych do określenia harmonik jest tym większa, im dłuższy jest szereg czasowy, na podstawie którego budowany jest model. Jednak opracowując model, nie trzeba zawsze uwzględniać wszystkich możliwych harmonik. Gdy znane są długości cykli wahań okresowych w prognozowanym zjawisku, w modelu można uwzględnić tylko odzwierciedlające je harmoniki. Inny sposób to ujęcie w modelu tych harmonik, których udział w wyjaśnieniu wariancji rozpatrywanej zmiennej prognozowanej jest największy. Części ogólnej zmienności zmiennej prognozowanej Y , które są wyjaśniane przez różne harmoniki, można sumować. Związane jest to z faktem, że żadne dwie harmoniki nie są ze sobą skorelowane, czyli nie uwzględniają tej samej części ogólnej wariancji zmiennej prognozowanej⁵⁰. Trzeba jednak pamiętać, że dla modeli z trendem pod uwagę bierze się wariancję zmiennej po wyeliminowaniu trendu⁵¹.

* W prognozowaniu na podstawie szeregów czasowych wykorzystuje się pojęcie szeregu czasowego słabo stacjonarnego, to znaczy takiego, którego średnia oraz wariancja są skończone i stałe w czasie, a autokowariancja zależy od odstępów czasu pomiędzy obserwacjami, a nie od momentu pomiaru (M. Gruszczyński, M. Podgórska, *Ekonometria*, Szkoła Główna Handlowa, Warszawa 1996, s. 181). Do popularnych metod sprowadzenia szeregu niestacjonarnego do stacjonarnego należą: estymacja trendu deterministycznego i odjęcie od oryginalnego szeregu wartości wynikających z trendu (metoda stosowana m.in. w niniejszym podręczniku) lub policzenie pierwszych różnic dla szeregu czasowego (w wypadku, gdy szereg nie jest stacjonarny ze względu na wariancję). Inne metody obejmują różne sposoby wygładzania szeregu czasowego.

⁵⁰ P. Dittmann, *Prognozowanie w przedsiębiorstwie...*, op. cit., s. 91.

⁵¹ Ibidem.

Wzory pozwalające oszacować część wariancji zmiennej Y objaśnianej przez i -tą harmonikę zamieszczono w tabeli 3.2. Symbol s^2 występujący we wzorach oznacza ocenę wariancji zmiennej Y , tzn.:

$$s^2 = \frac{\sum_{t=1}^n (y_t - \bar{y})^2}{n}; \quad \bar{y} = \frac{1}{n} \sum_{t=1}^n y_t. \quad (3.10)$$

Tab. 3.2. Część wariancji zmiennej Y objaśniana przez i -tą harmonikę

n nieparzyste	n parzyste	
$i = 1 \dots q$	$i = 1 \dots q-1$	$i = q$
$\frac{a_i^2 + b_i^2}{2s^2}$	(3.11)	$\frac{a_q^2 + b_q^2}{s^2}$ (3.12)

Źródło: opracowanie własne na podstawie: P. Dittmann, *Prognozowanie w przedsiębiorstwie. Metody i ich zastosowanie*, Oficyna Ekonomiczna, Kraków 2008, s. 91.

W identyfikacji składowej sezonowej w serii danych można wykorzystać też podstawowe narzędzie analizy spektralnej, tj. periodogram. Przedstawia on amplitudę dla wszystkich możliwych częstotliwości i może być interpretowany jako wartości całkowitej sumy kwadratów:

$\sum_{t=1}^n (y_t - \bar{y})^2$, objaśnianej przez harmoniki o danych częstotliwościach⁵².

Periodogram składa się z $q = n/2$ dla n parzystego lub $q = (n-1)/2$, gdy n jest nieparzyste, wielkości nazywanych intensywnościami przy częstotliwości f_i ⁵³:

$$I(f_i) = n(a_i^2 + b_i^2)/2, \quad (3.13)$$

gdzie: a_i i b_i – parametry modelu analizy harmoniczej.

Suma wartości periodogramu dla wszystkich rozpatrywanych częstotliwości jest równa całkowitej sumie kwadratów⁵⁴:

⁵² S.A. DeLurgio, *Forecasting Principles ...*, op. cit., s. 255.

⁵³ G.M. Box, G.E.P Jenkins., *Analiza szeregów czasowych. Prognozowanie i sterowanie*, Państwowe Wydawnictwo Naukowe, Warszawa 1983, s. 45.

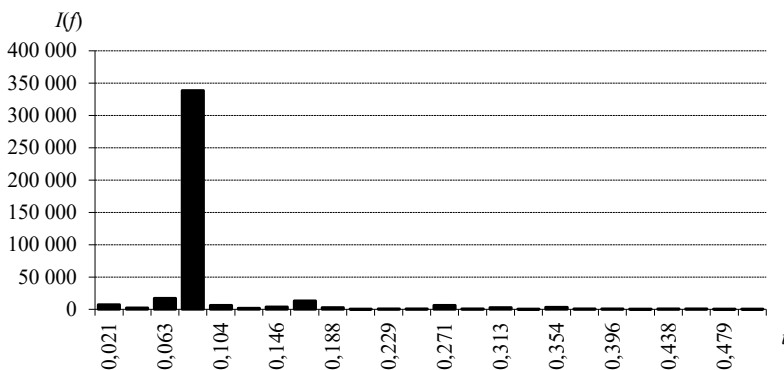
⁵⁴ S.A. DeLurgio, *Forecasting Principles ...*, op. cit., s. 256.

$$\sum_{i=1}^q I(f_i) = \sum_{t=1}^n (y_t - \bar{y})^2, \quad (3.14)$$

gdzie:

$$q = \begin{cases} (n-1)/2, & \text{gdy } n \text{ jest nieparzyste} \\ n/2, & \text{gdy } n \text{ jest parzyste.} \end{cases}$$

Na wykresie funkcję $I(f)$ przedstawić można w zależności od częstotliwości lub okresu. Hipotetyczny periodogram dla szeregu czasowego składającego się z $n = 48$ obserwacji i wykazującego silną sezonowość o okresie 12 (częstotliwości $f = 1/12 = 0,0833$) przedstawiono na rysunku 3.2.

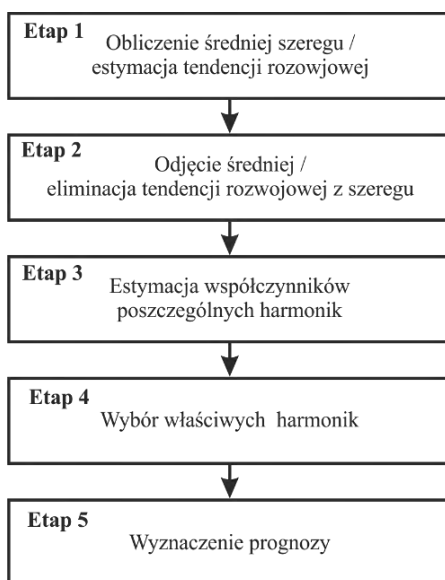


Rys. 3.2. Periodogram

Źródło: opracowanie własne na podstawie: S.A. DeLurgio, *Forecasting Principles and Applications*, Irwin/McGraw-Hill, Boston 1998, s. 258.

Definicja periodogramu zakłada, że rozpatrywane częstotliwości f są związane z długością szeregu n . W praktyce zdarza się, że żadna z branych pod uwagę częstotliwości nie „trafia” dokładnie w tę właściwą. Na przykład jeżeli w szeregu są wahania sezonowe o okresie 4, a do analizy zostanie pobranych $n=14$ obserwacji, to rzeczywista częstotliwość wynosi $f = 0,25$, a wartości periodogramu są wyliczane dla liczb będących wielokrotnością $1/n$, czyli dla 0,071, 0,143, 0,214, 0,286, ..., 0,500. Z tego powodu zdarzyć się może, że właściwa częstotliwość „przecieknie” do sąsiednich na periodogramie. Dodatkowo wartości periodogramu podlegają znacznym wahaniom losowym.

W praktyce podczas analizy zazwyczaj nie jest najważniejsze dokładne zidentyfikowanie częstotliwości podstawowych funkcji sinus czy cosinus. Wystarczające jest określenie częstotliwości o największych gęstościach widmowych, to znaczy obszarów częstotliwości składających się z wielu sąsiednich częstotliwości, które mają największy wkład w ogólną strukturę harmoniczną szeregu. Można to osiągnąć poprzez wygładzanie wartości periodogramu za pomocą średniej ruchomej ważonej.



Rys. 3.3. Etapy opracowania prognozy na podstawie modelu analizy harmonicznego

Źródło: opracowanie własne.

Podsumowując przedstawione rozważania, w procesie opracowywania prognozy na podstawie modelu analizy harmonicznego można wyróżnić pięć elementarnych etapów (rys. 3.3).

Prognozując na podstawie analizy harmonicznego powinno się wziąć pod uwagę pewne ograniczenia tej metody.

Pierwszym problemem jest nadmierne dopasowanie do danych (ang. *overfitting*) związane z liczbą parametrów uwzględnianych w modelu. Włączenie dodatkowego współczynnika do modelu powoduje,

że model staje się bardziej skomplikowany, estymacja jego parametrów wymaga więcej czasu, ale jednocześnie zwiększa się jego dopasowanie do danych historycznych. Modele zbyt rozbudowane mogą uwzględniać obserwacje odstające, czyli jednorazowe zjawiska, które pojawiły się w przeszłości, ale niekoniecznie muszą pojawić się w przyszłości. Bardzo ważny jest wybór i zawarcie w opracowanym modelu analizy harmonicznej tylko właściwych częstotliwości. *Overfitting* jednak nie jest problemem specyficznym tylko dla analizy harmonicznej, należy pamiętać o jego oddziaływaniu na jakość rezultatów procesu prognozowania, także stosując inne metody prognozowania⁵⁵.

Cechą przedstawionej metody jest przypisywanie jednakowych wag wszystkim obserwacjom. W analizie harmonicznej wykorzystuje się metodę najmniejszych kwadratów do obliczenia parametrów harmonik, która uwzględnia wszystkie reszty modelu i są one jednakowo ważne. Dlatego też w sporządzonym modelu każda obserwacja z szeregu czasowego ma ten sam wpływ na obliczoną prognozę. Zazwyczaj ostatnie obserwacje lepiej niż początkowe wskazują przyszłe trendy w danych⁵⁶. Związane jest to z faktem starzenia się informacji i danych prognostycznych.

Inną wadą analizy harmonicznej może okazać się też pewna trudność w odniesieniu otrzymanych parametrów modelu do powszechnie przyjętej lub intuicyjnej interpretacji przyczyn wahań okresowych w szeregu czasowym⁵⁷. Może się zdarzyć, że rzeczywisty okres drgań nie zostaje bezpośrednio uwzględniony w modelu, gdyż nie jest wielokrotnością $1/n$ i wówczas jest odwzorowywany przez sąsiednie częstotliwości.

Przykład 3.1. Prognoza zapotrzebowania mocy KSE metodą analizy harmonicznej

W szeregu czasowym ilustrującym zapotrzebowanie mocy Krajowego Systemu Energetycznego (KSE) występuje bardzo wyraźna sezonowość dobowa, tygodniowa oraz roczna. Na podstawie zmian dobowych zaobserwowanych 2, 9 i 16 września 2013 r. o godz. 6:00, 12:00, 18:00 i 00:00 wyznacz prognozę na 23 września 2013 r.

⁵⁵ Ibidem, s. 255.

⁵⁶ Ibidem.

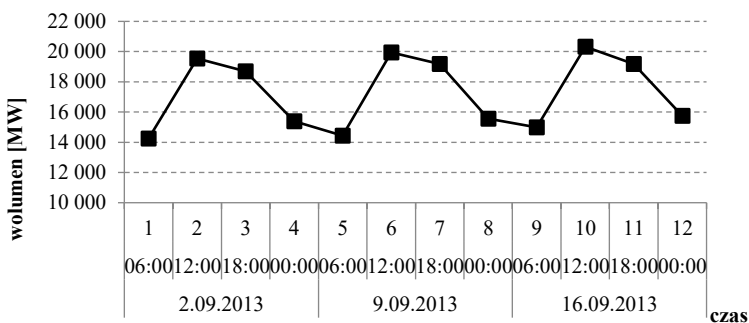
⁵⁷ Ibidem.

Tab. 3.3. Zapotrzebowanie mocy KSE

t	Dzień	Godzina	Wolumen [MW]
			y_t
1	2.09.2013	6:00	14251,7
2		12:00	19545,2
3		18:00	18705,0
4		00:00	15394,6
5	9.09.2013	6:00	14437,0
6		12:00	19938,2
7		18:00	19176,8
8		00:00	15566,4
9	16.09.2013	6:00	14979,3
10		12:00	20320,7
11		18:00	19189,3
12		00:00	15743,7

Źródło: opracowanie własne na podstawie strony internetowej Polskich Sieci Elektroenergetycznych S.A., <http://www.pse-operator.pl/index.php?dzid=77>, data wejścia 29.04.2016.

Analiza danych numerycznych przedstawionych w tabeli 3.3 i ich graficznej reprezentacji na rysunku 3.4 wskazuje, że w szeregu czasowym występują trend oraz wyraźne wahania sezonowe o okresie równym 4 (częstotliwości $f = 0,25$).



Rys. 3.4. Zapotrzebowanie mocy KSE w MW

Źródło: opracowanie własne na podstawie strony internetowej Polskich Sieci Elektroenergetycznych S.A., <http://www.pse-operator.pl/index.php?dzid=77>, data wejścia 29.04.2016.

Opracowanie prognozy zapotrzebowania mocy na kolejny poniedziałek miesiąca (23.09.2014) za pomocą metody analizy harmonicznej wymaga wyodrębnienia i eliminacji tendencji rozwojowej z szeregu (etap 1 i 2), estymacji parametrów poszczególnych harmonik (etap 3) oraz wyboru właściwych harmonik (etap 4). Prognozowana wartość zmiennej będzie sumą wartości wyodrębnionego trendu oraz wybranych harmonik (etap 5).

Etap 1. Estymacja funkcji trendu

Przyjęto liniową postać funkcji trendu: $f^*(t) = 167009 + 87,657t$. Do estymacji parametrów wykorzystano metodę najmniejszych kwadratów zaprezentowaną w rozdziale 1.

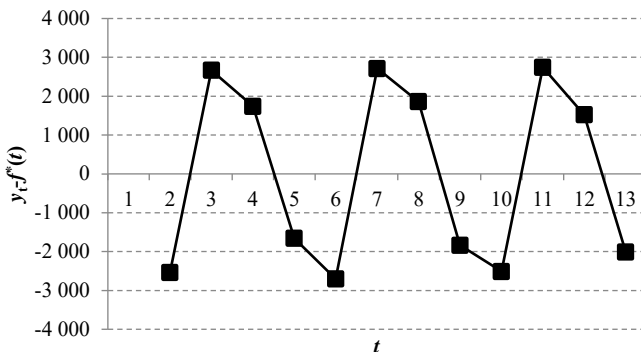
Etap 2. Eliminacja trendu z szeregu

Trend z szeregu wyeliminowano poprzez odjęcie od wartości szeregu wartości trendu, otrzymując nowy szereg, o wartościach oscylujących wokół 0 (tab. 3.4, rys. 3.5). Ten nowy szereg będzie wykorzystany do oszacowania parametrów poszczególnych harmonik modelu.

Tab. 3.4. Wartości zapotrzebowania na moc KSE, trendu oraz szeregu stacjonarnego

t	y_t	Trend $f^*(t)$	$y_t - f^*(t)$
1	14251,7	16788,6	-2536,9
2	19545,2	16876,2	2669,0
3	18705,0	16963,9	1741,1
4	15394,6	17051,5	-1656,9
5	14437,0	17139,2	-2702,2
6	19938,2	17226,8	2711,4
7	19176,8	17314,5	1862,3
8	15566,4	17402,1	-1835,7
9	14979,3	17489,8	-2510,5
10	20320,7	17577,4	2743,3
11	19189,3	17665,1	1524,2
12	15743,7	17752,7	-2009,0

Źródło: obliczenia własne.



Rys. 3.5. Wartości szeregu bez trendu

Źródło: opracowanie własne.

Etap 3. Estymacja parametrów poszczególnych harmonik

W wypadku analizowanego szeregu składającego się z $n = 12$ obserwacji maksymalna liczba harmonik q w modelu jest równa $n/2$, czyli 6. Przy czym pierwsza harmonika będzie miała okres równy n , czyli 12, częstotliwość $f = 0,083$, druga $n/2 = 6$, $f = 0,167$, trzecia $n/3 = 4$, $f = 0,25$, czwarta $n/4 = 3$, $f = 0,333$, piąta $n/5 = 2,4$, $f = 0,417$ i ostatnia $n/6 = 2$, $f = 0,5$.

Zgodnie ze wzorami na oszacowania parametrów równania modelu (patrz: tab. 3.1) obliczono najpierw wartości wyrażenia: $\sin\left(\frac{2\pi}{n}it\right)$

dla $n = 12$, $i = 1, 2, \dots, 5$, $t = 1, \dots, 12$ oraz $\cos\left(\frac{2\pi}{n}it\right)$ dla $t = 1, \dots, 12$.

Wyniki obliczeń zamieszczono w tabelach 3.5 i 3.6.

Tab. 3.5. Wartości funkcji sinus

		$\sin\left(\frac{2\pi}{n}it\right)$				
		i				
		1	2	3	4	5
t	1	0,5	0,866	1	0,866	0,5
	2	0,866	0,866	0	-0,866	-0,866
	3	1	0	-1	0	1
	4	0,866	-0,866	0	0,866	-0,866
	5	0,5	-0,866	1	-0,866	0,5
	6	0	0	0	0	0
	7	-0,5	0,866	-1	0,866	-0,5
	8	-0,866	0,866	0	-0,866	0,866
	9	-1	0	1	0	-1
	10	-0,866	-0,866	0	0,866	0,866
	11	-0,5	-0,866	-1	-0,866	-0,5
	12	0	0	0	0	0

Źródło: obliczenia własne.

Tab. 3.6. Wartości funkcji cosinus

		$\cos\left(\frac{2\pi}{n}it\right)$					$\cos(\pi)$
		i					
		1	2	3	4	5	
t	1	0,866	0,5	0	-0,5	-0,866	-1
	2	0,5	-0,5	-1	-0,5	0,5	1
	3	0	-1	0	1	0	-1
	4	-0,5	-0,5	1	-0,5	-0,5	1
	5	-0,866	0,5	0	-0,5	0,866	-1
	6	-1	1	-1	1	-1	1
	7	-0,866	0,5	0	-0,5	0,866	-1
	8	-0,5	-0,5	1	-0,5	-0,5	1
	9	0	-1	0	1	0	-1
	10	0,5	-0,5	-1	-0,5	0,5	1
	11	0,866	0,5	0	-0,5	-0,866	-1
	12	1	1	1	1	1	1

Źródło: obliczenia własne.

W celu oszacowania parametrów pięciu pierwszych harmonik należy obliczyć sumy iloczynów wartości obserwacji szeregu bez trendu (z tab. 3.4) i wartości odpowiadających im funkcji sinus bądź cosinus (zamieszczone w tab. 3.5 oraz 3.6) zgodnie ze wzorami z tabeli (3.1), to znaczy:

$$a_i = \frac{2}{n} \sum_{t=1}^n (y_t - f^*(t)) \sin\left(\frac{2\pi}{n} it\right) \quad \text{dla } n=12 \text{ oraz } i=1, \dots, 5$$

oraz

$$b_i = \frac{2}{n} \sum_{t=1}^n (y_t - f^*(t)) \cos\left(\frac{2\pi}{n} it\right) \quad \text{dla } n=12 \text{ oraz } i=1, \dots, 5.$$

W wypadku współczynnika a_1 oraz b_1 będzie to odpowiednio:

$$\begin{aligned} a_1 &= \frac{2}{12} \sum_{t=1}^{12} (y_t - f^*(x)) \sin\left(\frac{2\pi}{12} it\right) = \frac{2}{12} [(-2536,9) 0,5 + \\ &+ 2669,0 \cdot 0,866 + 1741,1 \cdot 1 + (-1656,9) \cdot 0,866 + (-2702,2) 0,5 + \\ &+ 2711,4 \cdot 0 + 1862,3 \cdot (-0,5) + (-1835,7) \cdot (-0,866) + \\ &+ (-2510,5) \cdot (-1) + 2743,3 \cdot (-0,866) + 1524,2 \cdot (-0,5) + \\ &+ (-2009,0) \cdot 0] = \frac{2}{12} \cdot 29,3 = 4,9; \end{aligned}$$

$$\begin{aligned} b_1 &= \frac{2}{12} \sum_{t=1}^{12} (y_t - f^*(t)) \cos\left(\frac{2\pi}{12} t\right) = \frac{2}{12} [(-2536,9) 0,866 + \\ &+ 2669,0 \cdot 0,5 + 1741,1 \cdot 0 + (-1656,9) \cdot (-0,5) + \\ &+ (-2702,2) \cdot (-0,866) + 2711,4 \cdot (-1) + 1862,3 \cdot (-0,866) + \\ &+ (-1835,7) \cdot (-0,5) + (-2510,5) 0 + 2743,3 \cdot 0,5 + 1524,2 \cdot 0,866 + \\ &+ (-2009,0) \cdot 1] = \frac{2}{12} \cdot (-417,6) = -69,6. \end{aligned}$$

W wypadku korzystania z arkusza kalkulacyjnego można wykorzystać wbudowane funkcje do obliczenia sumy iloczynów lub samodzielnie zsumować uprzednio wymnożone wartości szeregu bez trendu i funkcji sinus lub cosinus z odpowiadających sobie momentów lub okresów t . Na podstawie wyników cząstkowych obliczeń wartości szeregu czasowego po eliminacji trendu (patrz: tab. 3.4) oraz wartości funkcji sinus (tab. 3.5) obliczono wartość współczynnika a_1 . Wyniki obliczeń zawarto w tabeli 3.7. Podobnie wartości służące do obliczenia wartości b_1 zamieszczono w tabeli 3.8.

Tab. 3.7. Obliczenia sumy iloczynów dla a_1 ($i = 1; n = 12$)

t	$y_t - f^*(t)$	$\sin\left(\frac{2\pi}{n}it\right)$	$(y_t - f^*(t))\sin\left(\frac{2\pi}{n}it\right)$
1	-2536,9	0,5	$-2536,9 \cdot 0,5 = -1268,5$
2	2669,0	0,866	$2669,0 \cdot 0,866 = 2311,4$
3	1741,1	1	$\dots = 1741,1$
4	-1656,9	0,866	-1434,9
5	-2702,2	0,5	-1351,1
6	2711,4	0	0,0
7	1862,3	-0,5	-931,2
8	-1835,7	-0,866	1589,7
9	-2510,5	-1	2510,5
10	2743,3	-0,866	-2375,7
11	1524,2	-0,5	-762,1
12	-2009,0	0	0,0
			29,3
		$\sum_{t=1}^n (y_t - f^*(t))\sin\left(\frac{2\pi}{n}it\right)$	4,9

Źródło: obliczenia własne.

Tab. 3.8. Obliczenia sumy iloczynów dla b_1 ($i = 1; n = 12$)

t	$y_t - f^*(t)$	$\cos\left(\frac{2\pi}{n}it\right)$	$(y_t - f^*(t))\cos\left(\frac{2\pi}{n}it\right)$
1	-2536,9	0,866	$-2536,9 \cdot 0,866 = -2197,0$
2	2669,0	0,5	$2669,0 \cdot 0,5 = 1334,5$
3	1741,1	0	$\dots = 0,0$
4	-1656,9	-0,5	828,5
5	-2702,2	-0,866	2340,1
6	2711,4	-1	-2711,4
7	1862,3	-0,866	-1612,8
8	-1835,7	-0,5	917,9
9	-2510,5	0	0,0
10	2743,3	0,5	1371,7
11	1524,2	0,866	1320,0
12	-2009,0	1	-2009,0
		$\sum_{t=1}^n (y_t - f^*(t))\cos\left(\frac{2\pi}{n}it\right)$	-417,6
		$\frac{1}{6} \sum_{t=1}^n (y_t - f^*(t))\cos\left(\frac{2\pi}{n}it\right)$	-69,6

Źródło: obliczenia własne.

Obliczenia dla współczynników harmoniki drugiej a_2 i b_2 przedstawiono odpowiednio w tabelach 3.9 i 3.10.

Tab. 3.9. Obliczenia sumy iloczynów dla a_2 ($i = 2; n = 12$)

t	$y_t - f^*(t)$	$\sin\left(\frac{2\pi}{n}it\right)$	$(y_t - f^*(t))\sin\left(\frac{2\pi}{n}it\right)$
1	-2536,9	0,866	-2197,0
2	2669,0	0,866	2311,4
3	1741,1	0	0,0
4	-1656,9	-0,866	1434,9
5	-2702,2	-0,866	2340,1
6	2711,4	0	0,0
7	1862,3	0,866	1612,8
8	-1835,7	0,866	-1589,7
9	-2510,5	0	0,0
10	2743,3	-0,866	-2375,7
11	1524,2	-0,866	-1320,0
12	-2009,0	0	0,0
		$\sum_{t=1}^n (y_t - f^*(t))\sin\left(\frac{2\pi}{n}it\right)$	216,8
		$\frac{1}{6} \sum_{t=1}^n (y_t - f^*(t))\sin\left(\frac{2\pi}{n}it\right)$	36,1

Źródło: obliczenia własne.

Tab. 3.10. Obliczenia sumy iloczynów dla b_2 ($i = 2; n = 12$)

t	$y_t - f^*(t)$	$\cos\left(\frac{2\pi}{n}it\right)$	$(y_t - f^*(t))\cos\left(\frac{2\pi}{n}it\right)$
1	-2536,9	0,5	-1268,5
2	2669,0	-0,5	-1334,5
3	1741,1	-1	-1741,1
4	-1656,9	-0,5	828,5
5	-2702,2	0,5	-1351,1
6	2711,4	1	2711,4
7	1862,3	0,5	931,2
8	-1835,7	-0,5	917,9
9	-2510,5	-1	2510,5
10	2743,3	-0,5	-1371,7
11	1524,2	0,5	762,1
12	-2009,0	1	-2009,0
		$\sum_{t=1}^n (y_t - f^*(t))\cos\left(\frac{2\pi}{n}it\right)$	-414,3
		$\frac{1}{6} \sum_{t=1}^n (y_t - f^*(t))\cos\left(\frac{2\pi}{n}it\right)$	-69,1

Źródło: obliczenia własne.

Analogicznie, powtarzając obliczenia dla kolejnych harmonik, tzn. dla $i = 3, \dots, 5$, otrzymuje się oszacowania kolejnych parametrów modelu. Wartość współczynnika a_6 , zgodnie ze wzorem z tabeli 3.1, wynosi 0, natomiast b_6 otrzymuje się jako 1/12 sumy iloczynów szeregu

bez trendu oraz funkcji: $\cos(\pi t)$, tj. $b_6 = \frac{1}{n} \sum_{t=1}^n y_t \cos(\pi t)$;

$$\begin{aligned}
 b_6 &= \frac{1}{12} \sum_{t=1}^{12} (y_t - f^*(t)) \cos(\pi t) = \frac{1}{12} [(-2536,9)(-1) + 2669,0 \cdot 1 + \\
 &+ 1741,1 \cdot (-1) + (-1656,9) \cdot 1 + (-2702,2)(-1) + 2711,4 \cdot 1 + \\
 &+ 1862,3 \cdot (-1) + (-1835,7) \cdot 1 + (-2510,5)(-1) + 2743,3 \cdot 1 + \\
 &+ 1524,2 \cdot (-1) + (-2009,0) \cdot 1] = \frac{1}{12} \cdot 5244,1 = 437,0.
 \end{aligned}$$

Obliczenia dla b_6 przedstawiono w tabeli 3.11.

Tab. 3.11. Obliczenia sumy iloczynów dla b_6 ($i = 6; n = 12$)

t	$y_t - f^*(t)$	$\cos(\pi t)$	$\sum_{t=1}^{12} (y_t - f^*(t)) \cos(\pi t)$
1	-2536,9	-1	2536,9
2	2669,0	1	2669,0
3	1741,1	-1	-1741,1
4	-1656,9	1	-1656,9
5	-2702,2	-1	2702,2
6	2711,4	1	2711,4
7	1862,3	-1	-1862,3
8	-1835,7	1	-1835,7
9	-2510,5	-1	2510,5
10	2743,3	1	2743,3
11	1524,2	-1	-1524,2
12	-2009,0	1	-2009,0
		$\sum_{t=1}^{12} (y_t - f^*(t)) \cos(\pi t)$	5244,1
		$\frac{1}{12} \sum_{t=1}^{12} (y_t - f^*(t)) \cos(\pi t)$	437,0

Źródło: obliczenia własne.

Zestaw wartości wszystkich parametrów modelu zawarto w tabeli 3.12.

Tab. 3.12. Parametry modelu analizy harmonicznej

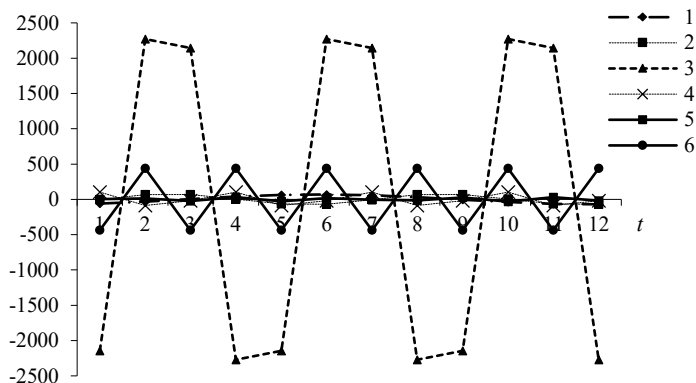
numer harmoniki i	a_i	b_i
1	$a_1 = 4,9$	$b_1 = -69,6$
2	$a_2 = 36,1$	$b_2 = -69,1$
3	$a_3 = -2146,2$	$b_3 = -2270,9$
4	$a_4 = 109,2$	$b_4 = -16,8$
5	$a_5 = -25,3$	$b_5 = -19,7$
6	$a_6 = 0$	$b_6 = 437,0$

Źródło: obliczenia własne.

Pełny model analizy harmoniczej otrzymany w wyniku przeprowadzonych powyżej obliczeń ma postać:

$$\begin{aligned}
 \hat{y}_t = & \underbrace{16700,9 + 87,654t + 4,9\sin\left(\frac{2\pi}{12}t\right) - 69,6\cos\left(\frac{2\pi}{12}t\right)}_{\text{Harmonika 1}} \\
 & + \underbrace{36,1\sin\left(\frac{4\pi}{12}t\right) - 69,1\cos\left(\frac{4\pi}{12}t\right)}_{\text{Harmonika 2}} - \underbrace{2146,2\sin\left(\frac{6\pi}{12}t\right) - 2270,9\cos\left(\frac{6\pi}{12}t\right)}_{\text{Harmonika 3}} \\
 & + \underbrace{109,2\sin\left(\frac{8\pi}{12}t\right) - 16,8\cos\left(\frac{8\pi}{12}t\right)}_{\text{Harmonika 4}} - \underbrace{25,3\sin\left(\frac{10\pi}{12}t\right) - 19,7\cos\left(\frac{10\pi}{12}t\right)}_{\text{Harmonika 5}} \\
 & + \underbrace{437,0\cos\left(\frac{12\pi}{12}t\right)}_{\text{Harmonika 6}}.
 \end{aligned}$$

Rozkłady poszczególnych harmonik tego modelu pokazano na rysunku 3.6.



Rys. 3.6. Rozkład harmonik

Źródło: opracowanie własne.

Etap 4. Oszacowanie udziału poszczególnych harmonik w wyjaśnianiu ogólnej zmienności szeregu

Korzystając z ilustracji rozkładu poszczególnych harmonik modelu (rys. 3.6), można zaobserwować, że największe odchylenia wartości zmiennej od linii trendu powodują wahania o okresie 4 (harmonika 3). Opisu jąca je harmonika ma największą amplitudę. Obliczenia amplitud wszystkich harmonik oraz ich udziału w wyjaśnianiu zmienności szeregu przedstawiono w tabeli 3.13.

Tab. 3.13. Wielkości amplitud oraz udział wariancji uwzględniany przez poszczególne harmoniki

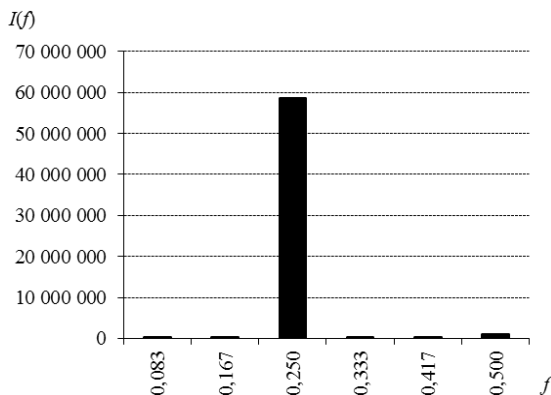
i	f	a_i	b_i	$a_i^2 + b_i^2$	$A_i = \sqrt{a_i^2 + b_i^2}$	$\frac{a_i^2 + b_i^2}{2s^2}$ [%]	$I(f_i) = n(a_i^2 + b_i^2)/2$
1	0,083	4,9	-69,6	4 868,2	69,8	0,05	29 209,0
2	0,167	36,1	-69,1	6 078,0	78,0	0,06	36 468,1
3	0,250	-2 146,2	-2 270,9	9 763 161,3	3 124,6	96,01	5 857 867,5
4	0,333	109,2	-16,8	12 206,9	110,5	0,12	73 241,3
5	0,417	-25,3	-19,7	1 028,2	32,1	0,01	6 169,1
						$\frac{a_6^2 + b_6^2}{s^2}$	
6	0,500	0	437,0	190 969,0	437,0	3,76	1 145 857,7

Źródło: obliczenia własne.

Do obliczeń wykorzystano obciążony estymator wariancji szeregu po wyeliminowaniu trendu:

$$s^2 = \frac{\sum_{t=1}^{12} ((y_t - f^*(t)) - \bar{y})^2}{12} = 5084605,98$$

W tabeli zamieszczono także obliczone wartości periodogramu. Na wykresie periodogramu (rys. 3.7) wyraźniej niż na rysunku z rozkładami poszczególnych harmonik (rys. 3.6) widać, że za przedstawienie sezonowej zmienności analizowanego szeregu odpowiada w zasadzie tylko harmonika 3.



Rys. 3.7. Periodogram

Źródło: opracowanie własne.

Ponieważ harmonika 3 wyjaśnia ponad 96% wariancji analizowanej zmiennej, model można uprościć, uwzględniając w nim tylko tę harmonikę:

$$\hat{y}_t = 167009 + 87,654t - 2146,2 \sin\left(\frac{6\pi}{12}t\right) - 2270,9 \cos\left(\frac{6\pi}{12}t\right).$$

Etap 5. Prognoza

Zakładając, że w okresie prognozowanym utrzyma się zaobserwowana tendencja rozwojowa, a rodzaj i siła wahań sezonowych nie ulegną zmianie, można obliczyć prognozę poprzez ekstrapolację zbudowanego modelu na następne momenty czasu. Na przykład wartość zmiennej 23.09.2013 r. o godz. 6:00, czyli dla $t = 13$, będzie równała się:

$$y_t^* = 167009 + 87,654t - 2146,2 \sin\left(\frac{6\pi}{12}13\right) - 2270,9 \cos\left(\frac{6\pi}{12}13\right) = 15694,2.$$

Prognozy na kolejne wybrane godziny 23.09.2013 r. wynoszą odpowiednio:

$$y_t^* = 167009 + 87,654t - 2146,2\sin\left(\frac{6\pi}{12}14\right) -$$

$$- 2270,9\cos\left(\frac{6\pi}{12}14\right) = 20198,9;$$

$$y_t^* = 167009 + 87,654t - 2146,2\sin\left(\frac{6\pi}{12}15\right) -$$

$$- 2270,9\cos\left(\frac{6\pi}{12}15\right) = 20161,9;$$

$$y_t^* = 167009 + 87,654t - 2146,2\sin\left(\frac{6\pi}{12}16\right) -$$

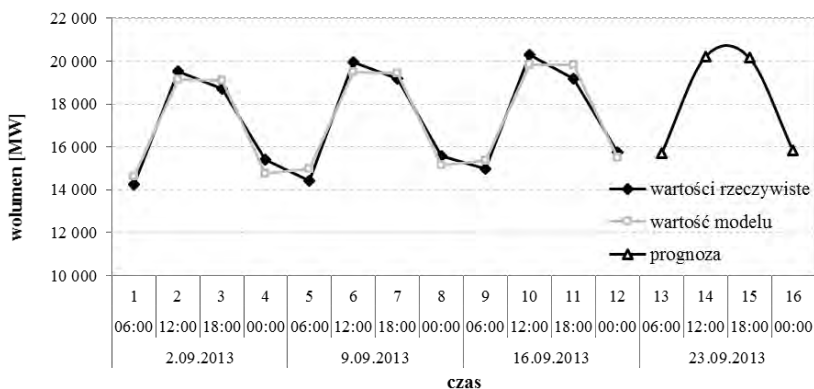
$$- 2270,9\cos\left(\frac{6\pi}{12}16\right) = 15832,5.$$

Wyniki obliczeń zebrano w tabeli 3.14. Wykres wartości rzeczywistych, wartości modelu oraz prognozy zamieszczono na rysunku 3.8.

Tab. 3.14. Prognoza wartości zmiennej

<i>t</i>	Dzień	Godzina	Wolumen [MW]	Model i prognoza	Model i prognoza (wszystkie harmoniki)
			y_t	$y_{t_{ij}}^*$	$y_{t_{ij}}^*$
1	2.09.2013	6:00	14251,7	14642,4	14251,6
2		12:00	19545,2	19147,1	19545,2
3		18:00	18705,0	19110,1	18705,0
4		00:00	15394,6	14780,6	15394,6
5	9.09.2013	6:00	14437,0	14993,0	14437,0
6		12:00	19938,2	19497,7	19938,2
7		18:00	19176,8	19460,7	19176,8
8		00:00	15566,4	15131,2	15566,4
9	16.09.2013	6:00	14979,3	15343,6	14979,3
10		12:00	20320,7	19848,3	20320,7
11		18:00	19189,3	19811,3	19189,3
12		00:00	15743,7	15481,9	15743,7
13	23.09.2013	6:00		15694,2	15303,5
14		12:00		20198,9	20597,0
15		18:00		20161,9	19756,8
16		00:00		15832,5	16446,5

Źródło: obliczenia własne.



Rys. 3.8. Wartości rzeczywiste, modelu i prognozy

Źródło: opracowanie własne.

Obliczone wartości zmiennej na podstawie modelu uwzględniającego wszystkie 6 harmonik (kolumna: model i prognoza wszystkie harmoniki) zaprezentowano w tabeli 3.13. Warto zwrócić uwagę na osiągnięte wówczas prawie idealne odwzorowanie szeregu czasowego – różnice między wartościami rzeczywistymi i wartościami modelu nie przekraczają 0,1 MW. Jednak, o czym była mowa powyżej, nie musi to oznaczać dokładniejszej prognozy na następny tydzień. Korzystanie z uproszczonego do jednej harmoniki modelu daje różnice maksymalnie rzędu 622 MW.

Kluczowe zagadnienia:

analiza harmoniczna, analiza Fouriera, harmonika, periodogram

Literatura podstawowa:

1. Dittmann P., *Prognozowanie w przedsiębiorstwie. Metody i ich zastosowanie*, Wolters Kluwer, Kraków 2009.

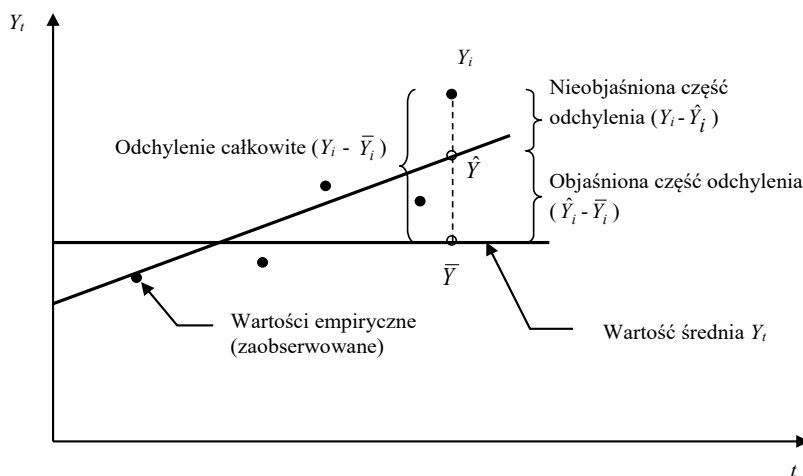
Literatura uzupełniająca:

1. Box G.M., Jenkins G.E.P., *Analiza szeregów czasowych. Prognozowanie i sterowanie*, Państwowe Wydawnictwo Naukowe, Warszawa 1983.
2. Bronsztejn I.N., Siemiendiajew K.A., *Matematyka. Poradnik encyklopedyczny*, Wydawnictwo Naukowe PWN, Warszawa 1995.
3. DeLurgio S.A., *Forecasting Principles and Applications*, Irwin/McGraw-Hill, Boston 1998.
4. Dym H., McKean H.P., *Fourier Series and Integrals (Probability and Mathematical Statistics)*, ACADEMIC PressINC 1985.
5. Fourier J.B.J., *Théorie analytique de la chaleur (The Analytical Theory of Heat)*, Firmin Didot, Paris 1822.
6. Gruszczyński M., Podgórska M., *Ekomonetria*, Szkoła Główna Handlowa, Warszawa 1996.
7. Pickover C.A., *The Math Book: From Pythagoras to the 57th Dimension, 250 Milestones in the History of Mathematics*, Sterling Publishing Company, Inc. 2009.
8. Zeliaś A, Pawełek B. Wanat S., *Prognozowanie ekonomiczne. Teoria, przykłady zadania*, Wydawnictwo Naukowe PWN, Warszawa 2004.
9. Zong-Chang Y., Modeling and forecasting monthly movement of annual average solar insolation based on the least-squares Fourier-model, *Energy Conversion and Management*, vol. 81, May 2014, s. 201–210.

4. OCENA MODELU ORAZ DOPUSZCZALNOŚCI PROGNOZ

4.1. Ocena dopasowania modelu

Po oszacowaniu parametrów potencjalnych postaci modeli analitycznych konieczna jest ocena ich dopasowania do danych empirycznych. Ocena ta ma na celu sprawdzenie, czy skonstruowany model przy założeniach przyjętych przez prognozę (patrz: rozdz. 1) spełnia stawiane wymagania jakościowe. Kluczowym miernikiem informującym o jakości doboru modelu do wartości rzeczywistych jest **współczynnik determinacji R^2** . Punktem wyjścia do zbudowania tej miary jest badanie sumy kwadratów odchyłeń poszczególnych obserwacji od średniej. Na rysunku 4.1 przedstawiono graficzną ilustrację idei oceny dopasowania modelu do wartości empirycznych prognozowanej zmiennej z wykorzystaniem R^2 .



Rys. 4.1. Graficzna ilustracja oceny dopasowania modelu do wartości empirycznych

Źródło: opracowanie własne.

Ocena dopasowania modelu do wartości rzeczywistych z wykorzystaniem współczynnika determinacji polega na sprawdzeniu, jaką część całkowitej zmienności zmiennej prognozowanej Y (odchylenia całkowitego) stanowi jej objaśniona część zmienności (objaśniona część odchylenia). Stwierdzenie to można przedstawić w postaci podanego poniżej prostego równania:

$$\begin{array}{ccccc} (Y_i - \bar{Y}) & = & (Y_i - \hat{Y}_i) & + & (\hat{Y}_i - \bar{Y}) \\ \uparrow & & \uparrow & & \uparrow \\ \text{odchylenie} & & \text{odchylenie} & & \text{odchylenie} \\ \text{całkowite} & & \text{nieobjaśnione} & & \text{objaśnione} \end{array}$$

Jeżeli przyjmiemy, że miarą reprezentującą ww. odchylenie jest wariancja, to powyższe równanie można sprowadzić do następującej postaci:

$$\begin{array}{ccccc} \sum (Y_i - \bar{Y})^2 & = & \sum (Y_i - \hat{Y}_i)^2 & + & \sum (\hat{Y}_i - \bar{Y})^2 \\ \uparrow & & \uparrow & & \uparrow \\ \text{wariancja} & & \text{wariancja} & & \text{wariancja} \\ \text{całkowita} & & \text{nieobjaśniona} & & \text{objaśniona} \end{array}$$

Korzystając z powyższego równania, można sformułować podstawą zależności prezentującą ideę współczynnika determinacji R^2 :

$$R^2 = \frac{\sum_{t=1}^n (\hat{y}_t - \bar{y})^2}{\sum_{t=1}^n (y_t - \bar{y})^2}, \quad (4.1)$$

gdzie:

y_t – wartość zmiennej prognozowanej w okresie/momentcie t ,

\hat{y}_t – teoretyczna wartość zmiennej prognozowanej w okresie / momencie t wynikająca z modelu,

\bar{y} – średnia wartość zmiennej prognozowanej Y w szeregu czasowym o długości n .

Współczynnik determinacji R^2 jest opisową miarą dopasowania liniowego modelu do danych rzeczywistych i przyjmuje wartości z przedziału $\langle 0; 1 \rangle$:

$$R^2 \in \langle 0; 1 \rangle. \quad (4.2)$$

W sytuacji, gdy model objaśni całkowite odchylenie zmiennej, współczynnik R^2 przyjmie wartość równą 1. W celu ułatwienia interpretacji wartości R^2 uzyskaną wartość mnoży się razy 100% i interpretuje jako procentowy udział całkowitej zmienności zmiennej, który został objaśniony przez model. Generalizując, można stwierdzić, że współczynnik determinacji wskazuje, w ilu procentach zbudowany przez prognozę model objaśnia zmienną prognozowaną. Zatem im wyższą wartość przyjmuje R^2 tym oceniany model jest lepiej dopasowany do wartości rzeczywistych⁵⁸. Wykorzystanie symbolu R^2 do oznaczenia współczynnika determinacji wynika z tego, że w modelu liniowym współczynnik ten jest równy kwadratowi współczynnika korelacji r .

Alternatywnym miernikiem oceny jakości modelu i jego dopasowania do wartości rzeczywistych jest **współczynnik interdeterminacji** φ^2 :

$$\varphi^2 = \frac{\sum_{t=1}^n (y_t - \hat{y}_t)^2}{\sum_{t=1}^n (y_t - \bar{y}_t)^2}. \quad (4.3)$$

Współczynnik interdeterminacji φ^2 , określane także w literaturze jako **współczynnik zbieżności**, wskazuje, jaką część odchylenia całkowitego stanowi odchylenie nieobjaśnione. Analogicznie, w celu ułatwienia interpretacji wartości φ^2 , uzyskaną wartość mnoży się razy 100% i interpretuje jako procentowy udział całkowitej zmienności zmiennej, który nie został objaśniony przez model. Suma wartości współczynników determinacji oraz interdeterminacji jest równa jedności:

$$R^2 + \varphi^2 = 1. \quad (4.4)$$

Współczynnik determinacji można wykorzystać do oceny dopasowania modeli nieliniowych do wartości empirycznych pod warunkiem ich sprowadzalności do postaci liniowej. Wyznaczany jest wówczas dla postaci zlinearyzowanej. Może on jednak przyjąć wartość spoza ww. przedziału, co uniemożliwia jego interpretację, a ponadto porównanie

⁵⁸ D.T. Larose, *Metody i modele eksploracji danych*, Wydawnictwo Naukowe PWN, Warszawa 2016, s. 42-46.

potencjalnych modeli. W takich sytuacjach należy wykorzystać **skorygowany współczynnik determinacji \tilde{R}^2** :

$$\tilde{R}^2 = 1 - \frac{n-1}{n-m-1} (1 - R^2), \quad (4.5)$$

gdzie:

n – liczba obserwacji w szeregu czasowym,

m – liczba zmiennych objaśniających (bez zmiennej stojącej przy wyrazie wolnym).

Stosując współczynnik determinacji do oceny dopasowania modeli do wartości empirycznych, należy zwrócić szczególną uwagę w trakcie jego dekompozycji na występowanie w analizowanym szeregu czasowym tzw. obserwacji nietypowych, punktów odstających (oddalonych). Wówczas gdy model istotnie zmienia się w zależności od ich występowania lub nieobecności w szeregu czasowym, obserwacja ta może być tzw. obserwacją wpływową. Zwykle cechą charakteryzującą taką obserwację jest duża wartość reszty z modelu. Niewyeliminowanie takich obserwacji na etapie gromadzenia i analizy danych⁵⁹ lub nieuwzględnienie ich wpływu na dobór modelu do wartości rzeczywistych może doprowadzić do sformułowania mylnych wniosków i rekomendacji.⁶⁰

Do oceny dopasowania modelu do wartości rzeczywistych można wykorzystać także **odchylenie standardowe reszt modelu s** :

$$s = \sqrt{\frac{1}{n-m-1} \cdot \sum_{t=1}^n (y_t - \hat{y}_t)^2}, \quad (4.6)$$

gdzie:

y_t – wartość zmiennej prognozowanej w okresie/momencie t ,

\hat{y}_t – teoretyczna wartość zmiennej prognozowanej w okresie/momencie t ,

n – liczba obserwacji w szeregu czasowym,

m – liczba zmiennych objaśniających (bez zmiennej stojącej przy wyrazie wolnym).

⁵⁹ J. Nazarko (red.), *Prognozowanie w zarządzaniu przedsiębiorstwem*. Cz. II., op. cit.

⁶⁰ A. Orwat, *Przykład zastosowania metody odpornej w modelowaniu finansowych szeregów czasowych*. Zeszyty Naukowe Szkoły Głównej Gospodarstwa Wiejskiego. Ekonomika i Organizacja Gospodarki Żywnościowej 60, 2006, s. 279-288.

Wartość odchylenia standardowego reszt modelu informuje o przeciętnych odchyleniach wartości rzeczywistych prognozowanej zmiennej od wartości teoretycznej uzyskanej z wykorzystaniem zbudowanego przez prognozę modelu. Miara ta jest wyrażona w tych samych jednostkach co zmienna prognozowana. Im mniejszą wartość przyjmuje odchylenie standardowe reszt modelu, tym jest on lepiej dopasowany do wartości empirycznych, a co za tym idzie – jest on wyższej jakości.

Komplementarną miarą oceny dopasowania modelu do wartości rzeczywistych jest **współczynnik wyrazistości**⁶¹:

$$w = \frac{s}{\bar{y}} \cdot 100. \quad (4.7)$$

Współczynnik ten ułatwia interpretację wartości odchylenia standardowego reszt modelu, informując, jaką część średniej wartości zmiennej prognozowanej stanowi to odchylenie. Podobnie jak poprzednio, im mniejsza jest wartość współczynnika wyrazistości, tym model lepiej odzwierciedla zmiany zmiennej prognozowanej.

Jeżeli wyniki oceny dopasowania modelu nie spełniają stawianych mu przez odbiorcę prognozy wymagań jakościowych, należy zbudować nowy model bądź rekomendować model alternatywny i poddać go dalszej weryfikacji. Na ocenę końcową modelu mają wpływ oceny istotności parametrów modelu oraz oceny estymatorów (patrz rozdz. 1).

4.2. Ocena dopuszczalności prognoz

Realizacja procesów decyzyjnych przedsiębiorstwa wymaga opracowania dokładnych prognoz w celu przygotowania i realizacji na ich podstawie właściwych działań w zakresie planowania produkcji, sprzedaży, budżetowania, wprowadzania nowych produktów, planowania promocji, harmonogramowania zasobów itp. Rola prognozy w procesie decyzyjnym i jej wpływ na jego efektywność powoduje, że poszukuje się metod i modeli dostarczających prognoz o wysokiej trafności, spełniających oczekiwania decydentów. Z uwagi na dostępność i koszt pozyskiwania danych prognostycznych większość przedsiębiorstw, budując prognozy,

⁶¹ M. Cieślak (red.), *Prognozowanie gospodarcze...*, op. cit., s. 47.

bazuje na posiadanych danych z przeszłości⁶². Dzięki analizie danych historycznych oraz gromadzeniu i ocenie przesłanek prognostycznych przedsiębiorstwa dążą do zdobycia informacji o przyszłości, wykorzystując do tego celu najczęściej metody mało kosztowne, proste w stosowaniu, łatwe pod względem interpretacyjnym, a przede wszystkim zapewniające zadowalające wyniki procesu prognozowania⁶³. W tym kontekście ważny jest dobór właściwych i obiektywnych miar oceny jakości prognoz, w tym jej trafności i dopuszczalności.

Modele analityczne należą do grupy stosunkowo prostych modeli prognostycznych. Jednym z najtrudniejszych etapów w procesie ich budowy, a następnie predykcji, jest oszacowanie ich parametrów tak, aby funkcja trendu wybrana przez prognozę „najlepiej” odzwierciedlała badaną zmienną (patrz: rozdz. 1). Modele analityczne stosowane są w przypadku, gdy prognozowane zjawisko można przedstawić w postaci modelu trendu na podstawie danych z przeszłości, które charakteryzowały się regularnymi zmianami. Wykorzystanie modeli analitycznych do budowy prognozy uzasadnione jest, jak już podkreślono wcześniej (patrz rozdz. 1), wówczas, gdy spełniona jest zasada *ceteris paribus*. Oznacza ona przyjęcie założenia o niezmienności uwarunkowań i czynników wpływających na badane zjawisko ekonomiczne, czyli stabilności uwarunkowań egzogenicznych czynników kształtujących zmienną prognozowaną. W tej sytuacji zakłada się, że zarówno postać modelu oraz wartości ocen jego parametrów nie ulegną zmianie w horyzoncie czasowym, na jaki sporządzana jest dana prognoza. Innymi słowy, oznacza to, że kierunek zmian zmiennej powinien zostać taki sam (wzrost lub spadek) oraz że postać analityczna modelu (wybrana funkcja trendu) nie zmieni się w czasie. Warunkiem koniecznym jest także akceptacja założenia o stabilności rozkładu składnika losowego⁶⁴. Oznacza to przyjęcie przez prognozę *pasywnej postawy prognostycznej* i wykorzystanie zazwyczaj reguły podstawowej (patrz: *Prognozowanie w zarządzaniu przedsiębiorstwem*. Cz. I, rozdz. 3, s. 27)⁶⁵. Przenoszenie (ekstrapolacja)

⁶² J. Nazarko (red.), *Prognozowanie w zarządzaniu przedsiębiorstwem*. Cz. II., op. cit., s. 5-18.

⁶³ J. Nazarko (red.), *Prognozowanie w zarządzaniu przedsiębiorstwem*. Cz. I., op. cit., s. 27-37.

⁶⁴ P. Dittmann, *Metody prognozowania sprzedaży...*, op. cit., s. 70-71.

⁶⁵ J. Nazarko (red.), *Prognozowanie w zarządzaniu przedsiębiorstwem*. Cz. I., op. cit., s. 27-34.

zaobserwowanych tendencji (prawidłowości) z przeszłości w przyszłość jest uzasadnione wówczas, gdy zostaną spełnione omówione powyżej założenia. Ich spełnienie jest bardziej prawdopodobne w przypadku nieodległej przyszłości (prognozy krótkoterminowe) aniżeli okresów bardzo odległych (prognozy długoterminowe).

Akceptacja przedstawionych powyżej założeń umożliwia prawidłową ocenę błędu prognozy *ex ante*. Można go ogólnie zdefiniować jako spodziewane w przyszłości odchylenie wartości rzeczywistej od wartości prognozy.

W celu omówienia zasad oceny błędów prognozy *ex ante* przyjęto następujące ogólne założenia⁶⁶:

- przedmiotem analizy jest model o postaci $y_t = \beta_0 + \beta_1 x_t + \xi_t$, przy czym składnik losowy ma rozkład normalny ($E(\xi) = 0$, $V(\xi) < \infty$);
- w celu wyznaczenia prognozy na okres $t+n$ dokonano estymacji parametrów powyższego modelu, uzyskując $\hat{y}_{t+n} = \beta_0 + \beta_1 \cdot x_{t+n}$, przy czym x_t jest zmienną objaśniającą w modelu⁶⁷ (w przypadku nieliniowej postaci funkcji trendu możliwe było jego sprowadzenie, transformacja do liniowej postaci modelu).

Przyjęcie ww. założeń pozwala na budowę prognozy poprzez ekstrapolację funkcji trendu (patrz rozdz. 1) i ocenę jej jakości za pomocą **błędu prognozy *ex ante***. Przez taki błąd prognozy należy zatem rozumieć różnicę pomiędzy wartością rzeczywistą y_{t+n} a jej prognozą \hat{y}_{t+n} . Przy czym błąd prognozy *ex ante* może być jedynie oszacowany, ponieważ w momencie/okresie wyznaczania prognozy nie jest znana jej rzeczywista realizacja y_t^* (w chwili konstruowania prognozy zmiennej nie są znane jej wartości rzeczywiste), zatem prognoza jest traktowana jako jedna z wielu realizacji zmiennej losowej Y_t (dla $t > n$).

Do oceny rozproszenia możliwych prognoz wokół wartości możliwych realizacji zmiennej w okresie/momentcie $t > n$ wykorzystuje się wariancję prognozy⁶⁸:

⁶⁶ A. Welfe, *Ekonometria*, Wydawnictwo PWE, Warszawa 1995, s. 177-183.

⁶⁷ Założenia te znajdują swoje odzwierciedlenie w modelach przedstawionych w rozdziale 1, w których parametry modelu oznaczono odpowiednio jako β_0 , β_1 , natomiast zmienną objaśniającą jako t .

⁶⁸ M. Cieślak (red.), *Prognozowanie gospodarcze...*, op. cit., s. 54-57.

$$V_t^2 = E\left(Y_t - Y_t^*\right)^2, t > n, \quad (4.8)$$

gdzie:

Y_t – zmienna losowa reprezentująca możliwe realizacje zmiennej prognozowanej Y w okresie/momentcie $t > n$,

Y_t^* – zmienna losowa reprezentująca możliwe prognozy zmiennej prognozowanej Y na okres/moment $t > n$.

Zgodnie z przyjętymi powyżej założeniami zależność opisującą wariancję prognozy można zapisać w postaci następującego równania⁶⁹:

$$V_t^2 = E\left[\left(\sum_{i=0}^m \beta_i t_{it}^* + \xi_t\right) - \sum_{i=0}^m \beta_i t_{it}^*\right]^2, t > n. \quad (4.9)$$

Uwzględniając powyższe założenia oraz założenia dotyczące klasycznego modelu regresji liniowej, przy założeniu normalności rozkładu reszt, można oszacować wariancję prognozy:

$$V_T^2 = \sum_{i=0}^m t_{iT}^*{}^2 D^2(\beta_i) + 2 \sum_{i=0}^{m-1} \sum_{j>1} t_{iT}^* t_{jT}^* \text{cov}(\beta_i, \beta_j) + s^2. \quad (4.10)$$

Przy czym:

$$s = \sqrt{\frac{1}{n-m-1} \sum_{t=1}^n (y_t - \hat{y}_t)^2}, \quad (4.11)$$

gdzie:

t_{iT}^*, t_{jT}^* – wartość i -tej (j -tej) zmiennej objaśniającej modelu na okres/moment T ,

$D^2(\beta_i)$ – ocena wariancji parametru β_i ,

$\text{cov}(\beta_i, \beta_j)$ – kowariancja oceny parametrów β_i oraz β_j ,

s^2 – ocena wariancji odchyłeń reszt z modelu,

s – odchylenie standardowe składnika resztowego,

m – liczba zmiennych objaśniających modelu.

⁶⁹ W celu zachowania spójności symboli w zależnościach związanych z oceną dopuszczalności prognoz wprowadzono zastosowane w rozdziale 1 oznaczenia zmiennej objaśniającej oraz oceny parametrów modelu. W literaturze przedmiotu stosowane są do ich oznaczenia odpowiednio x_t oraz a_i, a_j .

Analizując zależność opisującą wariancję prognozy w kontekście oceny dopuszczalności prognozy, można stwierdzić, że:

- wariancja błędu prognozy maleje wraz ze wzrostem liczebności próby (przy dużej liczbie obserwacji – wykorzystanych danych historycznych – wzrasta dokładność oszacowania parametrów modelu);
- wariancja błędu prognozy rośnie wraz ze wzrostem rozpiętości pomiędzy wartościami zmiennej objaśniającej a jej wartością średnią w próbie.

Korzystając z zależności 4.10, można zdefiniować zależność pozwalającą na ocenę **bezwzględnego błędu prognozy *ex ante*** w momencie/okresie $t > n$. Jest to pierwiastek kwadratowy z oszacowania wariancji prognozy.

$$V_t = \sqrt{V_t^2}, t > n. \quad (4.12)$$

Wartość **bezwzględnego błędu prognozy *ex ante*** informuje o przeciętnych oczekiwanych odchyleniach wartości rzeczywistej od prognozy w momencie $t > n$, czyli wskazuje, jakich można spodziewać się przeciętnych wahań zmiennej prognozowanej wokół jej wartości oczekiwanej. Przyjmuje się, że prognoza będzie tym dokładniejsza, im mniejsza jest wartość oszacowanego błędu. Jest on wyrażony w tych samych jednostkach miary co zmienna prognozowana. Błąd ten może być wykorzystywany przez prognostę do wyboru spośród kilku testowanych modeli tego, który ma najwyższą wartość prognostyczną, co oznacza, że może dostarczyć najlepszej pod względem jakościowym prognozy (wartość prognostyczna tworzona jest łącznie przez model i regułę prognozowania).

W celu odniesienia wielkości błędu do poziomu prognozy stosowane są **względne wartości błędu prognozy *ex ante***. Jest to podstawowa i łatwo interpretowalna miara oceny dopuszczalności prognozy:

$$\eta_t = \frac{V_t}{y_t} \cdot 100, t > n, \quad (4.13)$$

Zgodnie z zależnością 4.13 **względny błąd prognozy *ex ante*** jest równy ilorazowi bezwzględnego błędu prognozy *ex ante* i wartości prognozy wyznaczonej na okres/moment t . Względny błąd prognozy *ex ante* informuje, jaką procentową część wielkości prognozy stanowi przeciętny

błąd predykcji. Błąd ten jest wykorzystywany przy ocenie tzw. **dopuszczalności prognozy**.

Dopuszczalność prognozy jest ustalana przez zleceniodawcę na etapie formułowania zadania prognostycznego. Prognoza uznawana jest za dopuszczalną, gdy jest obdarzona przez jej odbiorcę takim poziomem zaufania, jaki jest niezbędny do realizacji celu, dla którego została przygotowana. Ocena dopuszczalności jest przeprowadzana w tym samym okresie, w którym sporządzana jest prognoza⁷⁰. Oczywiście im błędy *ex ante* są mniejsze, tym ocena dopuszczalności jest wyższa.

Kryterium dopuszczalności prognozy definiowane jest w postaci ograniczenia nakładanego na względny błąd prognozy *ex ante*. Wyróżnia się dwa rodzaje kryteriów dopuszczalności, tj. subiektywne i obiektywne. Kryteria te mogą być definiowane na etapie formułowania zadania prognostycznego głównie przez odbiorcę prognozy, który podaje tzw. wartość progową na przykład względnego błędu prognozy. Mówi się wówczas o tzw. subiektywnym kryterium dopuszczalności. Wartość kryterium zależy od istniejących uwarunkowań i przeznaczenia prognozy przejawiającego się określonym, wymaganym poziomem dokładności prognozy (celem prognozy, przeznaczeniem). Kryterium dopuszczalności może się więc różnie kształtować dla różnych zmiennych prognozowanych⁷¹. Prognozę można zatem uznać za dopuszczalną w przypadku, gdy spełniona jest nierówność:

$$V_t \leq V_t^* \quad (4.14)$$

lub

$$\eta_t \leq \eta_t^*, \quad (4.15)$$

gdzie:

V_t^* – wartość dopuszczalnego bezwzględnego błędu prognozy sporządzonej na okres/moment t zdefiniowanego przez jej odbiorcę/decydenta;

η_t^* – wartość dopuszczalnego względnego błędu prognozy sporządzonej na okres/moment t zdefiniowanego przez jej odbiorcę/decydenta.

⁷⁰ M. Cieślak (red.), *Prognozowanie gospodarcze...*, op. cit., s. 51.

⁷¹ A. Zeliaś, *Teoria prognozy*, PWE, Warszawa 1997, s. 71.

Drugim rodzajem kryterium dopuszczalności jest tzw. kryterium obiektywne, zgodnie z którym prognoza zostaje sklasyfikowana do określonej kategorii jakościowej na podstawie arbitralnie przyjętych wartości względnego błędu prognoz *ex ante*. Wytyczne wraz z arbitralnie przyjętymi wartościami progowymi wykorzystywanymi do oceny jakości prognoz według obiektywnego kryterium ich dopuszczalności przedstawiono w tabeli 4.1.

Tab. 4.1. Klasyfikacja jakościowa prognoz na podstawie obiektywnego kryterium dopuszczalności

Wartości progowe względnego błędu prognozy	Kategoria jakościowa prognozy
$\eta_t \leq 3\%$	bardzo dokładna
$3\% < \eta_t \leq 5\%$	dokładna
$5\% < \eta_t \leq 10\%$	dopuszczalna
$\eta_t > 10\%$	niedopuszczalna

Źródło: opracowanie własne.

Wartości progowe względnego błędu prognozy, umożliwiające klasyfikację jakościową zbudowanej prognozy, są przyjmowane przez prognozę na etapie formułowania zadania prognostycznego.

Ocena dopuszczalności prognozy, oprócz kryteriów bazujących na błędach prognoz *ex ante*, może być przeprowadzona także z wykorzystaniem:

- prawdopodobieństwa realizacji prognozy;
- przedziału ufności do prognozy;
- oceny eksperckiej uzasadniającej wiarygodność bądź dopuszczalność prognozy;
- błędów prognoz *ex post* (patrz: *Prognozowanie w zarządzaniu przedsiębiorstwem*. Cz. III, rozdz. 3, s. 36)⁷².

Prawdopodobieństwo realizacji prognozy jest alternatywnym sposobem oceny dopuszczalności prognozy. Jeżeli zmienna prognozowana ma charakter losowy, to prawdopodobieństwo realizacji prognozy wyznacza się w następujący sposób:

⁷² J. Nazarko (red.), *Prognozowanie w zarządzaniu przedsiębiorstwem*. Cz. III., op. cit.

– dla zmiennej skokowej:

$$P(Y_t = y_t^*) = \gamma_t, t > n, \quad (4.16)$$

– dla zmiennej ciągłej:

$$P(|Y_t - y_t^*| < \varepsilon) = \gamma_t, t > n, \quad (4.17)$$

gdzie:

γ_t – prawdopodobieństwo przyjęcia przez zmienną Y w czasie $t > n$ wartości y_t^* , tzw. wiarygodność prognozy;

ε – dowolnie mała liczba (np. krotność błędu *ex ante* $\varepsilon = uV_t$).

Prognozę uznaje się za dopuszczalną wówczas, gdy:

$$\gamma_t \leq \gamma_t^*, t > n, \quad (4.18)$$

przy czym γ_t^* jest to wartość progowa prawdopodobieństwa realizacji prognozy ustalana przez jej odbiorcę (powinna być bliska jedności)⁷³. Jest to jednak mało popularny sposób weryfikacji jakości prognozy, gdyż niewiele jest metod⁷⁴, które pozwalają na oszacowanie takiego prawdopodobieństwa γ_t .

Ocena dopuszczalności przeprowadzana jest również na podstawie wielkości przedziału ufności. Prognozę przedziałową i przedział ufności dla prognozy formułuje się, wykorzystując średni błąd prognozy *ex ante* V_t :

$$P\{y_t^* - u \cdot V_t \leq y_t \leq y_t^* + u \cdot V_t\} = p, \quad (4.19)$$

gdzie:

u – współczynnik charakteryzujący wiarygodność prognozy, wynikający z rozkładu zmiennej prognozowanej i długości szeregu czasowego (patrz: *Prognozowanie w zarządzaniu przedsiębiorstwem*. Cz. III, rozdz. 3, s. 46-47);

p – wiarygodność prognozy.

⁷³ M. Cieślak (red.), *Prognozowanie gospodarcze...*, op. cit., s. 54-57.

⁷⁴ S. Ostasiewicz, *Wykorzystanie metod dyskryminacyjnych w prognozowaniu dyskretnym*, Prace Naukowe Akademii Ekonomicznej we Wrocławiu nr 232, seria: Monografie i Opracowania nr 14, 1983.

Interpretując tak skonstruowany przedział ufności, można stwierdzić, że z określonym prawdopodobieństwem (wiarygodnością prognozy) realizacja zmiennej prognozowanej Y (przyszła wartość) w okresie/momentcie t będzie należała do tego przedziału. Szczegółowe zasady budowy prognoz przedziałowych i formułowania przedziału ufności do prognozy zostały omówione w trzeciej części podręcznika⁷⁵.

Ocena ekspercka dopuszczalności prognozy wymaga arbitralnego zdefiniowania kryterium jej uznania za dopuszczalną. Powinna być ona przeprowadzona zgodnie z zasadami obowiązującymi w metodach heurystycznych, np. w metodzie delfickiej. Ocena ekspercka stosowana jest wówczas, gdy przyjmuje się aktywną postawę prognozowania bądź mierniki *ex ante*, czy też *ex post* nie charakteryzują się wystarczająco wysoką wiarygodnością.

W sytuacji, gdy kryterium dopuszczalności (patrz: zależności 4.14 i 4.15; tab. 4.1) nie zostanie spełnione, należy zweryfikować przyjęte w zadaniu prognostycznym założenia, sprawdzić zasadność przyjęcia postawy pasywnej i reguły podstawowej prognozowania i/lub zmodyfikować pierwotny model prognostyczny (np. zweryfikować postać funkcji trendu).

Na podstawie przedstawionych zależności można zdefiniować zależności pozwalające oszacować bezwzględne błędy prognoz *ex ante* dla konkretnych postaci modelu analitycznego (postaci funkcji trendu). **Błąd bezwzględny prognoz *ex ante*** dla podstawowej postaci modelu analitycznego, jakim jest model liniowy, można przedstawić za pomocą zależności⁷⁶:

$$V_t = \sqrt{\frac{(T - \bar{t})^2}{\sum_{t=1}^n (t - \bar{t})^2} + \frac{1}{n} + 1} \cdot s, \quad (4.20)$$

gdzie:

- n – liczba wyrazów szeregu czasowego zmiennej prognozowanej;
- s – odchylenie standardowe reszt z modelu;
- T – okres, dla którego wyznaczana jest prognoza, reprezentowany przez kolejną liczbę naturalną;
- \bar{t} – średnia wartość zmiennej czasowej t , $t = 1, \dots, n$.

⁷⁵ J. Nazarko (red.), *Prognozowanie w zarządzaniu przedsiębiorstwem*. Cz. III., op. cit.

⁷⁶ M. Cieślak (red.), *Prognozowanie gospodarcze...*, op. cit., s. 54.

Wielkość błędu *ex ante* rośnie wraz z wydłużeniem horyzontu czasowego, na jaki sporządzana jest prognoza. Błąd ten rośnie także, gdy rośnie wariancja składnika reszt z modelu (patrz: zależność 4.20).

Na wartość błędu prognoz *ex ante*, a tym samym na jakość prognozy może mieć wpływ kilka istotnych składników związanych z procesem predykcji, m.in.⁷⁷:

- błędy w procesie szacowania parametrów modelu,
- błędnie wybrana postać funkcji trendu,
- pojawienie się zakłóceń postaci modelu,
- zafalszowanie danych źródłowych.

Dokonując ostatecznej oceny prognoz *ex ante*, należy pamiętać o kontekstowości i wielowymiarowości oceny procesu prognozowania. Aspekty te zostały omówione w pierwszej części podręcznika *Prognozowanie w zarządzaniu przedsiębiorstwem*⁷⁸.

W związku z powyższą zależnością, oceniając dopuszczalność prognozy, należy ją oceniać w kontekście czasu, tj. horyzontu czasu, na który prognoza jest przygotowywana. Wartość błędu *ex ante* jest wykorzystywana do określania **maksymalnego horyzontu czasowego**, przy którym prognoza nadal może być uznana za dopuszczalną. Przez **maksymalny horyzont czasowy** należy rozumieć najdalszy moment lub okres należący do przyszłości, dla którego prognoza (będąca wynikiem ekstrapolacji modelu) spełnia kryterium dopuszczalności. Należy podkreślić, że im dalszy jest horyzont czasowy, tym niepewność sądu o przyszłości rośnie, co przekłada się na niższe prawdopodobieństwo zaistnienia przewidywanego stanu zmiennej. Oprócz pojęcia maksymalnego horyzontu pojawia się także pojęcie **żądanego horyzontu prognozy**. Oznacza on moment lub okres należący do przyszłości, dla którego prognoza ma być zgodnie z zaleceniem odbiorcy przygotowana. Jeśli żądany horyzont prognozy jest dłuższy od maksymalnego, prognoza nie może być przygotowana z wykorzystaniem opracowanego modelu⁷⁹.

Przedstawione powyżej zależności pozwalają dokonać jedynie oceny błędu prognozy *ex ante* dla modelu liniowego. Zależności dla mo-

⁷⁷ A. Welfe, *Ekonometria*, Wydawnictwo PWE, Warszawa 1995, s. 187-189.

⁷⁸ J. Nazarko (red.), *Prognozowanie w zarządzaniu przedsiębiorstwem*. Cz. I., op. cit., s. 34-37.

⁷⁹ M. Cieślak (red.), *Prognozowanie gospodarcze...*, op. cit., s. 54.

delu liniowego, po ich odpowiednim przekształceniu, mogą być podstawą oszacowania **błędu prognoz *ex ante* dla modeli nieliniowych** sprowadzalnych (transponowanych) do postaci liniowej. Punktem wyjścia do zdefiniowania zależności pozwalających na oszacowanie tego błędu dla modeli nieliniowych jest określenie zależności między wariancją zmiennej pierwotnej Y (postać nieliniowa modelu) i wariancją zmiennej transponowanej \tilde{Y} . Przy czym funkcja $\tilde{Y} = f(Y)$ jest funkcją nielosową i różniczkowalną⁸⁰.

$$V_t = \sqrt{\frac{\tilde{V}_t^2}{\left(\frac{d\tilde{y}}{dy}\right)^2}}, \quad (4.21)$$

$$\tilde{V}_t^2 = \sum_{i=0}^m t_{it}^{*2} D^2(\beta_i) + 2 \sum_{i=0}^{m-1} \sum_{j>1} t_{iT}^* t_{jT}^* \text{cov}(\beta_i, \beta_j) + S^2, \quad (4.22)$$

gdzie:

- \tilde{V}_T – błąd *ex ante* prognozy na okres t (wyznaczonej dla transformaty);
- \tilde{V}_t^2 – ocena wariancji otrzymana z modelu transponowanego;
- $\frac{d\tilde{y}}{dy}$ – pochodna liczona w punkcie $y_t^* = E(Y_t)$;
- t_{iT}^*, t_{jT}^* – wartość i -tej (j -tej) zmiennej objaśniającej modelu transponowanego do postaci liniowej w okresie T ;
- $D^2(\beta_i)$ – wariancja oceny parametru β_i ;
- $\text{cov}(\beta_i, \beta_j)$ – kowariancja oceny parametrów β_i oraz β_j ;
- S^2 – ocena wariancji odchyłeń losowych;
- t – okres, dla którego wyznaczana jest wartość y na podstawie modelu;
- T – okres, dla którego wyznaczana jest prognoza, reprezentowany przez kolejną liczbę naturalną;
- m – liczba zmiennych objaśniających budowanego modelu bez zmiennej stojącej przy wyrazie wolnym.

⁸⁰ Ibidem.

Zależności umożliwiające ocenę wariancji prognozy zmiennej pierwotnej i oszacowanie błędu prognoz *ex ante* w przypadku modeli nieliniowych przedstawiono w tabeli 4.2.

Wyniki badań dotyczące metod monitorowania jakości prognoz w przedsiębiorstwach, przeprowadzone przez Business Forecast Systems⁸¹, wskazują, że ocena dopuszczalności i trafności prognoz jest najczęściej ustalana na podstawie przyjmowanego poziomu dopuszczalności oraz ocen średniego bezwzględnego błędu procentowego lub ważonego średniego bezwzględnego błędu procentowego.

Należy mieć świadomość, że pomimo posługiwania się adekwatnymi metodami oraz obiektywnymi i kompleksowymi miarami oceny jakości prognozy przy analizie uzyskanych wyników należy zwrócić uwagę na trzy zasadnicze czynniki⁸²:

- założenia uwzględniające zdarzenia i uwarunkowania egzogeniczne (zewnętrzne);
- jakość danych źródłowych;
- specyfikę wykorzystanego modelu, techniki prognozowania.

Jest to szczególnie ważne w przypadku nieosiągnięcia zakładanego poziomu jakości i niespełnienia kryterium dopuszczalności prognozy wskazanego przez jego odbiorcę.

⁸¹ Business Forecast Systems, *Forecast Accuracy. Trends 2009*, document elektroniczny <http://www.forecastpro.com/>, data wejścia 29.04.2016 r.

⁸² K. Featherstone, D. James, *Assessment of NZIER Forecast Accuracy*, *New Zealand Institute of Economic Research*, Research Working Paper 94/9, Wellington 1994.

Tab. 4.2. Oceny *ex ante* średniego błędu prognoz dla wybranych nieliniowych postaci funkcji trendu

Rodzaj trendu	Zapis funkcji trendu	Ocena błędu prognoz <i>ex ante</i>
1	2	3
Trend liniowy	$y_t = \beta_0 + \beta_1 t$	<p>Ocena średniego błędu <i>ex ante</i> prognozy (tzw. błąd bezwzględny prognoz <i>ex ante</i>):</p> $V_T = \sqrt{\frac{(T-\bar{t})^2}{\sum_{t=1}^n (t-\bar{t})^2} + \frac{1}{n} + 1} \cdot s.$ <p>Błąd względny prognoz <i>ex ante</i>:</p> $\eta_T = \frac{V_T}{y_T^*} \cdot 100, \quad T > n,$ <p>gdzie: s – odchylenie standardowe reszt z modelu, n – liczba wartości zmiennej (liczba wyrazów szeregu czasowego), T – okres na jaki sporządzana jest prognoza, y_T^* – wartość prognozy wyznaczonej na okres T.</p>
Trend wykładniczy	$y_t = \beta_0 \beta_1^t, \quad a_0 > 1$	<p>Ocena <i>ex ante</i> średniego błędu prognozy logarytmu zmiennej Y (dla postaci zlinearyzowanej):</p> $\tilde{V}_T = \tilde{S}^2 \sqrt{1 + \frac{1}{n} + \frac{(T-\bar{t})^2}{\sum_{t=1}^n (t-\bar{t})^2}},$ <p>gdzie:</p> $\sum_{t=1}^n (t-\bar{t})^2 = \sum_{t=1}^n t^2 - n \cdot (\bar{t})^2,$ <p>\tilde{S}^2 – wariancja reszt z modelu liniowego otrzymanego po transformacji prognozowanej zmiennej (postaci zlinearyzowanej).</p>

Rodzaj trendu	Zapis funkcji trendu	Ocena błędu prognoz <i>ex ante</i>
1	2	3
		<p>Wiedząc, że model został sprowadzony do postaci liniowej w celu oszacowania średniego błędu prognozy dla wykładniczej postaci modelu, dokonujemy adekwatnego przekształcenia przedstawionego poniżej.</p> <p>Ocena średniego błędu <i>ex ante</i> prognozy dla zmiennej Y (pierwotnej, tzn. dla wykładniczej postaci modelu):</p> $V_T = \exp(\tilde{V}_T).$ <p>Błąd względny prognozy <i>ex ante</i>:</p> $\eta_T = \frac{V_T}{y_T} \cdot 100, \quad T > n.$
<i>Trend logarytmiczny</i>	$y_t = \beta_0 + \beta_1 \ln t, \quad a_1 \neq 0$	<p>Błąd bezwzględny prognoz <i>ex ante</i>⁸³:</p> $V_T = \sqrt{S^2 \cdot [(\tilde{T}_T^*)^T (\tilde{T}_T^* \tilde{T}_T^*)^{-1} \tilde{T}_T^* + 1]}.$ <p>Błąd względny prognoz <i>ex ante</i>:</p> $\eta_T = \frac{V_T}{y_T} \cdot 100, \quad T > n,$ <p>gdzie:</p> <p>S^2 – wariancja reszt z modelu,</p> $\tilde{T}_T = \begin{bmatrix} 1 & \ln 1 \\ 1 & \ln 2 \\ 1 & \ln 3 \\ \vdots & \vdots \\ 1 & \ln n \end{bmatrix},$ $\tilde{T}_T^* = \begin{bmatrix} 1 \\ \ln T \end{bmatrix}; \quad (\tilde{T}_T^*)^T = [1 \quad \ln T].$
<i>Trend</i>	$y_t = \beta_0 t^{\beta_1} \quad \beta_1 > 1$	Błąd bezwzględny prognoz <i>ex ante</i> :

⁸³ W literaturze przedmiotu wektor dla okresu T , na który sporządzana jest prognoza, oznaczany jest za pomocą symbolu X . W celu zachowania spójności symboli w zależnościach związanych z oceną dopuszczalności prognoz wprowadzono oznaczenia zastosowane w rozdziale 1.

Rodzaj trendu	Zapis funkcji trendu	Ocena błędu prognoz <i>ex ante</i>
1	2	3
<i>potęgowy</i>		$V_T = Y_T^* \cdot \sqrt{\tilde{S}^2 \cdot [(\tilde{T}_T^*)^T (\tilde{T}^T \tilde{T})^{-1} \tilde{T}_T^* + 1]}.$ <p>Błąd względny prognoz <i>ex ante</i>:</p> $\eta_T = \frac{V_T}{y_T} \cdot 100, \quad T > n.$ <p>Wiedząc, że model został sprawdzony do postaci liniowej, należy przyjąć, że:</p> $\tilde{T} = \begin{bmatrix} 1 & \ln 1 \\ 1 & \ln 2 \\ 1 & \ln 3 \\ \vdots & \vdots \\ 1 & \ln n \end{bmatrix},$ $\tilde{T}_T^* = \begin{bmatrix} 1 \\ \ln T \end{bmatrix} - \text{wektor } \tilde{T}^* \text{ dla okresu } T,$ <p>na który jest sporządzana prognoza (za pomocą \tilde{T} i T w zależnościach ogólnych oznacza się zmienną czasową t),</p> $(\tilde{T}_T^*)^T = [1 \quad \ln T] - \text{transponowany wektor } \tilde{T}^*.$
<i>Trend paraboliczny – wielomian drugiego stopnia</i>	$y_t = a_0 + a_1 t + a_2 t^2 + \dots + a_k t^k$	<p>Błąd bezwzględny prognoz <i>ex ante</i>:</p> $V_T = \sqrt{S^2 \cdot [(T_T^*)^T (T^T T)^{-1} T_T^* + 1]}.$ <p>Błąd względny prognoz <i>ex ante</i>:</p> $\eta_T = \frac{V_T}{y_T} \cdot 100, \quad T > n,$

Rodzaj trendu	Zapis funkcji trendu	Ocena błędu prognoz <i>ex ante</i>
1	2	3
		gdzie: $T = \begin{bmatrix} 1 & 1 & 1 & \dots & 1 \\ 1 & 2 & 4 & \dots & 2^k \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & n & n^2 & \dots & n^k \end{bmatrix},$ $T_T^* = \begin{bmatrix} 1 \\ T \\ T^2 \\ \vdots \\ T^k \end{bmatrix}.$ <p>Wiedząc, że $T_1 = t, T_2 = t^2, \dots, T_k = t^k$.</p>
<i>Trend hiperboliczny</i>	$y_t = a_0 + a_1 \cdot \frac{1}{t}$	Błąd bezwzględny prognoz <i>ex ante</i> : $V_T = \sqrt{S^2 \cdot [(\tilde{T}_T^*)^T (\tilde{T}^T \tilde{T})^{-1} \tilde{T}_T^* + 1]}.$ <p>Błąd względny prognoz <i>ex ante</i>: $\eta_T = \frac{V_T}{y_T} \cdot 100, \quad T > n.$ <p>Wiedząc, że przekształcenie do postaci liniowej ma postać $\tilde{t} = \frac{1}{t}$, to:</p> $\tilde{T} = \begin{bmatrix} 1 & 1 \\ 1 & \frac{1}{2} \\ 1 & \frac{1}{3} \\ \vdots & \vdots \\ 1 & \frac{1}{n} \end{bmatrix}.$ </p>

Źródło: opracowanie własne na podstawie: A. Zeliaś, B. Pawełek, S. Wanat, *Prognozowanie ekonomiczne. Teoria. Przykłady. Zadania*, PWN, Warszawa 2004, s. 71-86.

Oprócz ilościowych miar weryfikujących dopuszczalność prognoz wskazana jest krytyczna analiza procesu i metodyki przygotowania prognozy odnosząca się do ww. aspektów przeprowadzona przez samego

prognostę. Zmienna i często trudno przewidywalna natura zjawisk społecznych, ekonomicznych powoduje, że nawet przy dołożeniu przez prognostę wszelkich starań na etapie budowy modelu uzyskane prognozy mogą być obciążone dużym błędem. Warto zatem wspomnieć o alternatywnym podejściu do podnoszenia jakości prognoz, jakim jest stosowanie tzw. *prognoz łączonych*⁸⁴. W dużym uproszczeniu podejście to zakłada budowę prognozy z wykorzystaniem na przykład dwóch różnych adekwatnych modeli regresji. Za ostateczną wartość prognozy można przyjmując wartość średnią z uzyskanych dzięki tym dwóm modelom wartościom prognoz „indywidualnych”. Uzyskana w ten sposób prognoza jest tzw. prognozą łączoną. Jednym z głównych problemów tego podejścia jest dobór wag prognoz indywidualnych.⁸⁵ Niemniej jednak prognoza łączona daje lepsze wyniki niż prognoza indywidualna z uwagi na dywersyfikację ryzyka. Wyniki prowadzonych badań wskazują, że stosowanie tego podejścia może zredukować błąd prognozy od 4% do 10%, a w niektórych przypadkach nawet o 58% (w zależności od uwarunkowań)⁸⁶.

Przykład 4.2.1. Ocena błędów prognozy *ex ante* modelu liniowego

Po przeprowadzonej analizie kształtowania się zmiennej i predykcji średniej miesięcznej wielkości produkcji energii elektrycznej w Polsce w latach 1998-2010, omówionej w rozdziale 1, można dokonać oceny błędów *ex ante* przygotowanej prognozy. Na podstawie dekompozycji szeregu czasowego i oceny wzrokowej wykresu ilustrującego kształtowanie się zmiennej do jej opisu wykorzystano model liniowy (rozdz. 1, przykł. 1.1).

W celu wyznaczenia wartości błędu należy oszacować zgodnie z zależnością 4.11 odchylenie standardowe składnika resztowego (reszt z modelu), wiedząc, że:

⁸⁴ Y. Yang, Combining forecasting procedures: some theoretical results, *Econometric Theory*, vol. 1, 2004, s. 176-222.

⁸⁵ J. M. Bates, C.J.W. Granger, The combination of forecasts, *Operational Research Quarterly* nr 20, 1969, s. 451-468; R.T. Clemen, Combining forecasts: A review and annotated bibliography, *International Journal of Forecasting* nr 5, North-Holland 1989, s. 559-583.

⁸⁶ A. Graefe, J.S. Armstrong, R.J. Jones Jr., A.G. Cuzan, Combining forecasts: An application to elections, *International Journal of Forecasting*, vol.30, Issue 1, 2014, s. 43-54; G. Elliott, Averaging and the optimal combination of forecasts, *Working Paper University of California*, San Diego 2011, s. 1-30.

- średnia arytmetyczna zmiennej $\bar{t} = 8,0$;
- liczba obserwacji $n = 15$;
- liczba zmiennych objaśniających m (bez zmiennej stojącej przy wyrazie wolnym) $m = 1$;
- suma kwadratów odchyłeń zmiennej prognozowanej i wartości teoretycznych wynikających z modelu $\sum_{t=1}^{15} (y_t - \hat{y}_t)^2 = 370,30$.

Obliczenia pomocnicze, ilustrujące sposób wyznaczenia wartości pośrednich wykorzystywanych w oszacowaniu błędów prognoz *ex ante* dla liniowej postaci modelu analitycznego, przedstawiono w tabeli 4.3.

Tab. 4.3. Wyniki obliczeń pomocniczych wykorzystywanych w oszacowaniu błędów *ex ante* prognoz dla liniowej postaci modelu analitycznego

t	$(t - \bar{t})^2$	y_t	\hat{y}_t	$(y_t - \hat{y}_t)^2$
1	49,00	122,00	133,09	123,03
2	36,00	138,00	135,20	7,83
3	25,00	136,00	137,31	1,72
4	16,00	139,00	139,42	0,18
5	9,00	145,00	141,53	12,01
6	4,00	146,00	143,65	5,54
7	1,00	144,00	145,76	3,08
8	0,00	152,00	147,87	17,08
9	1,00	154,00	149,98	16,18
10	4,00	157,00	152,09	24,13
11	9,00	161,00	154,20	46,26
12	16,00	159,00	156,31	7,24
13	25,00	155,00	158,42	11,70
14	36,00	152,00	160,53	72,78
15	49,00	158,00	162,64	21,55
Σ	280,00		Σ	370,30

Źródło: obliczenia własne.

Zatem, zgodnie z zależnością 4.11, odchylenie standardowe składnika resztowego w analizowanym przypadku wynosi $s = 5,337$.

Kolejnym krokiem jest ocena błędów bezwzględnych prognozy *ex ante*. Na podstawie wyników predykcji stwierdzono, że prognoza produkcji energii elektrycznej w Polsce w 2011 roku wyniesie 164,76 TWh,

w 2012 roku 166,87 TWh, natomiast w 2013 roku 168,98 TWh (patrz: rozdz. 1, tab. 1.3).

Bezwzględny błąd prognoz *ex ante* dla liniowej postaci modelu analitycznego szacowany jest zgodnie z zależnością 4.20:

$$V_{16} = \sqrt{\frac{(16-8)^2}{280,00} + \frac{1}{15}} + 1 \cdot 5,337 = 5,89;$$

$$V_{17} = \sqrt{\frac{(17-8)^2}{280,00} + \frac{1}{15}} + 1 \cdot 5,337 = 6,03;$$

$$V_{18} = \sqrt{\frac{(18-8)^2}{280,00} + \frac{1}{15}} + 1 \cdot 5,337 = 6,17.$$

Znając wartości prognozy oraz bezwzględnych błędów prognoz *ex ante*, wyznaczyć można zgodnie z zależnością 4.13 względny błąd prognoz *ex ante*:

$$\eta_{16} = \frac{V_{16}}{y_{16}^*} \cdot 100 = \frac{5,89}{164,76} \cdot 100 = 3,58\%;$$

$$\eta_{17} = \frac{6,03}{166,87} \cdot 100 = 3,61\%;$$

$$\eta_{18} = \frac{6,17}{168,98} \cdot 100 = 3,65\%.$$

Analizując uzyskane wyniki, można stwierdzić, że bezwzględny błąd prognoz *ex ante* dla okresu $t=16$ wynosi 5,89, co stanowi 3,58% wartości prognozy. Zgodnie z wytycznymi oceny dopuszczalności prognozy, jeżeli wartość względnego błędu prognoz *ex ante* nie przekracza poziomu ustalanego przez zleceniodawcę na etapie formułowania zadania prognostycznego, prognozę można uznać za dopuszczalną (spełnione zostaje subiektywne kryterium dopuszczalności). Ponadto, korzystając z kryteriów obiektywnych (patrz: tab. 4.1), można dokonać jakościowej klasyfikacji prognoz. Na podstawie wyników obliczeń można stwierdzić, że wyznaczone prognozy produkcji energii elektrycznej w Polsce w badanym okresie należą do kategorii prognoz dokładnych (błędy względne

prognoz *ex ante* mieszczą się w arbitralnie przyjętym przedziale, tj. $3\% < \eta_t \leq 5\%$).

Przykład 4.2.2. Ocena błędów prognozy *ex ante* modelu nieliniowego – funkcja logarytmiczna

Na podstawie wyników predykcji stwierdzono, że prognoza liczby gospodarstw domowych wyposażonych w komputery w Polsce w 2013 roku wyniesie 73,68% ogółu gospodarstw (patrz: rozdz. 4.1, przykł. 1.6). Bazując na wynikach przedstawionych w przykładzie 1.6, można przystąpić do oszacowania błędów prognozy *ex ante* zbudowanej z wykorzystaniem modelu nieliniowego logarytmicznego. Do oceny dopuszczalności prognozy w wypadku modelu nieliniowego bazującego należy wykorzystać zależności dotyczące błędu prognoz *ex ante* dedykowane funkcji logarytmicznej (patrz: tab. 4.2). Do rozwiązania mniejszego zadania wykorzystano funkcje matematyczne macierz iloczynów, macierz odwrotna – dostępne w arkuszu kalkulacyjnym MS Excel.

W celu oceny dopuszczalności prognozy sporządzonej z wykorzystaniem modelu nieliniowego należy dokonać stosowanych przekształceń, wyznaczając na początku macierz \tilde{T} (obliczenia pomocnicze do przykładu – patrz: tab. 1.9) oraz wektor T_{11}^* (ocena prognozy sporządzonej na okres $t=11$).

$$\tilde{T} = \begin{bmatrix} 1 & \ln 1 \\ 1 & \ln 2 \\ 1 & \ln 3 \\ \vdots & \vdots \\ 1 & \ln n \end{bmatrix}.$$

Wiedząc, że liczba obserwacji $n = 10$, to:

$$\tilde{T} = \begin{bmatrix} 1 & 0,000 \\ 1 & 0,693 \\ 1 & 1,099 \\ 1 & 1,386 \\ 1 & 1,609 \\ 1 & 1,791 \\ 1 & 1,946 \\ 1 & 2,079 \\ 1 & 2,197 \\ 1 & 2,303 \end{bmatrix}.$$

Wektor T_T^* (dla $T = 11$, $\ln(11) = 2,398$):

$$T_{11}^* = \begin{bmatrix} 1 \\ 2,398 \end{bmatrix}.$$

Po wyznaczeniu macierzy \tilde{T} należy dokonać transpozycji macierzy \tilde{T}^T oraz wektora T_{11}^* . W tym celu można skorzystać z opcji dostępnych w arkuszu kalkulacyjnym, tj. Wklej specjalnie, opcja *Wartości i Transpozycja*.

$$\tilde{T}^T = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0,000 & 0,693 & 1,099 & 1,386 & 1,609 & 1,791 & 1,946 & 2,079 & 2,197 & 2,303 \end{bmatrix},$$

$$(T_{11}^*)^T = [1 \quad 2,398].$$

Kolejny krok polega na wyznaczeniu macierzy $\tilde{X}^T \cdot \tilde{X}$ (będzie to macierz o wymiarach 2x2, taki też obszar należy zaznaczyć jako obszar wynikowy obliczeń w arkuszu kalkulacyjnym). Do wykonania tych obliczeń można skorzystać z funkcji matematycznej `MACIERZ.ILOCZYN()`.

$$\tilde{T}^T \cdot \tilde{T} = \begin{bmatrix} 10,000 & 15,104 \\ 15,104 & 27,650 \end{bmatrix}.$$

Uzyskaną ten sposób macierz należy odwrócić, wyznaczając tym samym macierz $(\tilde{T}^T \cdot \tilde{T})^{-1}$. Do wykonania tej operacji można skorzystać z funkcji matematycznej MACIERZ.ODW() (będzie to, jak poprzednio, macierz o wymiarach 2x2).

$$(\tilde{T}^T \cdot \tilde{T})^{-1} = \begin{bmatrix} 0,572 & -0,312 \\ -0,312 & 0,207 \end{bmatrix}.$$

W celu wyznaczenia bezwzględnego błędu prognoz *ex ante* należy wyznaczyć następnie iloczyn macierzy $(X_{11}^*)^T$ i $(\tilde{X}^T \cdot \tilde{X})^{-1}$ (uzyskany wynik będzie macierzą 1x1, taki też obszar należy zaznaczyć jako obszar wynikowy obliczeń w arkuszu kalkulacyjnym) oraz iloczyn macierzy $(T_{11}^*)^T \cdot (\tilde{T}^T \cdot \tilde{T})^{-1} \cdot T_{11}^*$:

$$(T_{11}^*)^T \cdot (\tilde{T}^T \cdot \tilde{T})^{-1} = [-0,177 \quad 0,184];$$

$$(T_{11}^*)^T \cdot (\tilde{T}^T \cdot \tilde{T})^{-1} \cdot T_{11}^* = 0,263.$$

Wiedząc, że odchylenie standardowe składnika resztowego w analizowanym przypadku zgodnie z zależnością 4.11 wynosi $s = 3,631$, można wyznaczyć bezwzględny błąd prognozy *ex ante* na okres $t=11$:

$$V_{11} = s \cdot \sqrt{[(T_{11}^*)^T (\tilde{T}^T \tilde{T})^{-1} T_{11}^* + 1]} = 3,631 \cdot \sqrt{0,263 + 1} = 4,08.$$

Znając wartość prognozy oraz bezwzględnego błędu prognoz *ex ante*, wyznaczyć można zgodnie z zależnością 4.13 względny błąd prognoz *ex ante*:

$$\eta_{11} = \frac{V_{11}}{y_{11}^*} \cdot 100 = \frac{4,08}{73,68} \cdot 100 = 5,54\%.$$

Bezwzględny błąd prognozy *ex ante* dla okresu $t=11$ wyniósł 4,08 – co stanowi 5,54% wartości prognozy. Jeżeli wartość względnego błędu prognoz *ex ante* nie przekracza poziomu ustalanego przez zleceniodawcę na etapie formułowania zadania prognostycznego, prognozę można uznać za dopuszczalną (spełnione zostaje subiektywne kryterium dopuszczalności). Według przyjętych kryteriów obiektywnych jakościowej klasyfikacji prognoz (patrz: tab.4.1) można stwierdzić, że wyznaczona

prognoza liczby gospodarstw domowych wyposażonych w komputery w Polsce w 2013 roku do ogółu gospodarstw należy do kategorii prognoz dopuszczalnych (błędy względne prognoz *ex ante* mieszczą się w arbitralnie przyjętym przedziale, tj. $5\% < \eta_t \leq 10\%$). W przedstawiony w przykładzie sposób można oszacować błędy prognoz *ex ante* dla $t = 12$ i $t = 13$.

Kluczowe zagadnienia:

dopuszczalność prognoz, bezwzględny i względny błąd prognozy *ex ante*, maksymalny horyzont czasowy prognozy, kryteria dopuszczalności prognoz

Literatura podstawowa:

1. Cieślak M. (red.), *Prognozowanie gospodarcze. Metody i zastosowanie*, Wydawnictwo Naukowe PWN, Warszawa 2011.
2. Dittmann P., *Metody prognozowania sprzedaży w przedsiębiorstwie*, Wydawnictwo Akademii Ekonomicznej im. O. Langego we Wrocławiu, Wrocław 2000.
3. Zeliaś A., Pawełek B., Wanat S., *Prognozowanie ekonomiczne. Teoria, przykłady, zadania*, Wydawnictwo Naukowe PWN, Warszawa 2004.

Literatura uzupełniająca:

1. Bates J.M., Granger C.J.W., The combination of forecasts, *Operational Research Quarterly* nr 20, 1969, s. 451-468.
2. Clemen R.T., Combining forecasts: A review and annotated bibliography, *International Journal of Forecasting* nr 5, North-Holland 1989, s. 559-583.
3. Elliott G., Averaging and the optimal combination of forecasts, *Working Paper University of California*, San Diego 2011, s. 1-30.

4. Featherstone K., James D., Assessment of NZIER Forecast Accuracy, *New Zealand Institute of Economic Research, Research Working Paper 94/9*, Wellington 1994.
5. Graefe A., Armstrong J.S., Jones Jr. R.J., Cuzan A.G., Combining forecasts: An application to elections, *International Journal of Forecasting*, vol. 30 Issue 1, 2014, s. 43–54.
6. Larose D.T., *Metody i modele eksploracji danych*, Wydawnictwo Naukowe PWN, Warszawa 2016.
7. Nazarko J. (red.), *Prognozowanie w zarządzaniu przedsiębiorstwem. Cz. I. Wprowadzenie do metodyki prognozowania*, Wydawnictwo Politechniki Białostockiej, Białystok 2004.
8. Nazarko J. (red.), *Prognozowanie w zarządzaniu przedsiębiorstwem. Cz. II. Prognozowanie na podstawie szeregów czasowych*, Wydawnictwo Politechniki Białostockiej, Białystok 2004.
9. Nazarko J. (red.), *Prognozowanie w zarządzaniu przedsiębiorstwem. Cz. III. Prognozowanie na podstawie modeli adaptacyjnych*, Wydawnictwo Politechniki Białostockiej, Białystok 2005.
10. Orwat A., Przykład zastosowania metody odpornej w modelowaniu finansowych szeregów czasowych. *Zeszyty Naukowe Szkoły Głównej Gospodarstwa Wiejskiego. Ekonomika i Organizacja Gospodarki Żywnościowej* 60, 2006, s. 279-288.
11. Ostasiewicz S., *Wykorzystanie metod dyskryminacyjnych w prognozowaniu dyskretnym*, Prace Naukowe Akademii Ekonomicznej we Wrocławiu nr 232, seria: Monografie i Opracowania nr 14, 1983.
12. Welfe A., *Ekonometria*, Wydawnictwo PWE, Warszawa 1995.
13. Yang Y., Combining forecasting procedures: some theoretical results, *Econometric Theory*, vol. 1, 2004, s. 176-222.
14. Zeliaś A., *Teoria prognozy*, PWE, Warszawa 1997.

5. PRZYKŁADY ZASTOSOWAŃ

5.1. Prognoza liczby gospodarstw domowych w Polsce posiadających dostęp do szerokopasmowego Internetu (wyrażona w procentach ogółu gospodarstw domowych)

I. Sformułowanie problemu

W Polsce w ostatnich latach zarówno rynek usług szerokopasmowych, jak i niezbędna infrastruktura telekomunikacyjna dynamicznie się rozwijają. Najbardziej rozpowszechnione połączenia szerokopasmowe w Polsce to przede wszystkim łącza DSL przez linię telefoniczną, sieci kablowe oraz komórkowy Internet LTE. Dostępność sieci szerokopasmowych dla użytkowników uzależniona jest od dostępu do sieci kablowych, telefonicznych, światłowodów lub od odległości od stacji bazowych operatorów komórkowych. Prognoza liczby gospodarstw domowych w Polsce posiadających dostęp do szerokopasmowego Internetu może być przydatna przedsiębiorstwom świadczącym usługi telekomunikacyjne, operatorom sieci komórkowych, czy też właścicielom telewizji kablowych. Może być również wykorzystywana przez przedsiębiorstwa produkujące sprzęt telekomunikacyjny. Ocena dostępności Internetu jest też kluczowym elementem diagnozy rozwoju społeczeństwa informacyjnego w Polsce przeprowadzonej systematycznie przez Główny Urząd Statystyczny.

Jedno z dużych przedsiębiorstw telekomunikacyjnych zwróciło się do instytucji badającej rynek z prośbą o opracowanie modelu do prognozowania liczby gospodarstw domowych w Polsce posiadających dostęp do Internetu szerokopasmowego. Prognoza wyznaczona przy wykorzystaniu opracowanego modelu miała stanowić źródło informacji zarządczych potrzebnych do opracowania planu rozwoju usług świadczonych przez to przedsiębiorstwo w trzyletniej perspektywie czasowej.

Celem zadania jest wyznaczenie prognozy liczby gospodarstw domowych posiadających dostęp do Internetu szerokopasmowego na trzy kolejne lata, tj. na 2016, 2017 i 2018 rok. Wszelkie obliczenia, wykresy i analizy przeprowadzono w arkuszu kalkulacyjnym Microsoft Excel.

Zgodnie ze schematem procesu budowy prognozy (rys. 1.1) prognozowanie na podstawie modelu analitycznego przebiega w czterech następujących po sobie etapach: specyfikacja, estymacja, weryfikacja, predykcja.

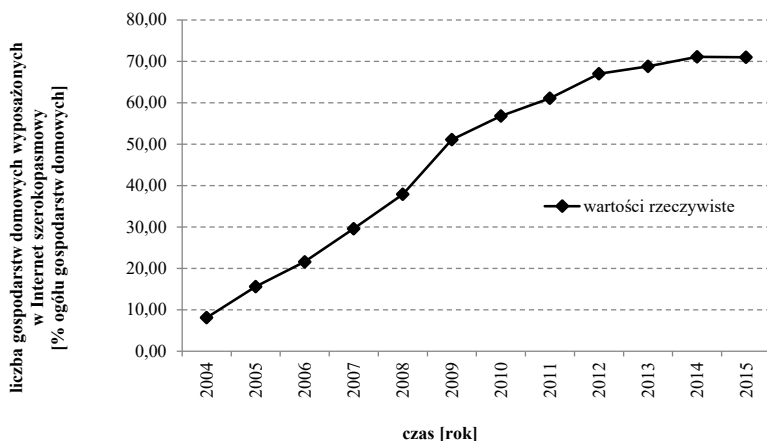
II. Specyfikacja

Dane dotyczące liczby gospodarstw domowych posiadających dostęp do Internetu szerokopasmowego, wyrażone w procentach ogółu gospodarstw domowych, przedstawiono w tabeli 5.1.1 oraz na wykresie na rysunku 5.1.1. Zgodnie z przyjętą metodyką, na podstawie zgromadzonych danych i ich wizualizacji, przeprowadzono analizę graficzną przebiegu zmiennej, określono charakter i zidentyfikowano prawidłowości zmian poziomu badanego zjawiska w czasie.

Tab. 5.1.1. Dane dotyczące liczby gospodarstw domowych wyposażonych w Internet [% ogółu gospodarstw domowych] w latach 2004-2015

t	Rok	y_t
1	2004	8,10
2	2005	15,60
3	2006	21,60
4	2007	29,60
5	2008	37,90
6	2009	51,10
7	2010	56,80
8	2011	61,10
9	2012	67,00
10	2013	68,80
11	2014	71,10
12	2015	71,00

Źródło: opracowanie własne na podstawie rocznych wskaźników makroekonomicznych, dokument elektroniczny http://www.stat.gov.pl/gus/wskazniki_makroekon_PLK_HTML.htm, data wejścia 29.04.2016.



Rys. 5.1.1. Liczba gospodarstw domowych wyposażonych w Internet w Polsce w latach 2004-2015 [% ogółu gospodarstw domowych]

Źródło: opracowanie własne.

W trakcie wstępnej analizy wizualnej przebiegu badanej zmiennej (rys. 5.1.1) zauważano początkowo szybki wzrost jej wartości (od 2004 do 2011 roku), przy czym jego tempo wyraźnie zwalniało w miarę upływu czasu (od 2012 do 2015 roku). Zaobserwowana prawidłowość może wskazywać, że najlepiej odzwierciedlającymi przebieg badanej zmiennej mogą być funkcje nieliniowe o malejącym tempie wzrostu. Do wyznaczenia prognozy liczby gospodarstw domowych posiadających dostęp do Internetu szerokopasmowego arbitralnie wybrano trzy modele prognostyczne – zbudowane z wykorzystaniem odpowiednio funkcji liniowej, logarytmicznej, potęgowej. Biorąc pod uwagę wyniki oceny wzrokowej przebiegu zmiennej oraz specyfikę zjawiska, funkcje te powinny dobrze opisywać tendencję do malejącego tempa wzrostu liczby gospodarstw domowych w Polsce posiadających dostęp do Internetu szerokopasmowego.

Zaproponowane funkcje, jak już wspomniano, mają charakter nieliniowy. Dlatego też zastosowanie klasycznej metody najmniejszych kwadratów do estymacji parametrów tych funkcji będzie możliwe po transformacji do postaci liniowej.

III. Estymacja

III a. Funkcja liniowa

Pierwszy model prognostyczny zbudowano z wykorzystaniem funkcji liniowej o postaci: $y_t = \beta_0 + \beta_1 t$.

Do wyznaczenia paramentów tego modelu klasyczną metodą najmniejszych kwadratów konieczne jest obliczenie średnich arytmetycznych następujących wartości:

$$\bar{y} = \frac{1}{n} \sum_{t=1}^n y_t = \frac{559,70}{12} = 46,642;$$

$$\bar{t} = \frac{1}{n} \sum_{t=1}^n t_i = \frac{78}{12} = 6,500.$$

Obliczenia pośrednie, niezbędne do oszacowania wartości parametrów modelu, zawarto w tabeli 5.1.2.

Tab. 5.1.2. Obliczenia parametrów modelu metodą najmniejszych kwadratów*

Rok	t	y_t	$t - \bar{t}$	$(t - \bar{t})^2$	$(t - \bar{t})y_t$
2004	1	8,10	1-6,500=-5,50	$(-5,50)^2=30,25$	$-5,50 \cdot 8,10=-44,55$
2005	2	15,60	2-6,500=-4,50	$(-4,50)^2=20,25$	$-4,50 \cdot 15,60=-70,20$
2006	3	21,60	3-6,500=-3,50	$(-3,50)^2=12,25$	$-3,50 \cdot 21,60=-75,60$
2007	4	29,60	...=-2,50	...=6,25	...=-74,00
2008	5	37,90	-1,50	2,25	-56,85
2009	6	51,10	-0,50	0,25	-25,55
2010	7	56,80	0,50	0,25	28,40
2011	8	61,10	1,50	2,25	91,65
2012	9	67,00	2,50	6,25	167,50
2013	10	68,80	3,50	12,25	240,80
2014	11	71,10	4,50	20,25	319,95
2015	12	71,00	5,50	30,25	390,50
Σ		559,70		143,00	892,05

Źródło: obliczenia własne.

* Wartości w tabeli zostały zaokrąglone do dwóch miejsc po przecinku.

Korzystając z obliczeń pośrednich umieszczonych w tabeli 5.1.2 oraz biorąc pod uwagę wartości średnie \bar{y} i \bar{t} , obliczono wartość parametrów β_0 , β_1 :

$$\beta_1 = \frac{\sum_{t=1}^n (t - \bar{t}) y_t}{\sum_{t=1}^n (t - \bar{t})^2} = \frac{892,05}{143,00} = 6,238;$$

$$\beta_0 = \bar{y} - \beta_1 \bar{t} = 46,642 - (6,238 \cdot 6,500) = 6,095.$$

Model oszacowany metodą najmniejszych kwadratów ostatecznie można przedstawić w następujący sposób:

$$\hat{y}_t = \beta_0 + \beta_1 t = 6,095 + 6,238 t.$$

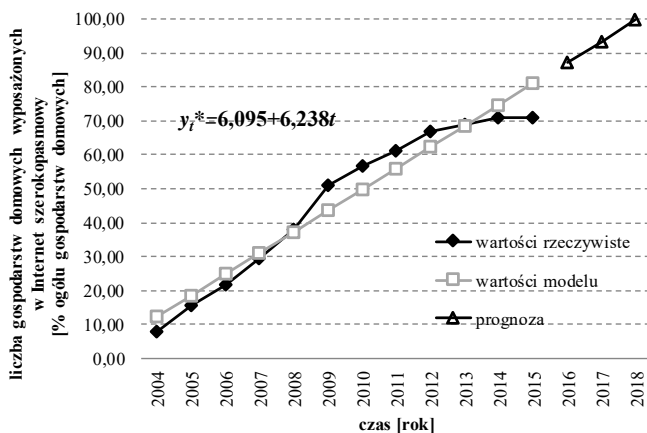
Wartości rzeczywiste oraz wartości modelu zbudowanego z wykorzystaniem funkcji liniowej zostały przedstawione w tabeli 5.1.3.

Tab. 5.1.3. Wartości rzeczywiste oraz wartości modelu zbudowanego z wykorzystaniem funkcji liniowej

Rok	t	y_t	$\hat{y}_t = 6,095 + 6,238t$
2004	1	8,10	$6,095 + 6,238 \cdot 1 = 12,33$
2005	2	15,60	$6,095 + 6,238 \cdot 2 = 18,57$
2006	3	21,60	$6,095 + 6,238 \cdot 3 = 24,81$
2007	4	29,60	$\dots = 31,05$
2008	5	37,90	37,29
2009	6	51,10	43,52
2010	7	56,80	49,76
2011	8	61,10	56,00
2012	9	67,00	62,24
2013	10	68,80	68,48
2014	11	71,10	74,71
2015	12	71,00	80,95

Źródło: obliczenia własne.

Wykres wartości rzeczywistych liczby gospodarstw domowych w Polsce posiadających dostęp do Internetu szerokopasmowego oraz wartości modelu opracowanego z wykorzystaniem funkcji liniowej przedstawiono na rysunku 5.1.2.



Rys. 5.1.2. Wartości rzeczywiste oraz wartości modelu zbudowanego z wykorzystaniem funkcji liniowej

Źródło: opracowanie własne.

III b. Funkcja logarytmiczna

Kolejny model prognostyczny zbudowano z wykorzystaniem nieliniowej funkcji logarytmicznej w postaci: $y_t = \beta_0 + \beta_1 \ln t$, $\beta_1 > 0$. W celu oszacowania parametrów β_0 , β_1 dokonano linearyzacji tej funkcji. Uzyskanie liniowej postaci modelu możliwe jest po dokonaniu przekształceń: $t' = \ln t$. Ostatecznie uzyskano model liniowy w postaci: $y_t = \beta_0 + \beta_1 t'$, $\beta_1 > 0$.

Do wyznaczenia paramentów tego modelu klasyczną metodą najmniejszych kwadratów konieczne jest obliczenie średnich arytmetycznych następujących wartości:

$$\bar{y} = \frac{1}{n} \sum_{t=1}^n y_t = \frac{559,70}{12} = 46,642;$$

$$\bar{t}' = \frac{1}{n} \sum_{t=1}^n t'_i = \frac{19,99}{12} = 1,666.$$

Obliczenia pośrednie niezbędne do oszacowania wartości parametrów modelu przedstawiono w tabeli 5.1.4.

Tab. 5.1.4. Obliczenia parametrów modelu metodą najmniejszych kwadratów

Rok	t	y_t	$t' = \ln t$	$t' - \bar{t}'$	$(t' - \bar{t}')^2$	$(t' - \bar{t}') y_t$
2004	1	8,10	$\ln 1 = 0,00$	$1 - 1,666 = -1,67$	$(-1,67)^2 = 2,79$	$1,67 \cdot 8,10 = -13,53$
2005	2	15,60	$\ln 2 = 0,69$	$2 - 1,666 = -0,98$	$(-0,98)^2 = 0,96$	$0,98 \cdot 15,60 = -15,29$
2006	3	21,60	$\ln 3 = 1,10$	$3 - 1,666 = -0,57$	$(0,57)^2 = 0,32$	$0,57 \cdot 21,60 = -12,31$
2007	4	29,60	$\dots = 1,39$	-0,28	0,08	-8,29
2008	5	37,90	1,61	-0,06	0,00	-2,27
2009	6	51,10	1,79	0,12	0,01	6,13
2010	7	56,80	1,95	0,28	0,08	15,90
2011	8	61,10	2,08	0,41	0,17	25,05
2012	9	67,00	2,20	0,53	0,28	35,51
2013	10	68,80	2,30	0,63	0,40	43,34
2014	11	71,10	2,40	0,73	0,53	51,90
2015	12	71,00	2,48	0,81	0,66	57,51
Σ		559,70	19,99		6,28	183,65

Źródło: obliczenia własne.

Korzystając z obliczeń pośrednich umieszczonych w tabeli 5.1.4 oraz biorąc pod uwagę wartości średnie \bar{y} i \bar{t}' , obliczono wartość parametrów β_0 , β_1 :

$$\beta_1 = \frac{\sum_{t=1}^n (t' - \bar{t}') y_t}{\sum_{t=1}^n (t' - \bar{t}')^2} = \frac{183,65}{6,28} = 29,244;$$

$$\beta_0 = \bar{y} - \beta_1 \bar{t}' = 46,642 - 29,244 \cdot 1,666 = -2,079.$$

Model oszacowany metodą najmniejszych kwadratów ostatecznie można przedstawić w następujący sposób:

$$\hat{y}_t = \beta_0 + \beta_1 \ln t = -2,079 + 29,244 \ln t.$$

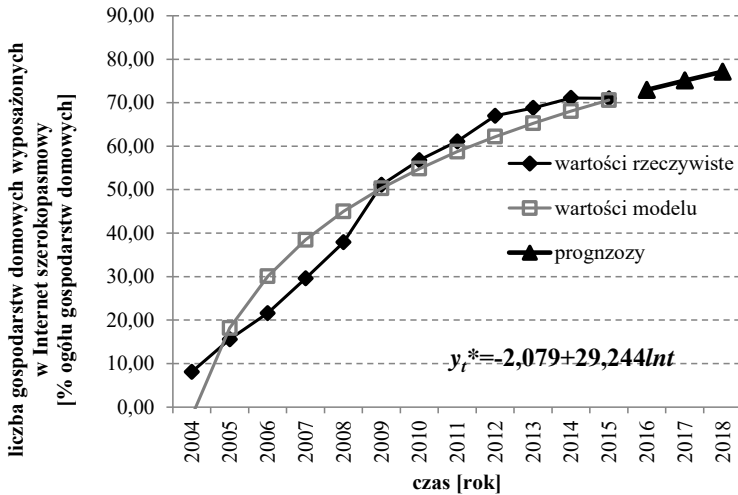
Wartości rzeczywiste oraz wartości modelu zbudowanego z wykorzystaniem funkcji logarytmicznej zostały przedstawione w tabeli 5.1.5

Tab. 5.1.5. Wartości rzeczywiste oraz wartości modelu zbudowanego z wykorzystaniem funkcji logarytmicznej

Rok	t	y_t	$\hat{y}_t = -2,079 + 29,244 \ln t$
2004	1	8,10	$-2,079 + 29,244 \ln 1 = -2,08$
2005	2	15,60	$-2,079 + 29,244 \ln 2 = 18,19$
2006	3	21,60	$-2,079 + 29,244 \ln 3 = 30,05$
2007	4	29,60	$\dots = 38,46$
2008	5	37,90	44,99
2009	6	51,10	50,32
2010	7	56,80	54,83
2011	8	61,10	58,73
2012	9	67,00	62,18
2013	10	68,80	65,26
2014	11	71,10	68,05
2015	12	71,00	70,59

Źródło: obliczenia własne.

Wykres wartości rzeczywistych liczby gospodarstw domowych w Polsce wyposażonych w Internet szerokopasmowy oraz wartości modelu opracowanego z wykorzystaniem funkcji nieliniowej logarytmicznej został przedstawiony na rysunku 5.1.3.



Rys. 5.1.3. Wartości rzeczywiste oraz wartości modelu zbudowanego z wykorzystaniem funkcji logarytmicznej

Źródło: opracowanie własne.

III c. Funkcja potęgowa

Kolejny model prognostyczny zbudowano z wykorzystaniem nieliniowej funkcji potęgowej o postaci: $y_t = \beta_0 t^{\beta_1}$, $0 < \beta_1 < 1$. Dokonując przekształceń: $y'_t = \ln y_t$, $\beta'_0 = \ln \beta_0$, $t' = \ln t$, otrzymano następującą liniową postać modelu: $y'_t = \beta'_0 + \beta_1 t'$, $0 < \beta_1 < 1$.

Do wyznaczenia paramentów tego modelu klasyczną metodą najmniejszych kwadratów konieczne jest obliczenie średnich arytmetycznych następujących wartości:

$$\bar{y}' = \frac{1}{n} \sum_{t=1}^n y'_t = \frac{43,96}{12} = 3,663;$$

$$\bar{t}' = \frac{1}{n} \sum_{t=1}^n t'_t = \frac{19,99}{12} = 1,666.$$

Obliczenia pośrednie niezbędne do oszacowania wartości parametrów modelu umieszczono w tabeli 5.1.6.

Tab. 5.1.6. Obliczenia parametrów modelu metodą najmniejszych kwadratów*

Rok	t	y_t	$t' = \ln t$	$y_t' = \ln y_t$	$t' - \bar{t}'$	$(t' - \bar{t}')^2$	$(t' - \bar{t}')y_t'$
2004	1	8,10	$\ln 1=0,00$	$\ln(8,10)=2,09$	$1-1,666=-1,67$	$(-1,67)^2=2,79$	$-1,67 \cdot 2,09=-3,49$
2005	2	15,60	$\ln 2=0,69$	$\ln(15,60)=2,75$	$2-1,666=-0,98$	$(-0,98)^2=0,96$	$-0,98 \cdot 2,75=-2,70$
2006	3	21,60	$\ln 3=1,10$	$\ln(21,60)=3,07$	$3-1,666=-0,57$	$(0,57)^2=0,32$	$-0,57 \cdot 2,16=-1,75$
2007	4	29,60	$\dots=1,39$	$\dots=3,39$	-0,28	0,08	-0,95
2008	5	37,90	1,61	3,63	-0,06	0,00	-0,22
2009	6	51,10	1,79	3,93	0,12	0,01	0,47
2010	7	56,80	1,95	4,04	0,28	0,08	1,13
2011	8	61,10	2,08	4,11	0,41	0,17	1,69
2012	9	67,00	2,20	4,20	0,53	0,28	2,23
2013	10	68,80	2,30	4,23	0,63	0,40	2,66
2014	11	71,10	2,40	4,26	0,73	0,53	3,11
2015	12	71,00	2,48	4,26	0,81	0,66	3,45
Σ		559,70	19,99	43,96		6,28	5,63

Źródło: obliczenia własne.

Korzystając z obliczeń pośrednich umieszczonych w tabeli 5.1.6 oraz biorąc pod uwagę wartości średnie \bar{y}' i \bar{t}' , obliczono wartość parametrów β_0 , β_1 :

$$\beta_1 = \frac{\sum_{t=1}^n (t' - \bar{t}') y_t'}{\sum_{t=1}^n (t' - \bar{t}')^2} = \frac{5,63}{6,28} = 0,896;$$

$$\beta_1 t' = \bar{y}' - \beta_1 t' = 3,663 - (0,896 \cdot 1,666) = -2,170;$$

$$\beta_0 = e^{\beta_1} = e^{-2,170} = 8,758.$$

* Wartości w tabeli zostały zaokrąglone do dwóch miejsc po przecinku.

Model oszacowany metodą najmniejszych kwadratów ostatecznie można przedstawić w następujący sposób:

$$\hat{y}_t = \beta_0 t^{\beta_1} = 8,758 t^{0,896}.$$

Wartości rzeczywiste oraz wartości modelu zbudowanego z wykorzystaniem funkcji potęgowej zostały przedstawione w tabeli 5.1.7.

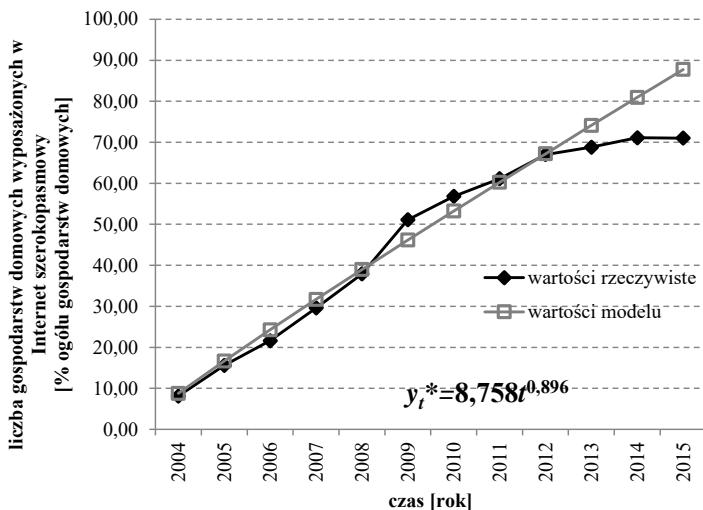
Tab. 5.1.7. Wartości rzeczywiste oraz wartości modelu zbudowanego z wykorzystaniem funkcji potęgowej*

Rok	t	y_t	$\hat{y}_t = 8,758 t^{0,896}$
2004	1	8,10	$8,758 \cdot 1^{0,896} = 8,76$
2005	2	15,60	$8,758 \cdot 2^{0,896} = 16,66$
2006	3	21,60	$8,758 \cdot 3^{0,896} = 24,26$
2007	4	29,60	... = 31,68
2008	5	37,90	38,96
2009	6	51,10	46,14
2010	7	56,80	53,22
2011	8	61,10	60,24
2012	9	67,00	67,19
2013	10	68,80	74,10
2014	11	71,10	80,94
2015	12	71,00	87,74

Źródło: obliczenia własne.

Wykres wartości rzeczywistych liczby gospodarstw domowych w Polsce posiadających dostęp do Internetu szerokopasmowego oraz wartości modelu opracowanego z wykorzystaniem funkcji nieliniowej potęgowej przedstawiono na rysunku 5.1.4.

* Wartości w tabeli zostały zaokrąglone do dwóch miejsc po przecinku.



Rys. 5.1.4. Wartości rzeczywiste oraz wartości modelu zbudowanego z wykorzystaniem funkcji potęgowej

Źródło: opracowanie własne.

IV. Weryfikacja

Weryfikacja obejmuje sprawdzenie, czy model(e) jest(są) zgodny(e) z danymi empirycznymi i formułowanymi założeniami dotyczącymi kształtowania się analizowanego zjawiska w czasie.

Tab. 5.1.8. Zestawienie porównawcze wyników obliczeń

Miara oceny	Liniowa postać funkcji trendu	Logarytmiczna postać funkcji trendu	Potęgowa postać funkcji trendu
Współczynnik determinacji R^2	0,948	0,938	0,985
Odchylenie standardowe reszt z modelu s	5,544	6,042	0,09 (dla liniowej postaci funkcji trendu)

Źródło: obliczenia własne.

Biorąc pod uwagę wartości współczynnika determinacji R^2 , rekomendacji można udzielić modelowi zbudowanemu z wykorzystaniem potęgowej funkcji trendu. W tym przypadku model ten odzwierciedla 98,5% zmienności zmiennej prognozowanej. Wartość ta wskazuje na wysoką jakość dopasowania modelu potęgowego do danych. Zbliżone oceny dopasowania uzyskał dużo prostszy model liniowy. Dobre odzwierciedlenie zmienności zmiennej potwierdza niższa, w porównaniu z logarytmiczną postacią trendu, wartość odchylenia standardowego reszt z modelu.

V. Predykcja

Wyznaczenie prognozy, z wykorzystaniem funkcji analitycznych, sprowadza się do obliczenia wartości funkcji liniowej, logarytmicznej i potęgowej dla przyszłego momentu lub okresu czasu, tzn. dla 2016, 2017 i 2018 roku.

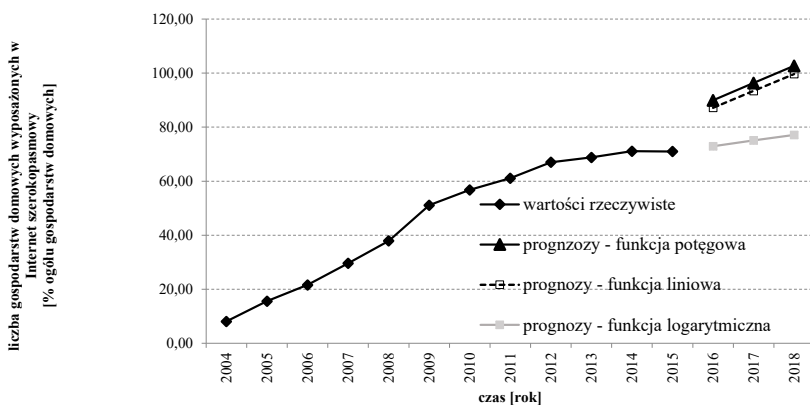
Wartości teoretyczne wynikające z modeli oraz oszacowania prognoz zostały przedstawione w tabeli 5.1.9.

Tab. 5.1.9. Wartości rzeczywiste oraz wartości modelu zbudowanego z wykorzystaniem funkcji liniowej, logarytmicznej i potęgowej

Rok	t	$\hat{y}_t = 6,095 + 6,238t$	$y_t^* = -2,079 + 29,244 \ln t$	$y_t^* = 8,758t^{0,896}$
2016	13	$6,095 + 6,238 \cdot 13 = 87,19$	$-2,079 + 29,44 \ln 13 = 72,93$	$8,758 \cdot 13^{0,896} = 89,95$
2017	14	$6,095 + 6,238 \cdot 14 = 93,43$	$-2,079 + 29,44 \ln 14 = 75,10$	$8,758 \cdot 14^{0,896} = 96,35$
2018	15	$6,095 + 6,238 \cdot 15 = 99,67$	$-2,079 + 29,44 \ln 15 = 77,12$	$8,758 \cdot 15^{0,896} = 102,71$

Źródło: obliczenia własne.

Wykres wartości rzeczywistych oraz prognozy liczby gospodarstw domowych w Polsce wyposażonych w Internet szerokopasmowy wyznaczone z wykorzystaniem funkcji liniowej, logarytmicznej i potęgowej, został przedstawiony na rysunku 5.1.5.



Rys. 5.1.5. Wartości rzeczywiste oraz prognozy liczby gospodarstw domowych posiadających dostęp do Internetu szerokopasmowego w Polsce wyznaczone z wykorzystaniem funkcji liniowej, logarytmicznej i potęgowej

Źródło: opracowanie własne.

VI. Ocena porównawcza dopuszczalności prognoz

W celu oceny błędów prognoz *ex ante* dla liniowej postaci (część III a) modelu analitycznego wyznaczono odchylenie standardowe reszt z modelu (s), wiedząc, że:

- średnia arytmetyczna zmiennej $\bar{t} = 6,5$;
- liczba obserwacji $n = 12$;
- liczba zmiennych objaśniających (bez zmiennej stojącej przy wyrazie wolnym) $m = 1$;
- suma kwadratów odchylenia zmiennej prognozowanej i wartości teoretycznych wynikających z modelu $\sum_{t=1}^{15} (y_t - \hat{y}_t)^2 = 307,32$.

$$\sum_{t=1}^{15} (y_t - \hat{y}_t)^2 = 307,32.$$

Zgodnie z zależnością 4.11 odchylenie standardowe reszt z modelu w analizowanym przypadku wynosi $s = 5,554$. Zatem bezwzględny błąd prognoz *ex ante* dla liniowej postaci modelu analitycznego (patrz: zależność 4.20) badanej zmiennej wynosi:

$$V_{13} = \sqrt{\frac{(13 - 6,5)^2}{307,315} + \frac{1}{12}} + 1 \cdot 5,554 = 6,510;$$

$$V_{14} = \sqrt{\frac{(14-6,5)^2}{307,315} + \frac{1}{12}} + 1 \cdot 5,554 = 6,737;$$

$$V_{15} = \sqrt{\frac{(15-6,5)^2}{307,315} + \frac{1}{12}} + 1 \cdot 5,554 = 6,988.$$

Znając wartości prognozy (patrz: tab. 5.1.8) oraz bezwzględnych błędów prognoz *ex ante*, wyznaczono względne błędy prognoz *ex ante*:

$$\eta_{13} = \frac{V_{13}}{y_{13}^*} \cdot 100 = \frac{6,51}{87,19} \cdot 100 = 7,47\%;$$

$$\eta_{14} = \frac{V_{14}}{y_{14}^*} \cdot 100 = \frac{6,737}{93,43} \cdot 100 = 7,21\%;$$

$$\eta_{15} = \frac{V_{15}}{y_{15}^*} \cdot 100 = \frac{6,988}{99,67} \cdot 100 = 7,01\%.$$

Analizując uzyskane wyniki obliczeń, można stwierdzić, że bezwzględny błąd prognozy *ex ante* liczby gospodarstw domowych wyposażonych w Internet w Polsce (wyznaczonej w stosunku do ogółu gospodarstw domowych) wyniósł 6,51 – co stanowi 7,47% wartości prognozy przygotowanej na 2016 rok.

W celu oceny dopuszczalności prognozy sporządzonej z wykorzystaniem modelu nieliniowego (logarytmicznego i potęgowego) należy dokonać stosowanych przekształceń, wyznaczając na początku macierz \tilde{T} (obliczenia pomocnicze do przykładu – patrz: tab. 1.9) oraz wektor X_T^* (ocena prognozy sporządzonej na okres T).

$$\tilde{T} = \begin{bmatrix} 1 & \ln 1 \\ 1 & \ln 2 \\ 1 & \ln 3 \\ \vdots & \vdots \\ 1 & \ln n \end{bmatrix}.$$

Wiedząc, że liczba obserwacji $n = 12$, to:

$$\tilde{T} = \begin{bmatrix} 1 & 0 \\ 1 & 0,693 \\ 1 & 1,099 \\ 1 & 1,386 \\ 1 & 1,609 \\ 1 & 1,792 \\ 1 & 1,946 \\ 1 & 2,079 \\ 1 & 2,197 \\ 1 & 2,303 \\ 1 & 2,398 \\ 1 & 2,485 \end{bmatrix}.$$

W kolejnym kroku wyznaczany wektor T_T^* (dla $T = 13$ i $T = 14$ oraz $T = 15$):

$$T_{13}^* = \begin{bmatrix} 1 \\ 2,565 \end{bmatrix},$$

$$T_{14}^* = \begin{bmatrix} 1 \\ 2,639 \end{bmatrix},$$

$$T_{15}^* = \begin{bmatrix} 1 \\ 2,708 \end{bmatrix}.$$

Po wyznaczeniu macierzy \tilde{T} należy dokonać transpozycji macierzy przedstawionych powyżej macierzy, uzyskując odpowiednio \tilde{T}^T oraz wektor $(T_T^*)^V$ (opis niezbędnych operacji związanych z transpozycją macierzy opisano w przykł. 4.2):

$$\tilde{T}^T = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0,693 & 1,099 & 1,386 & 1,609 & 1,792 & 1,946 & 2,079 & 2,197 & 2,303 & 2,398 & 2,485 \end{bmatrix};$$

$$(T_{13}^*)^T = [1 \quad 2,565];$$

$$(T_{14}^*)^T = [1 \quad 2,639];$$

$$(T_{15}^*)^T = [1 \quad 2,708].$$

W analogiczny sposób wyznaczamy wektory T_T^* dla $T = 14$ i $T = 15$ i przeprowadzamy ich transpozycję, tworząc wektor \tilde{T}^T dla każdego z analizowanych okresów.

Wyznaczamy iloczyn macierzy $\tilde{T}^T \cdot \tilde{T}$, a następnie uzyskaną macierz odwracamy, uzyskując $(\tilde{T}^T \cdot \tilde{T})^{-1}$ (w obu przypadkach będą to macierze o wymiarach 2x2):

$$\tilde{T}^T \cdot \tilde{T} = \begin{bmatrix} 12,000 & 19,987 \\ 19,987 & 39,575 \end{bmatrix};$$

$$(\tilde{T}^T \cdot \tilde{T})^{-1} = \begin{bmatrix} 0,525 & -0,265 \\ -0,265 & 0,159 \end{bmatrix}.$$

W celu oszacowania bezwzględnego błędu prognoz *ex ante* wyznaczamy iloczyn macierzy $(T_{13}^*)^T$ i $(\tilde{T}^T \cdot \tilde{T})^{-1}$ oraz iloczyn macierzy $(T_{13}^*)^T \cdot (\tilde{T}^T \cdot \tilde{T})^{-1} \cdot X_{13}^*$. Analogicznie postępujemy dla $T = 14$ i $T = 15$. Użytko w ten sposób następujące wyniki obliczeń:

$$(T_{13}^*)^T \cdot (\tilde{T}^T \cdot \tilde{T})^{-1} = [-0,155 \quad 0,143];$$

$$(T_{14}^*)^T \cdot (\tilde{T}^T \cdot \tilde{T})^{-1} = [-0,174 \quad 0,155];$$

$$(T_{15}^*)^T \cdot (\tilde{T}^T \cdot \tilde{T})^{-1} = [-0,193 \quad 0,166];$$

$$(T_{13}^*)^T \cdot (\tilde{T}^T \cdot \tilde{T})^{-1} \cdot T_{13}^* = 0,212;$$

$$(T_{14}^*)^T \cdot (\tilde{T}^T \cdot \tilde{T})^{-1} \cdot T_{14}^* = 0,235;$$

$$(T_{15}^*)^T \cdot (\tilde{T}^T \cdot \tilde{T})^{-1} \cdot T_{13}^* = 0,257.$$

Wiedząc, że odchylenie standardowe reszt z modelu w przypadku modelu z logarytmiczną postacią funkcji trendu wynosi $s = 6,042$, można wyznaczyć bezwzględny błąd prognozy *ex ante* na okres $T = 13$:

$$V_{13} = s \cdot \sqrt{[(T_{13}^*)^T (\tilde{T}^T \tilde{T})^{-1} T_{13}^* + 1]} = 6,042 \cdot \sqrt{0,212 + 1} = 6,51.$$

Znając wartość prognozy (patrz: tab. 5.1.3) oraz bezwzględny błąd prognoz *ex ante*, wyznaczamy względny błąd prognoz *ex ante*:

$$\eta_{13} = \frac{V_{13}}{y_{13}^*} \cdot 100 = \frac{6,51}{87,19} \cdot 100 = 7,47\%.$$

Korzystając z powyższych zależności, można wyznaczyć bezwzględne i względne błędy prognoz *ex ante* na kolejne okresy T .

W celu oszacowania bezwzględnych i względnych błędów prognoz *ex ante* dla modelu analitycznego bazującego na **potęgowej postaci funkcji trendu** dokonujemy oceny wariancji parametru β_i . Macierz wariancji i kowariancji w tej sytuacji wyznacza się według poniższej zależności⁸⁷:

$$D^2(\beta_i) = \tilde{S}^2 \cdot (\tilde{T}^T \cdot \tilde{T})^{-1}.$$

Wiedząc, że $\tilde{S}^2 = 0,008$ oraz korzystając z wcześniej podanych wyników obliczeń pośrednich, wyznaczmy oceny wariancji parametru β_i . Należy zaznaczyć, że wariancję składnika resztowego szacujemy dla zlinearyzowanej postaci funkcji (\tilde{S}^2). Macierz wariancji i kowariancji ma zatem podaną poniżej postać:

$$D^2(\beta_i) = \begin{bmatrix} 0,0043 & -0,0021 \\ -0,0021 & 0,0013 \end{bmatrix}.$$

⁸⁷ W ogólnych zależnościach parametry modelu oznaczone są jako a_i .

Korzystając z zależności podanej w tabeli 4.2, wyznaczamy oceny *ex ante* błędów prognoz przygotowanych z wykorzystaniem modelu bazującego na potęgowej postaci funkcji trendu. Do obliczeń wykorzystujemy adekwatną zależność przedstawioną w tabeli 4.2 oraz wyznaczone wcześniej wektory T_T^* i \tilde{T}^T dla kolejnych okresów T , dla których ma być sporządzona prognoza. Wyniki oszacowania bezwzględnych błędów prognoz *ex ante* przedstawiają się następująco:

$$V_{13} = 8,60;$$

$$V_{14} = 9,59;$$

$$V_{15} = 10,31.$$

Przyjmując na podstawie przeprowadzonych obliczeń wartość prognozy na okres $T = 13$ (patrz: tab. 5.1.8) oraz bezwzględnego błędu prognoz *ex ante*, wyznaczamy względny błąd prognoz *ex ante*:

$$\eta_{13} = \frac{V_{13}}{y_{13}^*} \cdot 100 = \frac{8,60}{87,19} \cdot 100 = 9,86\%.$$

Postępując analogicznie, można wyznaczyć pozostałe wartości ww. błędów prognoz na kolejne okresy.

W celu dokonania ostatecznej oceny dopuszczalności przygotowanych prognoz w tabeli 5.1.10 przedstawiono porównanie ocen *ex ante* średnich błędów predykcji dla trzech analizowanych postaci modeli analitycznych.

Tab. 5.1.10. Oceny *ex ante* średnich błędów predykcji dla analizowanych postaci modeli analitycznych

Rok	T	Liniowa postać funkcji trendu			Logarytmiczna postać funkcji trendu			Potęgowa postać funkcji trendu		
		y_T^*	V_T	η_T	y_T^*	V_T	η_T	y_T^*	V_T	η_T
2016	13	87,19	6,510	7,467%	73,26	6,651	9,079%	89,95	8,60	9,86%
2017	14	93,43	6,737	7,211%	75,46	6,715	8,899%	96,35	9,59	9,95%
2018	15	99,67	6,988	7,011%	77,50	6,773	8,739%	102,72	10,31	10,04%

Źródło: opracowano na podstawie obliczeń własnych.

Wyniki końcowej oceny jakości opracowanych prognoz na podstawie trzech analizowanych postaci funkcji trendu zostały wykorzystane do sformułowania rekomendacji i zaleceń, które mogą być przedstawione odbiorcy prognoz.

VII. Konkluzja

Analizując wartości przedstawione w tabeli 5.1.10, można zaobserwować, że najlepsze oceny *ex ante* średnich błędów predykcji uzyskał model liniowy oraz model bazujący na logarytmicznej postaci funkcji trendu (najniższe wartości błędów). Biorąc pod uwagę okres $T = 13$, można stwierdzić, że błąd *ex ante* prognozy stanowi odpowiednio około 7,47% i 9,08% wielkości prognozy. Korzystając z arbitralnie przyjętych wartości progowych dopuszczalności (patrz: tab. 4.1), uzyskane prognozy wykorzystaniem ww. modeli można uznać za dopuszczalne. Prognozy skonstruowane z wykorzystaniem modelu bazującego na potęgowej postaci funkcji trendu są obarczone największymi błędami. Biorąc pod uwagę wartości błędów względnych, przygotowane prognozy, szczególnie na okres $T = 15$, należy uznać, zgodnie z przyjętymi kryteriami jakościowej klasyfikacji prognoz, za niedopuszczalne (patrz: tab. 4.1).

W modelach liniowych, których parametry oszacowano metodą najmniejszych kwadratów, można także wyznaczyć prognozę przedziałową (patrz: zależność 4.12). W analizowanym przykładzie nie weryfikowano hipotezy o normalności rozkładu reszt modelu, zatem:

$$u = \sqrt{\frac{1}{1-p}} = \sqrt{\frac{1}{1-0,95}} = 4,472136$$

Ponadto z wybranym przez prognozę prawdopodobieństwem $p = 0,95$, zwanym wiarygodnością prognozy, i przyjmując poziom istotności $\alpha = 0,05$ przy $n - k = 10$ (liczba stopni swobody), odczytujemy z tablicy wartości krytycznych dla testu *t*-Studenta wartość $u = 2,228$. Dla modelu liniowego przedział prognozy (dla $T = 13$) ma zatem postać:

$$[87,19 - 4,472136 \cdot 6,51; 1075 + 4,472136 \cdot 6,51].$$

Interpretując tak skonstruowany przedział ufności, można stwierdzić, że liczba gospodarstw domowych wyposażonych w Internet szerokopasmowy w Polsce w 2016 roku (realizacja zmiennej prognozowanej Y) będzie zawierała się w przedziale $[58,1; 116,3]$ z prawdopodobieństwem 0,95.

Zgodnie z przyjętymi zasadami wyniki oceny ilościowej powinny być skonfrontowane z wynikami oceny eksperckiej, uzasadniającej wiarygodność bądź dopuszczalność prognoz. Warto także wziąć pod uwagę znaczenie poszczególnych parametrów modeli analitycznych oraz ich interpretację w kontekście przyjętych założeń i przesłanek prognostycznych.

5.2. Prognoza zainteresowania turystyką online na podstawie liczby zapytań w Google Trends

I. Sformułowanie problemu

Dla branży turystycznej Internet to nie tylko przestrzeń tradycyjnej reklamy. Coraz częściej osoby korzystające z e-commerce w kontekście turystyki doceniają wygodę zakupów online. Google Trends jest źródłem danych na temat częstości zapytań związanych z wyszukiwanym zagadnieniem. Ponieważ zapytania wpisywane w wyszukiwarki nie tylko odzwierciedlają zainteresowania, ale są też użytecznymi wskaźnikami wyprzedzającymi decyzje konsumenckie, prognoza aktywności użytkowników Internetu może dać im wskazówkę, np.: kiedy warto rozpocząć działania marketingowe i jakiego zainteresowania usługami można się spodziewać?

II. Budowa przesłanek prognostycznych

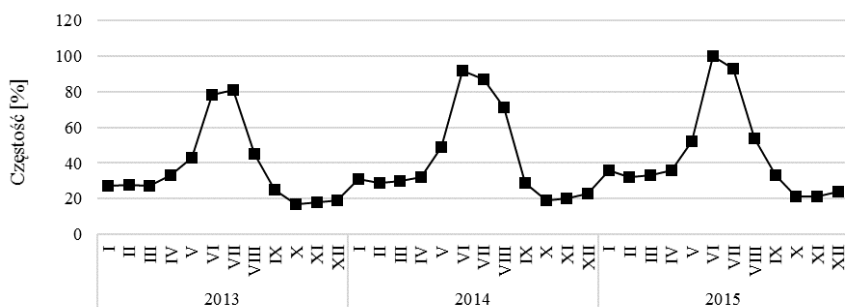
Na podstawie częstości zapytań dla hasła wakacje w okresie od stycznia 2013 r. do grudnia 2015 r. (tab. 5.2.1, wykres 5.2.1) spróbowano określić zainteresowanie wakacjami w wyszukiwarce w roku 2016.

Tab. 5.2.1. Google Trends: zainteresowanie w ujęciu czasowym (styczeń 2013 – grudzień 2015); wyszukiwane hasło: wakacje

t	Rok	Miesiące	y_t	t	Rok	Miesiące	y_t	t	Rok	Miesiące	y_t
1	2013	I	27	13	2014	I	31	25	2015	I	36
2		II	28	14		II	29	26		II	32
3		III	27	15		III	30	27		III	33
4		IV	33	16		IV	32	28		IV	36
5		V	43	17		V	49	29		V	52
6		VI	78	18		VI	92	30		VI	100
7		VII	81	19		VII	87	31		VII	93
8		VIII	45	20		VIII	71	32		VIII	54
9		IX	25	21		IX	29	33		IX	33
10		X	17	22		X	19	34		X	21
11		XI	18	23		XI	20	35		XI	21
12		XII	19	24		XII	23	36		XII	24

Źródło: strona internetowa Google Trends,

<https://www.google.pl/trends/explore#q=wakacje&date=1%2F2013%2037m&cmpt=q&tz=Etc%2FGMT-1>, data wejścia 18.02.2016.



czas

Rys. 5.2.1. Google Trends: zainteresowanie w ujęciu czasowym (styczeń 2013 – grudzień 2015); wyszukiwane hasło: wakacje

Źródło: opracowanie własne na podstawie strony internetowej Google Trends, <https://www.google.pl/trends/explore#q=wakacje&date=1%2F2013%2037m&cmpt=q&tz=Etc%2FGMT-1>, data wejścia 18.02.2016.

Dane Google Trends określają popularność względną, tj. każdy punkt danych (liczba operacji wyszukiwania) jest dzielony przez łączną liczbę wyszukiwań w danym regionie i czasie. Uzyskane wartości są skalowane do przedziału 0-100⁸⁸.

⁸⁸ Strona internetowa Google Trends, https://support.google.com/trends/answer/4365533?hl=pl&ref_topic=4365599, data wejścia 01.03.2016.

III. Wybór metody i reguły prognozowania

Na podstawie wykresu (rys. 5.2.1) dokonano wizualnej dekompozycji analizowanego szeregu czasowego. W szeregu występuje nieznaczny trend rosnący, bardzo wyraźna składowa sezonowa oraz wahania przypadkowe. Założono, że zależność między składowymi jest multiplikatywna, gdyż amplituda wahań rośnie.

IV. Wyznaczenie prognozy

Prognozę przyszłego zainteresowania wakacjami w Internecie za pomocą metody wskaźników sezonowości rozpoczyna się od wyodrębnienia i eliminacji trendu z szeregu. Korzystając z metody najmniejszych kwadratów i przyjmując model regresji liniowej (rozdz. 1), otrzymano następujące równanie funkcji trendu:

$$f^*(t) = 0,186t + 37,89.$$

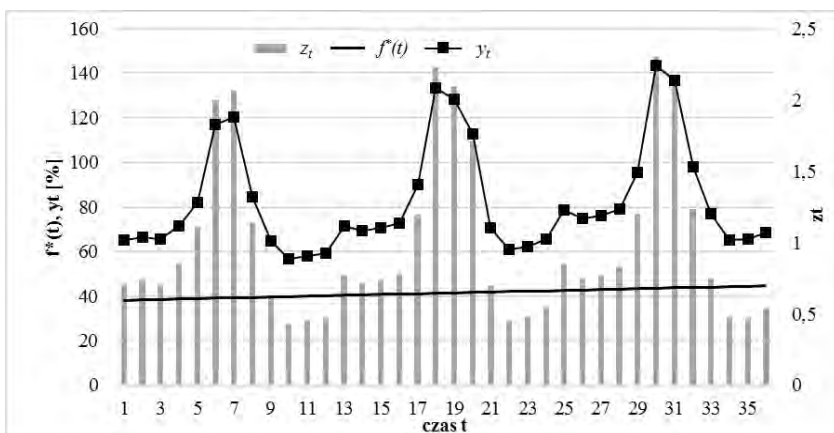
Uwzględniając silne wahania sezonowe w prognozowanym szeregu, porzucono ocenę i interpretację słabego dopasowania modelu trendu do danych. Przy tak nieznacznym trendzie i dużej amplitudzie wahań sezonowych współczynnik kierunkowy z dużym prawdopodobieństwem będzie parametrem nieistotnym modelu regresji. Jednakże w ostatecznym modelu wskaźników sezonowości funkcja trendu zostanie wykorzystana jako czynnik w iloczynie składowych systematycznych, aby uwzględnić zmianę amplitudy wahań, tj. założony multiplikatywny charakter zależności pomiędzy składowymi.

Eliminacja trendu z szeregu w wypadku budowy modelu multiplikatywnego polega na obliczeniu ilorazów wartości rzeczywistych prognozowanej zmiennej – y_t i wartości oszacowanych modelem regresji – $f^*(t)$. Wyniki obliczeń przedstawiono w tabeli 5.2.2 w kolumnie szeregu bez trendu, a zilustrowano na wykresie na rysunku 5.2.2.

Tab. 5.2.2. Google Trends: zainteresowanie w ujęciu czasowym (styczeń 2013 – grudzień 2015); wyszukiwane hasło: wakacje; trend i wskaźniki sezonowości

t	Rok (l)	Mie- siące j	y_t	Model tendencji rozwojowej: $f^*(t) = 0,186t + 37,89$	Szereg bez trendu: $z_t = \frac{y_t}{f^*(t)}$	t_{lj}	Wskaźniki sezonowości $s'_j = \frac{1}{N} \sum_{l=1}^N z_{t_{lj}}$
1	2013 (1)	I	27	0,186·1+37,89=38,26	27/38,26=0,71	1 ₁₁	
2		II	28	0,186·2+37,89=38,45	28/38,45=0,73	2 ₁₂	
3		III	27	0,186·3+37,89=38,63	27/38,63=0,70	3 ₁₃	
4		IV	33	...=38,82	...=0,85	4 ₁₄	
5		V	43	39,01	1,10	5 ₁₅	
6		VI	78	39,19	1,99	6 ₁₆	
7		VII	81	39,38	2,06	7 ₁₇	
8		VIII	45	39,56	1,14	8 ₁₈	
9		IX	25	39,75	0,63	9 ₁₉	
10		X	17	39,94	0,43	10 ₁₁₀	
11		XI	18	40,12	0,45	12 ₁₁₁	
12		XII	19	40,31	0,47	12 ₁₁₂	
13	2014 (2)	I	31	40,49	0,77	13 ₂₁	
14		II	29	40,68	0,71	14 ₂₂	
15		III	30	40,87	0,73	15 ₂₃	
16		IV	32	41,05	0,78	16 ₂₄	
17		V	49	41,24	1,19	17 ₂₅	
18		VI	92	41,42	2,22	18 ₂₆	
19		VII	87	41,61	2,09	19 ₂₇	
20		VIII	71	41,80	1,70	20 ₂₈	
21		IX	29	41,98	0,69	21 ₁₉	
22		X	19	42,17	0,45	22 ₂₁₀	
23		XI	20	42,35	0,47	23 ₂₁₁	
24		XII	23	42,54	0,54	24 ₂₁₂	
25	2015 (3)	I	36	42,73	0,84	25 ₃₁	$s'_5 = 1/3(z_{11}+z_{13}+z_{25}) =$ $= (0,71+0,77+0,84)/3 =$ $= 2,324/3 = 0,775$
26		II	32	42,91	0,75	26 ₃₂	$s'_2 = 1/3(z_{22}+z_{14}+z_{26}) =$ $= (0,73+0,71+0,75)/3 =$ $= 2,197/3 = 0,732$
27		III	33	43,10	0,77	27 ₃₃	$s'_3 = 1/3(z_{33}+z_{15}+z_{27}) =$ $= (0,70+0,73+0,77)/3 =$ $= 2,209/3 = 0,736$
28		IV	36	43,28	0,83	28 ₃₄	$s'_4 = 0,824$
29		V	52	43,47	1,20	29 ₃₅	$s'_5 = 1,168$
30		VI	100	43,66	2,29	30 ₃₆	$s'_6 = 2,177$
31		VII	93	43,84	2,12	31 ₃₇	$s'_7 = 2,099$
32		VIII	54	44,03	1,23	32 ₃₈	$s'_8 = 1,360$
33		IX	33	44,21	0,75	33 ₃₉	$s'_9 = 0,692$
34		X	21	44,40	0,47	34 ₃₁₀	$s'_{10} = 0,452$
35		XI	21	44,59	0,47	35 ₃₁₁	$s'_{11} = 0,466$
36		XII	24	44,77	0,54	36 ₃₁₂	$s'_{12} = 0,518$

Źródło: opracowanie własne.



Rys. 5.2.2. Wykres wartości zmiennej prognozowanej y_t , trendu $f^*(t)$ oraz szeregu bez trendu z_t

Źródło: opracowanie własne.

Kolejny etap budowy modelu prognostycznego polega eliminacji wahań przypadkowych poprzez wyznaczenia tzw. surowych wskaźników sezonowości. W tym celu niezbędne jest dokładne określenie liczby cykli oraz liczby faz w cyklu. Szereg, na podstawie którego tworzona jest prognoza, liczy $n = 36$ obserwacji miesięcznych, tj. liczba cykli, w tym wypadku lat, wynosi 3, a każdy cykl składa się z 12 miesięcy, czyli faz. Inaczej mówiąc, $m = 12$ faz w $N = 3$ cyklach.

Wskaźnik sezonowości obliczono dla każdej fazy, licząc średnią arytmetyczną wartości szeregu po eliminacji trendu tego samego miesiąca (fazy) w kolejnych latach (cyklach).

Obliczenie surowych wskaźników sezonowości przedstawiono w tabeli 5.2.2 w kolumnie wskaźniki sezonowości. Do tabeli wprowadzono dodatkową kolumnę zawierającą numer obserwacji t z dodatkowymi indeksami, określającymi numer cyklu: l – numer roku oraz fazę: j – miesiąc, tj. t_{lj} .

Następnie przeprowadzono sprawdzenie, czy konieczna jest korekta surowych wskaźników poprzez obliczenie sumy obliczonych surowych wskaźników i ocenę, czy otrzymana suma różni się od liczby faz tworzących cykl, tj. 12:

$$\sum_{j=1}^m s'_j = 0,775 + 0,732 + 0,736 + 0,824 + 1,168 + 2,177 + 2,099 + 1,360 + 0,692 + 0,452 + 0,466 + 0,518 = 11,999.$$

Ponieważ nie ma konieczności korekty, można przystąpić do obliczenia wartości modelu oraz prognozy jako iloczynu ekstrapolowanej funkcji trendu oraz odpowiedniego surowego wskaźnika sezonowości:

$$y_{tj}^* = f^*(t_{tj})s_j. \text{ Obliczenia przedstawiono w tabeli 5.2.3.}$$

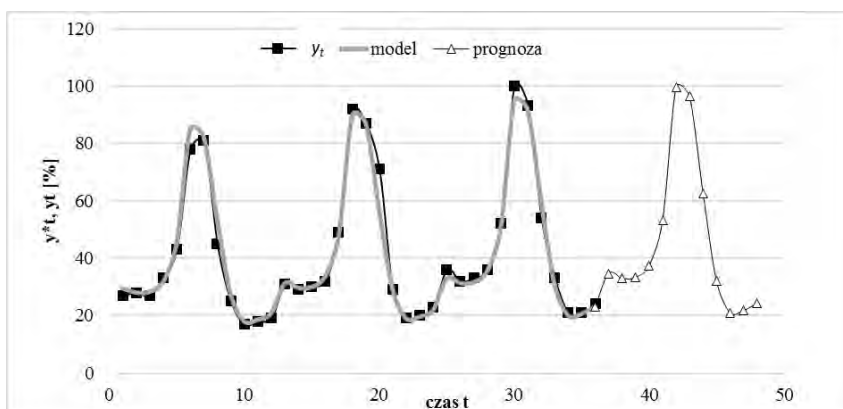
Tab. 5.2.3. Google Trends: zainteresowanie w ujęciu czasowym (styczeń 2013 r. – grudzień 2015 r.); wyszukiwane hasło: wakacje; wartości modelu i prognozy

t	Rok (I)	Miesiące j	y_t	Model tendencji rozwojowej: $f^*(t) = 0,186t + 37,89$	t_{tj}	Wskaźniki sezonowości $s'_j = \frac{1}{N} \sum_{l=1}^N z_{ltj}$	Model i prognoza $y_{tj}^* = f^*(t_{tj})s_j$
1	2013 (1)	I	27	38,26	1 ₁₁	0,775	$f^*(1_{11})s_1 =$ $= 38,26 \cdot 0,775 = 29,50$
2		II	28	38,45	2 ₁₂	0,732	$f^*(2_{12})s_2 =$ $= 38,45 \cdot 0,732 = 28,02$
3		III	27	38,63	3 ₁₃	0,736	$f^*(3_{13})s_3 =$ $= 38,63 \cdot 0,736 = 28,31$
4		IV	33	38,82	4 ₁₄	0,824	$\dots = 31,84$
5		V	43	39,01	5 ₁₅	1,168	45,32
6		VI	78	39,19	6 ₁₆	2,177	84,92
7		VII	81	39,38	7 ₁₇	2,099	82,27
8		VIII	45	39,56	8 ₁₈	1,360	53,56
9		IX	25	39,75	9 ₁₉	0,692	27,37
10		X	17	39,94	10 ₁₀	0,452	17,96
11		XI	18	40,12	12 ₁₁₁	0,466	18,61
12		XII	19	40,31	12 ₁₁₂	0,518	20,80
13	2014 (2)	I	31	40,49	13 ₂₁	0,775	$f^*(13_{21})s_1 =$ $= 40,49 \cdot 0,775 = 31,23$
14		II	29	40,68	14 ₂₂	0,732	$f^*(14_{22})s_2 =$ $= 40,68 \cdot 0,732 = 29,65$
15		III	30	40,87	15 ₂₃	0,736	$\dots = 29,95$
16		IV	32	41,05	16 ₂₄	0,824	33,68
17		V	49	41,24	17 ₂₅	1,168	47,93
18		VI	92	41,42	18 ₂₆	2,177	89,78
19		VII	87	41,61	19 ₂₇	2,099	86,95

t	Rok (l)	Miesiące j	y_t	Model tendencji rozwojowej: $f^*(t) = 0,186t + 37,89$	t_{ij}	Wskaźniki sezonowości $s'_j = \frac{1}{N} \sum_{i=1}^N z_{t_{ij}}$	Model i prognoza $y_{t_{ij}}^* = f^*(t_{ij})s_j$
20	2014 (2)	VIII	71	41,80	20 ₂₈	1,360	56,60
21		IX	29	41,98	21 ₁₉	0,692	28,91
22		X	19	42,17	22 ₂₁₀	0,452	18,96
23		XI	20	42,35	23 ₂₁₁	0,466	19,65
24		XII	23	42,54	24 ₂₁₂	0,518	21,95
25	2015 (3)	I	36	42,73	25 ₃₁	0,775	32,96
26		II	32	42,91	26 ₃₂	0,732	31,29
27		III	33	43,10	27 ₃₃	0,736	31,59
28		IV	36	43,28	28 ₃₄	0,824	35,52
29		V	52	43,47	29 ₃₅	1,168	50,54
30		VI	100	43,66	30 ₃₆	2,177	94,64
31		VII	93	43,84	31 ₃₇	2,099	91,64
32		VIII	54	44,03	32 ₃₈	1,360	59,64
33		IX	33	44,21	33 ₃₉	0,692	30,46
34		X	21	44,40	34 ₃₁₀	0,452	19,97
35		XI	21	44,59	35 ₃₁₁	0,466	20,69
36		XII	24	44,77	36 ₃₁₂	0,518	23,11
37	2016 (4)	I		44,77	37 ₄₁	0,775	$f^*(37_{41})s_1 =$ $=44,77 \cdot 0,775 = 34,69$
38		II		44,96	38 ₄₂	0,732	$f^*(38_{42})s_2 =$ $=44,96 \cdot 0,732 = 32,92$
39		III		45,14	39 ₄₃	0,736	$f^*(39_{43})s_3 =$ $=45,14 \cdot 0,736 = 33,24$
40		IV		45,33	40 ₄₄	0,824	$\dots = 37,36$
41		V		45,52	41 ₄₅	1,168	53,14
42		VI		45,70	42 ₄₆	2,177	99,49
43		VII		45,89	43 ₄₇	2,099	96,32
44		VIII		46,07	44 ₄₈	1,360	62,67
45		IX		46,26	45 ₄₉	0,692	32,00
46		X		46,45	46 ₄₁₀	0,452	20,98
47		XI		46,63	47 ₄₁₁	0,466	21,73
48		XII		46,82	48 ₄₁₂	0,518	24,27

Źródło: opracowanie własne.

Wykres danych rzeczywistych, wartości modelu oraz prognozy na 2016 rok przedstawiono na rysunku 5.2.3.



Rys. 5.2.3. Wartości rzeczywiste, modelu i prognozy Google Trends: zainteresowanie w ujęciu czasowym (styczeń 2013 r. – grudzień 2015 r.); wyszukiwane hasło: wakacje

Źródło: opracowanie własne.

Według prognozy w szczycie sezonu, czyli w czerwcu i w lipcu 2016 roku (tj. $t = 42$ i $t = 43$), można spodziewać odpowiednio o 0,51 punktu procentowego mniejszego i o 3,32 punktu procentowego większego niż w 2015 roku zainteresowania wacjami online w wyszukiwarkach. Jednocześnie można będzie obserwować wzrost wyszukiwania hasła o 1,36 i 1,14 punktu procentowego w miesiącach poprzedzających sezon wakacyjny: w kwietniu ($t = 40$) i maju ($t = 41$). Przeciętnie w całym 2016 roku prognozowana aktywność na tym polu internautów wzrośnie o 1,15 punktu procentowego.

V. Ocena prognozy i weryfikacja metody

Ocena *ex ante* modeli nieliniowych nietransformowanych do liniowych przekracza zakres podręcznika. Generalnie weryfikację otrzymanych prognoz na dany okres wyznaczonych metodą wskaźników sezonowości można przeprowadzić *ex post* w okresie następnym, tj. po ich realizacji (za pomocą miar omawianych w rozdziale „Ocena trafności prognoz”,

w 3. części podręcznika⁸⁹). Jednak w przedstawionym przykładzie w roku 2017 oceny *ex post* prognozy dla 2016 roku wyznaczonej na podstawie wartości z lat 2013-2015 przeprowadzić bezpośrednio nie można. W 2016 roku został zmieniony (poprawiony) system zbierania danych⁹⁰. Warto też podkreślić, że nawet w sytuacji, gdyby były dostępne porównywalne dane, ocenę *ex post* należałoby poprzedzić transformacją danych. Udostępniane w na stronach Google Trends liczby przedstawiają, jak często hasło było wyszukiwane w odniesieniu do najwyższego punktu w danym czasie, czyli wartość 100 oznacza najwyższą popularność hasła. Reasumując, gdyby były dostępne dane dla kolejnych miesięcy roku 2016 w systemie z lat 2013-2015, należałoby najpierw przeskalać dane, a dopiero wówczas korzystać ze wzorów oceny *ex post*. Z uwagi na brak adekwatnych danych przedstawiono ocenę jedynie modelu. Przedstawione miary określają jakość dopasowania do danych oszacowanego modelu nieliniowego. Przedział weryfikacji modelu jest tożsamy z całym zbiorem uczącym (wykorzystanym do budowy modelu). Szczegółowe wyjaśnienie wykorzystanych wzorów oraz ich interpretację zawiera część trzecia podręcznika⁹¹.

⁸⁹ J. Nazarko (red.), *Prognozowanie w zarządzaniu przedsiębiorstwem*. Cz. III..., op. cit.

⁹⁰ Na podstawie informacji na stronie internetowej Google Trends, <https://www.google.pl/trends/explore#q=wakacje&date=1%2F2013%2037m&cmpt=q&tz=Etc%2FGMT-1>, data wejścia 20.07.2017.

⁹¹ J. Nazarko (red.), *Prognozowanie w zarządzaniu przedsiębiorstwem*. Cz. III..., op. cit.

Tab. 5.2.4. Ocena modelu z wykorzystaniem błędów względnych i bezwzględnych

t	y _t	Model	Błąd (bezwzględny)	Moduł błędu	Kwadrat błędu	Procentowy (względny) błąd	Moduł względnego błędu
		y _t [*]	q _t = y _t - y _t [*]	q _t	q _t ²	ψ _t = $\frac{q_t}{y_t} \cdot 100\%$	ψ _t
1	27	29,50	-2,50	2,50	6,26	-9,27%	9,27%
2	28	28,02	-0,02	0,02	0,00	-0,07%	0,07%
3	27	28,31	-1,31	1,31	1,71	-4,84%	4,84%
4	33	31,84	1,16	1,16	1,34	3,51%	3,51%
5	43	45,32	-2,32	2,32	5,40	-5,41%	5,41%
6	78	84,92	-6,92	6,92	47,85	-8,87%	8,87%
7	81	82,27	-1,27	1,27	1,61	-1,56%	1,56%
8	45	53,56	-8,56	8,56	73,35	-19,03%	19,03%
9	25	27,37	-2,37	2,37	5,61	-9,47%	9,47%
10	17	17,96	-0,96	0,96	0,91	-5,63%	5,63%
11	18	18,61	-0,61	0,61	0,37	-3,39%	3,39%
12	19	20,80	-1,80	1,80	3,22	-9,45%	9,45%
13	31	31,23	-0,23	0,23	0,05	-0,75%	0,75%
14	29	29,65	-0,65	0,65	0,43	-2,25%	2,25%
15	30	29,95	0,05	0,05	0,00	0,17%	0,17%
16	32	33,68	-1,68	1,68	2,82	-5,25%	5,25%
17	49	47,93	1,07	1,07	1,14	2,18%	2,18%
18	92	89,78	2,22	2,22	4,95	2,42%	2,42%
19	87	86,95	0,05	0,05	0,00	0,05%	0,05%
20	71	56,60	14,40	14,40	207,35	20,28%	20,28%
21	29	28,91	0,09	0,09	0,01	0,30%	0,30%
22	19	18,96	0,04	0,04	0,00	0,19%	0,19%
23	20	19,65	0,35	0,35	0,12	1,75%	1,75%
24	23	21,95	1,05	1,05	1,10	4,56%	4,56%
25	36	32,96	3,04	3,04	9,24	8,44%	8,44%
26	32	31,29	0,71	0,71	0,51	2,22%	2,22%
27	33	31,59	1,41	1,41	1,98	4,26%	4,26%
28	36	35,52	0,48	0,48	0,23	1,33%	1,33%
29	52	50,54	1,46	1,46	2,14	2,81%	2,81%
30	100	94,64	5,36	5,36	28,78	5,36%	5,36%
31	93	91,64	1,36	1,36	1,86	1,46%	1,46%
32	54	59,64	-5,64	5,64	31,77	-10,44%	10,44%
33	33	30,46	2,54	2,54	6,47	7,71%	7,71%
34	21	19,97	1,03	1,03	1,05	4,89%	4,89%
35	21	20,69	0,31	0,31	0,10	1,48%	1,48%
36	24	23,11	0,89	0,89	0,79	3,71%	3,71%
			Błąd średni w przedziale weryfikacji (1, 36) $\bar{q} = \frac{1}{36} \sum_{t=1}^{36} q_t$	Średnia wartość modułu błędu $ \bar{q} = \frac{1}{36} \sum_{t=1}^{36} q_t $	Średnia wartość kwadratów błędów $\bar{q}^2 = \frac{1}{36} \sum_{t=1}^{36} q_t^2$		Średni procentowy błąd $\bar{\psi}_t = \frac{1}{36} \sum_{t=1}^{36} \psi_t $
			0,06	2,11	12,51		4,85%
					Średnio kwadratowy błąd $s^* = \sqrt{\bar{q}^2}$		Mediana modułów procentowych błędów $Me \psi_t $
					3,54		3,61%

Źródło: opracowanie własne.

Biorąc pod uwagę wartości błędów, można uzyskane dopasowanie modelu do danych uznać za zadowalające. Błąd średni, wynoszący 0,06, jest bliski zeru i wskazuje, że nie występuje ani systematyczne przeszacowanie, ani niedoszacowanie. Średniokwadratowy błąd, równy 3,54, można traktować jako przeciętne odchylenie wartości modelu od wartości rzeczywistych. Otrzymane wartości błędów wyrażone w procentach, z zakresu od -19,03% do 20,28%, ze średnią wartością modułów 4,85%, potwierdzają dobre dopasowanie modelu do danych.

VII. Konkluzja

Zarówno działalność operacyjna, jak też przede wszystkim decyzje strategiczne, inwestycyjne poprzedzone są szeregiem działań związanych z uzyskiwaniem i weryfikacją informacji. Prognoza zainteresowania turystyką online na podstawie popularność danego hasła w procentach pozwala z wyprzedzeniem dostosować ofertę do usług. Otrzymana prognoza byłaby jeszcze bardziej użyteczna, gdyby poszerzono ją o analizę i porównanie popularności różnych haseł oraz ocenę zmiany trendów.

5.3. Prognoza wskaźnika ogólnego klimatu koniunktury w budownictwie

I. Sformułowanie problemu

„Koniunktura” jest pojęciem abstrakcyjnym i nieobserwowalnym bezpośrednio. Cykl koniunkturalny, definiowany jest jako fluktuacje aktywności gospodarczej, o okresach trwania pomiędzy 2 a 8 lat⁹². Publikowane przez GUS badania wskaźników koniunktury gospodarczej, na przykład w przemyśle, budownictwie, handlu czy usługach, realizo-

⁹² *Równoległy oraz wyprzedzający zagregowany wskaźnik koniunktury, zegar koniunktury. Wprowadzenie*, GUS, Departament Studiów Makroekonomicznych I Finansów, http://stat.gov.pl/download/gfx/portalinformacyjny/pl/defaultaktualnosci/5732/3/1/5/zk_wskazniki_i_zegar_koniunktury_wprowadzenie.pdf, data wejścia 23.02.2016.

wane są za pomocą ankiet zawierających pytania o podstawowe wielkości ekonomiczne istotne dla oceny sytuacji przedsiębiorstw⁹³. Wskaźnik ogólnego klimatu koniunktury jest średnią arytmetyczną wskaźników prostych odnoszących się do wybranych pytań i odzwierciedlających bieżącą i przewidywaną ogólną sytuację gospodarczą przedsiębiorstwa. Przyjmuje on wartości od –100 do 100⁹⁴.

W przykładzie zaprezentowano budowę prognozy ogólnego wskaźnika klimatu koniunktury w budownictwie na 2016 rok na podstawie danych historycznych z 3 lat poprzednich. Prognoza wskaźników koniunktury może stanowić narzędzie wspomagające przy wysnuwaniu wniosków na temat przyszłego stanu aktywności gospodarczej nie tylko w sektorze budownictwa, gdyż spowolnienie lub przyspieszenie rozwoju w branży budowlanej ma wpływ również na inne sektory gospodarki.

II. Budowa przesłanek prognostycznych

Kształtowanie się wskaźnika koniunktury w budownictwie w latach 2013-2015 przedstawiono w tabeli 5.3.1 oraz zilustrowano na wykresie na rysunku 5.3.1.

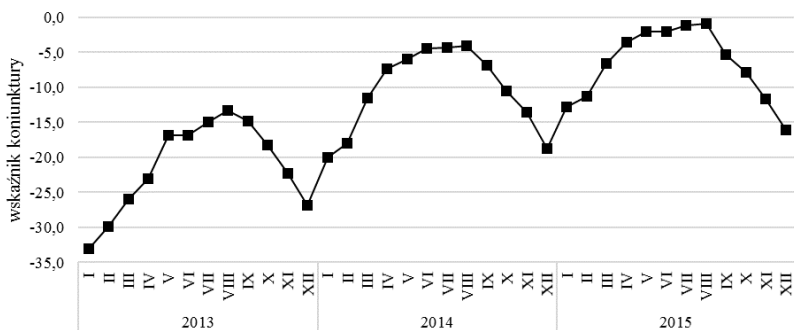
⁹³ Koniunktura w przemyśle, budownictwie, handlu i usługach 2000-2016, strona internetowa GUS, <http://stat.gov.pl/obszary-tematyczne/koniunktura/koniunktura/koniunktura-w-przemysle-budownictwie-handlu-i-uslugach-2000-2016,4,8.html>, data wejścia 23.02.2016.

⁹⁴ Strona internetowa serwisu Bankier.pl, <http://www.bankier.pl/gospodarka/wskazniki-makroekonomiczne/koniunktura-w-budownictwie-pol>, data wejścia 23.02.2016.

Tab. 5.3.1. Wskaźnik ogólnego klimatu koniunktury w budownictwie

t	Rok	Miesiąc	y_t	t	Rok	Miesiąc	y_t	t	Rok	Miesiąc	y_t
1	2013	I	-33,0	13	2014	I	-20,0	25	2015	I	-12,8
2		II	-29,9	14		II	-18,0	26		II	-11,3
3		III	-26,0	15		III	-11,5	27		III	-6,6
4		IV	-23,1	16		IV	-7,3	28		IV	-3,5
5		V	-16,8	17		V	-6,0	29		V	-2,0
6		VI	-16,9	18		VI	-4,5	30		VI	-2,0
7		VII	-14,9	19		VII	-4,3	31		VII	-1,2
8		VIII	-13,3	20		VIII	-4,0	32		VIII	-0,9
9		IX	-14,8	21		IX	-6,9	33		IX	-5,3
10		X	-18,3	22		X	-10,5	34		X	-7,8
11		XI	-22,3	23		XI	-13,5	35		XI	-11,6
12		XII	-26,8	24		XII	-18,8	36		XII	-16,1

Źródło: strona internetowa GUS, <http://stat.gov.pl/download/gfx/portalinformacyjny/pl/defaultaktualnosci/5516/4/8/12/bm-bud.xls>, data wejścia 23.02.2016.



czas

Rys. 5.3.1. Wskaźniki ogólnego klimatu koniunktury

Źródło: opracowanie własne na podstawie strony internetowej GUS, <http://stat.gov.pl/download/gfx/portalinformacyjny/pl/defaultaktualnosci/5516/4/8/12/bm-bud.xls>, data wejścia 23.02.2016.

III. Wybór metody i reguły prognozowania

Wartości wskaźnika ogólnego klimatu koniunktury w budownictwie odzwierciedlają sezonowość w branży budowlanej. Na podstawie oceny wizualnej szeregu przedstawionego na rysunku 5.3.1 do oceny perspektyw rozwoju rynku zaproponowano model analizy harmoniczej.

IV. Wyznaczenie prognozy

Proces opracowywania prognozy na podstawie modelu analizy harmoniczej należy rozpocząć od estymacji i eliminacji tendencji rozwojowej z szeregu. Przyjmując liniową postać modelu trendu i szacując jego parametry metodą najmniejszych kwadratów, otrzymano następującą funkcję:

$$f^*(t) = 0,574t + 23,469.$$

Wartość statystyki F równa 35,10 i na poziomie prawdopodobieństwa $\alpha = 0,05$ przekracza wartość krytyczną wynoszącą 4,13 co pozwala odrzucić hipotezę o braku zależności między zmienną objaśnianą a objaśniającą t . Wartości bezwzględne statystyki t -Studenta wynoszące odpowiednio: 5,92 oraz -11,41 dla współczynnika kierunkowego i wyrazu wolnego są większe od wartości krytycznej równej -1,691 dla $\alpha = 0,05$ i 34 stopni swobody i wskazują, że oba parametry modelu regresji są statystycznie istotne. Jednocześnie niski, mający wartość 0,508 współczynnik determinacji R^2 określa z kolei, że dopasowanie modelu jest słabe i nie powinno się wnioskować o przyszłej wartości współczynnika koniunktury tylko na podstawie modelu trendu.

W kolejnym etapie budowy modelu za pomocą analizy harmoniczej usunięto trend z szeregu poprzez odjęcie od wartości zmiennej prognozowanej wartości obliczonej na podstawie modelu regresji. Następnie przystąpiono do identyfikacji składowej sezonowej. Na podstawie analizy wizualnej szeregu można stwierdzić, że okres wahań sezonowych wynosi 12 miesięcy (częstotliwości $f = 1/12 = 0,08333$).

Ponieważ liczba danych $n = 36$ jest parzysta, można oszacować maksymalnie $q = n/2 = 18$ harmonik. Jednak znana jest długość cyklu wahań i nie ma potrzeby żmudnego obliczania parametrów wszystkich harmonik. Można założyć *a priori* postać modelu i uwzględnić tylko harmonikę odzwierciedlającą dany okres wahań. Pierwsza harmonika w modelu ma okres równy n , druga $n/2$, trzecia $n/3$ itd. Inaczej mówiąc,

do oszacowania modelu wskaźników koniunktury na podstawie danych miesięcznych na pewno potrzebne są parametry trzeciej harmoniki.

Tab. 5.3.2. Wskaźnik ogólnego klimatu koniunktury w budownictwie oraz obliczenia pomocnicze

t	y_t	Trend $f^*(t)$	Szereg bez trendu $z_t = y_t - f^*(t)$	$\sin\left(\frac{2\pi}{n} \cdot it\right)$ $i=3$	$\cos\left(\frac{2\pi}{n} \cdot it\right)$ $i=3$	$(z_t - \bar{z})^2$	$\sin\left(\frac{2\pi}{n} \cdot it\right)$ $i=1$	$\cos\left(\frac{2\pi}{n} \cdot it\right)$ $i=1$
1	-33,0	-22,90	-10,11	0,500	0,866	102,17	0,174	0,985
2	-29,9	-22,32	-7,58	0,866	0,500	57,48	0,342	0,940
3	-26,0	-21,75	-4,25	1,000	0,000	18,11	0,500	0,866
4	-23,1	-21,17	-1,93	0,866	-0,500	3,72	0,643	0,766
5	-16,8	-20,60	3,80	0,500	-0,866	14,41	0,766	0,643
6	-16,9	-20,03	3,13	0,000	-1,000	9,75	0,866	0,500
7	-14,9	-19,45	4,55	-0,500	-0,866	20,69	0,940	0,342
8	-13,3	-18,88	5,58	-0,866	-0,500	31,07	0,985	0,174
9	-14,8	-18,30	3,50	-1,000	0,000	12,25	1,000	0,000
10	-18,3	-17,73	-0,57	-0,866	0,500	0,33	0,985	-0,174
11	-22,3	-17,16	-5,15	-0,500	0,866	26,50	0,940	-0,342
12	-26,8	-16,58	-10,22	0,000	1,000	104,48	0,866	-0,500
13	-20,0	-16,01	-3,99	0,500	0,866	15,97	0,766	-0,643
14	-18,0	-15,43	-2,57	0,866	0,500	6,60	0,643	-0,766
15	-11,5	-14,86	3,36	1,000	0,000	11,26	0,500	-0,866
16	-7,3	-14,29	6,99	0,866	-0,500	48,75	0,342	-0,940
17	-6,0	-13,71	7,71	0,500	-0,866	59,42	0,174	-0,985
18	-4,5	-13,14	8,64	0,000	-1,000	74,55	0,000	-1,000
19	-4,3	-12,56	8,26	-0,500	-0,866	68,23	-0,174	-0,985
20	-4,0	-11,99	7,99	-0,866	-0,500	63,78	-0,342	-0,940
21	-6,9	-11,42	4,52	-1,000	0,000	20,36	-0,500	-0,866
22	-10,5	-10,84	0,34	-0,866	0,500	0,11	-0,643	-0,766
23	-13,5	-10,27	-3,23	-0,500	0,866	10,47	-0,766	-0,643
24	-18,8	-9,69	-9,11	0,000	1,000	82,99	-0,866	-0,500
25	-12,8	-9,12	-3,68	0,500	0,866	13,57	-0,940	-0,342
26	-11,3	-8,55	-2,76	0,866	0,500	7,61	-0,985	-0,174
27	-6,6	-7,97	1,37	1,000	0,000	1,87	-1,000	0,000
28	-3,5	-7,40	3,90	0,866	-0,500	15,16	-0,985	0,174
29	-2,0	-6,82	4,82	0,500	-0,866	23,23	-0,940	0,342
30	-2,0	-6,25	4,25	0,000	-1,000	18,03	-0,866	0,500
31	-1,2	-5,68	4,48	-0,500	-0,866	20,00	-0,766	0,643
32	-0,9	-5,10	4,20	-0,866	-0,500	17,63	-0,643	0,766
33	-5,3	-4,53	-0,77	-1,000	0,000	0,60	-0,500	0,866
34	-7,8	-3,95	-3,85	-0,866	0,500	14,82	-0,342	0,940
35	-11,6	-3,38	-8,22	-0,500	0,866	67,63	-0,174	0,985
36	-16,1	-2,81	-13,30	0,000	1,000	176,83	0,000	1,000

Źródło: opracowanie własne.

W celu oszacowania parametrów trzeciej harmoniki obliczono wartości $\sin\left(\frac{2\pi}{n}it\right)$ i $\cos\left(\frac{2\pi}{n}it\right)$ dla $i = 3, n = 36$ i $t = 1, 2, \dots, 36$. Wyniki obliczeń zamieszczono w tabeli 5.3.2 wraz z wartościami szeregu bez trendu z_t . Następnie policzono sumy iloczynów zgodnie ze wzorami:

$$a_3 = \frac{2}{36} \sum_{t=1}^{36} z_t \sin\left(\frac{2\pi}{36}3t\right) \text{ oraz } b_3 = \frac{2}{36} \sum_{t=1}^{36} z_t \cos\left(\frac{2\pi}{36}3t\right).$$

Otrzymano następującą wartość parametrów: $a_3 = -1,284$ oraz $b_3 = -7,187$. Wykonując obliczenia w arkuszu kalkulacyjnym Microsoft Excel, warto wykorzystać funkcję `SUMA.ILOCZYNÓW()`, której parametrami do obliczenia parametru a są odpowiednio: wartości szeregu bez trendu oraz wartości funkcji $\sin\left(\frac{2\pi}{n}it\right)$, zaś do obliczenia parametru b także wartości szeregu bez trendu oraz $\cos\left(\frac{2\pi}{n}it\right)$. W celu otrzymania wartości a oraz b wynik działania funkcji `SUMA.ILOCZYNÓW()` należy tylko pomnożyć przez $\frac{2}{n}$ tj. $\frac{2}{36}$.

Zanim przystąpiono do budowy modelu, obliczono udział wariancji harmoniki trzeciej w wyjaśnieniu wariancji rozpatrywanej zmiennej prognozowanej s^2 , wykorzystując wzór $\frac{a_3^2 + b_3^2}{2s^2}$. Oszacowanie całkowitej wariancji zmiennej prognozowanej wymaga obliczenia średniej zróżnicowanego szeregu oraz sumy kwadratów odchyłeń od obliczonej średniej, tj. $s^2 = \frac{1}{n} \sum_{t=1}^n (z_t - \bar{z})^2$, gdzie $\bar{z} = \frac{1}{n} \sum_{t=1}^n z_t$. Średnie różnicowanych szeregów powinny oscylować wokół 0, w tym przypadku otrzymano wartość $\bar{z} = 0,0028$ i tę wykorzystano do obliczeń. Suma kwadratów odchyłeń wynosi 1240,453, a iloraz tej wartości i liczby obserwacji daje wariancję równą 34,457. Natomiast suma kwadratów parametrów trzeciej harmoniki, tj. $a_3^2 + b_3^2$ to 53,306, po podzieleniu przez dwukrotność wariancji całego szeregu daje 0,774. Wartość ta określa jaką część wariancji zmiennej prognozowanej jest objaśniana przez trzecią harmonikę. Nie jest ona zbyt wysoka, dlatego też obliczono kolejną harmonikę, tym

razem przyjmując 3-letni okres wahań ($f = \frac{1}{36}$) czyli harmonikę pierwszą, i analogicznie powtórzono obliczenia, zaczynając od oszacowania wartości $\sin\left(\frac{2\pi}{n}it\right)$ i $\cos\left(\frac{2\pi}{n}it\right)$ dla $i = 1, n = 36$ i $t = 1, 2, \dots, 36$. Wyniki obliczeń zamieszczono także w tabeli 5.3.1. Suma iloczynów wartości szeregu bez trendu i odpowiadającej jej wartości funkcji sinus wynosi: -8,900, zaś suma iloczynów wartości szeregu bez trendu i odpowiadającej jej funkcji cosinus wynosi: -6,483. Po pomnożeniu otrzymanych wartości przez $\frac{2}{36}$ otrzymano $a_1 = -0,494$ oraz $b_1 = -3,138$. Udział pierwszej harmoniki, czyli suma kwadratów parametrów a_1 i b_1 w dwukrotności wariancji całego szeregu, to: 0,146. Łącznie obie harmoniki wyjaśniają $0,774 + 0,146 = 0,920$ zmienności szeregu bez trendu. Przyjmując, że jest to procent wystarczający, przystąpiono do budowy modelu prognostycznego składającego się z modelu trendu oraz sumy wybranych harmonik:

$$\begin{aligned}
 y_t^* &= f^*(t) + a_3 \sin\left(\frac{2\pi}{36}3t\right) + b_3 \cos\left(\frac{2\pi}{36}3t\right) + \\
 &+ a_1 \sin\left(\frac{2\pi}{36}t\right) + b_1 \sin\left(\frac{2\pi}{36}t\right) = \\
 &= 0,574t - 23,469 - 1,284\sin\left(\frac{2\pi}{36}3t\right) - 7,187\cos\left(\frac{2\pi}{36}3t\right) - \\
 &- 0,494\sin\left(\frac{2\pi}{36}t\right) - 3,138\cos\left(\frac{2\pi}{36}t\right).
 \end{aligned}$$

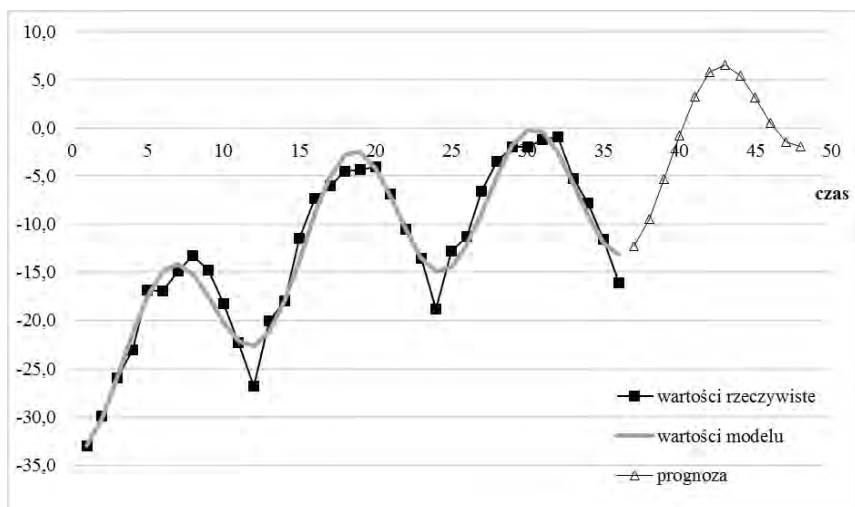
W prognozowaniu na kolejne okresy wykorzystano ekstrapolację zbudowanego modelu na przedział czasu, dla którego wyznaczana jest prognoza. Wynik obliczeń przedstawiono w tabeli 5.3.3. i zilustrowano na wykresie na rysunku 5.3.2.

Zgodnie z prognozą, jeżeli utrzyma się doczasowy trend, już w maju 2016 roku wskaźnik koniunktury w branży budowlanej przyjmie wartość dodatnią. Jednakże składnik sezonowy sprawi, że ocena koniunktury w listopadzie i grudniu 2016 roku spadnie, choć nieznacznie, poniżej zera.

Tab. 5.3.3. Wartości rzeczywiste, modelu i prognozy wskaźnika ogólnego klimatu koniunktury w budownictwie

t	y_t	Trend $f^*(t)$	$\sin\left(\frac{2\pi}{n} \cdot it\right)$ $i=3$	$\cos\left(\frac{2\pi}{n} \cdot it\right)$ $i=3$	$\sin\left(\frac{2\pi}{n} \cdot it\right)$ $i=1$	$\cos\left(\frac{2\pi}{n} \cdot it\right)$ $i=1$	y_t^* $a_3 = -1,284$ $b_3 = -7,187$ $a_1 = -0,494$ $b_1 = -3,138$
1	-33,0	-22,90	0,500	0,866	0,174	0,985	-22,90-1,284·0,5- -7,187·0,866- -0,494·0,174- -3,138·0,985= =-32,942
2	-29,9	-22,32	0,866	0,500	0,342	0,940	-22,32-1,284·0,866- -7,187·0,5- -0,494·0,342- -3,138·0,940= =-30,144
3	-26,0	-21,75	1,000	0,000	0,500	0,866	...=-25,996
4	-23,1	-21,17	0,866	-0,500	0,643	0,766	-21,413
5	-16,8	-20,60	0,500	-0,866	0,766	0,643	-17,412
6	-16,9	-20,03	0,000	-1,000	0,866	0,500	-14,835
7	-14,9	-19,45	-0,500	-0,866	0,940	0,342	-14,123
8	-13,3	-18,88	-0,866	-0,500	0,985	0,174	-15,203
9	-14,8	-18,30	-1,000	0,000	1,000	0,000	-17,514
10	-18,3	-17,73	-0,866	0,500	0,985	-0,174	-20,153
11	-22,3	-17,16	-0,500	0,866	0,940	-0,342	-22,129
12	-26,8	-16,58	0,000	1,000	0,866	-0,500	-22,628
13	-20,0	-16,01	0,500	0,866	0,766	-0,643	-21,235
14	-18,0	-15,43	0,866	0,500	0,643	-0,766	-18,053
15	-11,5	-14,86	1,000	0,000	0,500	-0,866	-13,673
16	-7,3	-14,29	0,866	-0,500	0,342	-0,940	-9,024
17	-6,0	-13,71	0,500	-0,866	0,174	-0,985	-5,124
18	-4,5	-13,14	0,000	-1,000	0,000	-1,000	-2,812
19	-4,3	-12,56	-0,500	-0,866	-0,174	-0,985	-2,521
20	-4,0	-11,99	-0,866	-0,500	-0,342	-0,940	-4,166
21	-6,9	-11,42	-1,000	0,000	-0,500	-0,866	-7,166
22	-10,5	-10,84	-0,866	0,500	-0,643	-0,766	-10,601
23	-13,5	-10,27	-0,500	0,866	-0,766	-0,643	-13,454
24	-18,8	-9,69	0,000	1,000	-0,866	-0,500	-14,883
25	-12,8	-9,12	0,500	0,866	-0,940	-0,342	-14,447
26	-11,3	-8,55	0,866	0,500	-0,985	-0,174	-12,219
27	-6,6	-7,97	1,000	0,000	-1,000	0,000	-8,760
28	-3,5	-7,40	0,866	-0,500	-0,985	0,174	-4,973
29	-2,0	-6,82	0,500	-0,866	-0,940	0,342	-1,849
30	-2,0	-6,25	0,000	-1,000	-0,866	0,500	-0,202
31	-1,2	-5,68	-0,500	-0,866	-0,766	0,643	-0,447
32	-0,9	-5,10	-0,866	-0,500	-0,643	0,766	-2,481
33	-5,3	-4,53	-1,000	0,000	-0,500	0,866	-5,713
34	-7,8	-3,95	-0,866	0,500	-0,342	0,940	-9,214
35	-11,6	-3,38	-0,500	0,866	-0,174	0,985	-11,966
36	-16,1	-2,81	0,000	1,000	0,000	1,000	-13,130
37		-2,23	0,50	0,87	0,174	0,985	-12,273
38		-1,66	0,87	0,50	0,342	0,940	-9,480
39		-1,08	1,00	0,00	0,500	0,866	-5,332
40		-0,51	0,87	-0,50	0,643	0,766	-0,749
41		0,06	0,50	-0,87	0,766	0,643	3,252
42		0,64	0,00	-1,00	0,866	0,500	5,829
43		1,21	-0,50	-0,87	0,940	0,342	6,541
44		1,79	-0,87	-0,50	0,985	0,174	5,461
45		2,36	-1,00	0,00	1,000	0,000	3,150
46		2,94	-0,87	0,50	0,985	-0,174	0,511
47		3,51	-0,50	0,87	0,940	-0,342	-1,465
48		4,08	0,00	1,00	0,866	-0,500	-1,964

Źródło: opracowanie własne.



Rys. 5.3.2. Wartości rzeczywiste, modelu i prognozy wskaźnika ogólnego klimatu koniunktury w budownictwie

Źródło: opracowanie własne.

VI. Ocena prognozy i weryfikacja metody

Dla modelu nieliniowego analizy harmonicznej, opracowanego dla wskaźnika ogólnego koniunktury w budownictwie, przeprowadzono ocenę *ex post* jakości wyznaczonych prognoz. W tabeli 5.3.4 przedstawiono wartości prognozowane za pomocą modelu oraz opublikowane na stronie GUS dla roku 2016, wraz z wartościami obliczonych różnic. Zastosowano wzory przedstawione w trzeciej części podręcznika⁹⁵.

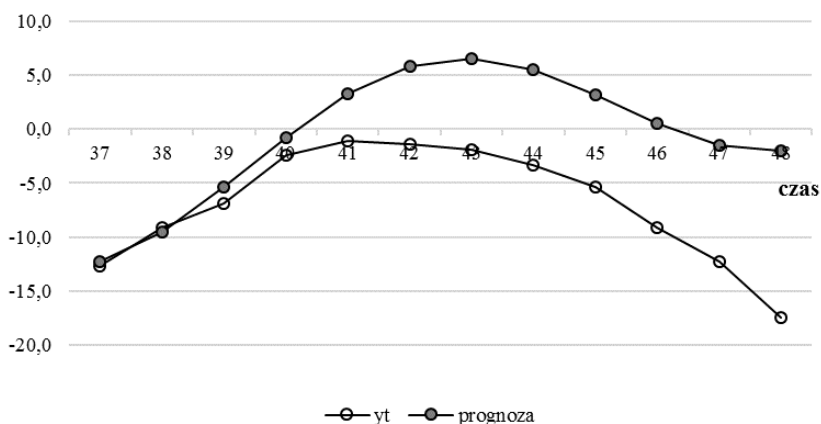
Analiza wyznaczonych miar zamieszczonych w tabeli 5.3.4 nie pozwala na pozytywną ocenę jakości otrzymanych prognoz. Model zdecydowanie przeszacował wartości (błąd średni jest ujemny). O niesymetrycznym rozkładzie błędów świadczy też duża różnica pomiędzy średnią a medianą procentowych błędów. Przeciętny procentowy błąd wynoszący 180% jest nieakceptowalny. Na podstawie oszacowanego modelu poprawne byłoby tylko krótkookresowe prognozowanie. Zauważalne od połowy roku rosnące rozbieżności powinny skłonić do ponownego opracowania modelu i uwzględnienia zmiany trendu koniunktury poprzez włączenie do jego budowy najnowszych dostępnych danych.

⁹⁵ J. Nazarko (red.), *Prognozowanie w zarządzaniu przedsiębiorstwem*. Cz. III..., op. cit.

Tab. 5.3.4. Wartości rzeczywiste oraz prognozowane wskaźnika ogólnego klimatu koniunktury w budownictwie wraz z wartościami błędów

t	y_t	Prognoza	Błąd (bezwzględny)	Moduł błędu	Kwadrat błędu	Procentowy (względny) błąd	Moduł względnego błędu
		y_t^*	$q_t = y_t - y_t^*$	$ q_t $	q_t^2	$\psi_t = \frac{q_t}{y_t} \cdot 100\%$	$ \psi_t $
37	-12,7	-12,273	-0,43	0,43	0,18	3,36%	3,36%
38	-9,1	-9,48	0,38	0,38	0,14	-4,18%	4,18%
39	-6,9	-5,332	-1,57	1,57	2,46	22,72%	22,72%
40	-2,4	-0,749	-1,65	1,65	2,73	68,79%	68,79%
41	-1,1	3,252	-4,35	4,35	18,94	395,64%	395,64%
42	-1,4	5,829	-7,23	7,23	52,26	516,36%	516,36%
43	-1,9	6,541	-8,44	8,44	71,25	444,26%	444,26%
44	-3,3	5,461	-8,76	8,76	76,76	265,48%	265,48%
45	-5,4	3,15	-8,55	8,55	73,10	158,33%	158,33%
46	-9,1	0,511	-9,61	9,61	92,37	105,62%	105,62%
47	-12,3	-1,465	-10,84	10,84	117,40	88,09%	88,09%
48	-17,4	-1,964	-15,44	15,44	238,27	88,71%	88,71%
			Błąd średni w przedziale weryfikacji $\bar{q} = \frac{1}{12} \sum_{37}^{48} q_t$	Średnia wartość modułu błędu $ \bar{q} = \frac{1}{36} \sum_{37}^{48} q_t $	Średnia wartość kwadratów błędów $\bar{q}^2 = \frac{1}{12} \sum_{37}^{48} q_t^2$		Średni procentowy błąd $\bar{\psi}_t = \frac{1}{12} \sum_{37}^{48} \psi_t $
			-6,37	6,44	62,15		180%
					Średnio-kwadratowy błąd $s^* = \sqrt{\bar{q}^2}$		Mediana modułów procentowych błędów $Me \psi_t $
					7,88		97,16%

Źródło: opracowanie własne na podstawie strony internetowej GUS, <http://stat.gov.pl/download/gfx/portalinformacyjny/pl/defaultaktualnosci/5516/4/8/12/bm-bud.xlsx>, data wejścia 20.07.2017.



Rys. 5.3.3. Wartości rzeczywiste oraz prognozy wskaźnika ogólnego klimatu koniunktury w budownictwie

Źródło: opracowanie własne.

Na rysunku 5.3.3 można zaobserwować systematyczny wzrost rozbieżności pomiędzy prognozą a wartościami obserwowanymi. Począwszy od okresu 41 (maja 2016 rok), model zdecydowanie przeszacowuje wartości. Rozbieżność ta osiąga maksimum w ostatnich miesiącach roku.

VII. Konkluzja

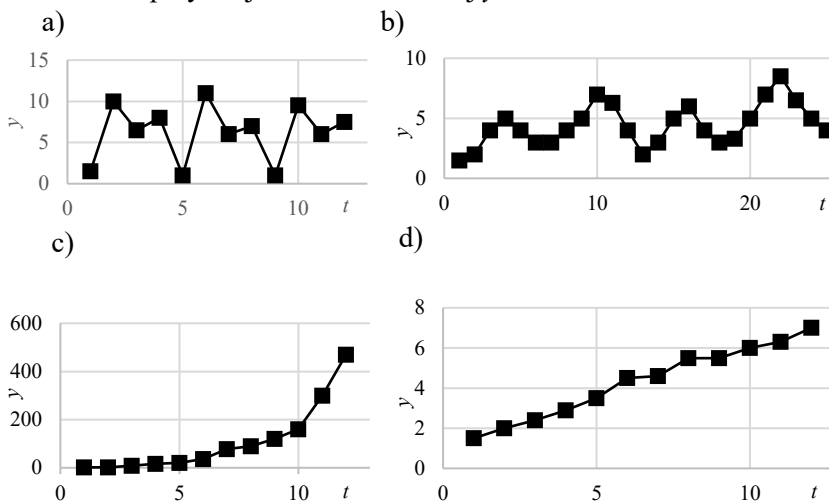
Prognoza wskaźników koniunktury jest jednym z narzędzi statystycznych służących do obserwacji i oceny zjawisk ekonomicznych, wspomagających wnioskowanie na temat przyszłego stanu aktywności gospodarczej. Należy pamiętać jednak, że analiza prognozy stanu gospodarki wymaga uwzględnienia z jednej strony losowości zjawisk gospodarczych a z drugiej także ewentualnych zmian strukturalnych. Warto odnieść się do szerszego zbioru zmiennych reprezentujących stan gospodarki, a także wykorzystać inne metody analizy.

Inny przykład wykorzystania metody analizy harmonicznej i metody wskaźników sezonowości do prognozowania danych charakteryzujących się wyraźnymi wahaniami sezonowymi wraz z oceną dopasowa-

nia zbudowanych modeli oraz oceną *ex post* jakości prognoz można odnaleźć w publikacji: E. Chodakowska, M. Komuda. Metody predykcji ekonomicznej na podstawie szeregu czasowego z wahaniami sezonowym, rozdz. w: L. Kiełtyka, J. Nazarko (red. nauk.), *Metody i procesy usprawniania zarządzania przedsiębiorstwem: wybrane zagadnienia*, Wydawnictwo Menadżerskie PTM, Warszawa 2006, s. 28-37. W artykule porównano prognozy miesięcznego zapotrzebowania na energię ciepłą w jednej z miejskich kotłowni województwa podlaskiego, uzyskane różnymi metodami w celu stworzenia planu grzewczego.

6. PROBLEMY DO SAMODZIELNEGO ROZWIĄZANIA

1. Na podstawie wizualnej oceny szeregów przedstawionych na rysunku 6.1 określ, jakie znane Ci metody prognozowania można zastosować do określenia przyszłej wartości zmiennej y .



Rys. 6.1. Szeregi czasowe

Źródło: opracowanie własne.

- Określ liczbę faz i cykli w szeregu przedstawionym na rysunku 6.1a.
- Zaproponuj metody, które mogą być wykorzystywane do prognozowania danych przedstawionych na rysunku 6.1c. Uzasadnij swój wybór.
- Wielkość sprzedaży produktu X w kolejnych miesiącach 2014 i 2015 roku przedstawiono tabeli 6.1.

Tab. 6.1. Szereg czasowy i jego wskaźniki sezonowości

Miesiące	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII	I	II	III
Rok	2014												2015		
Wielkość sprzedaży [tys. sztuk]	38	115	156	183	216	241	246	273	280	298	312	327	334	342	353

Źródło: opracowanie własne.

Wyznacz prognozę dla kolejnych czterech miesięcy, stosując następujące modele analityczne:

- A – liniowy,
- B – logarytmiczny,
- C – potęgowy.

5. Wyznacz i zinterpretuj miary oceny jakości modeli zbudowanych w zadaniu 4.
6. Wykonaj wykres danych rzeczywistych i prognoz wyznaczonych z wykorzystaniem trzech modeli zbudowanych w zadaniu 4.
7. Wskaż, spośród trzech modeli zbudowanych w zadaniu 4, model najlepiej wyrażający szereg czasowy analizowanej zmiennej. Uzasadnij swój wybór.
8. Dopasuj wartości wskaźników sezonowości do poszczególnych faz szeregu danych kwartalnych przedstawionego w tabeli 6.2.

Tab. 6.2. Szereg czasowy i jego wskaźniki sezonowości

t	1	2	3	4	5	6	7	8	9	10	11	12
Kwartały	I	II	III	IV	I	II	III	IV	I	II	III	IV
y	1,5	10,0	6,5	8,0	1,0	11,0	6,0	7,0	1,0	9,5	6,0	7,5
Wskaźnik sezonowości	10,1	7,5	6,2	1,2								

Źródło: opracowanie własne.

9. Na podstawie danych z tabeli 6.2 wyznacz prognozę wartość zmiennej y dla $t = 13, \dots, 16$ metodą wskaźników sezonowości.
10. W szeregu dwuletnich danych codziennych występuje sezonowość tygodniowa oraz miesięczna. Harmoniki o jakich częstotliwościach powinny być uwzględnione w modelu analizy harmoniczej?

LITERATURA

1. Aczel A.D., *Statystyka w zarządzaniu*, Wydawnictwo Naukowe PWN, Warszawa 2000.
2. Bates J.M., Granger C.J.W., The combination of forecasts, *Operational Research Quarterly* nr 20, 1969, s. 451-468.
3. Box G.M., Jenkins G.E.P., *Analiza szeregów czasowych. Prognozowanie i sterowanie*, Państwowe Wydawnictwo Naukowe, Warszawa 1983.
4. Bronsztejn I.N., Siemiendajew K.A., *Matematyka. Poradnik encyklopedyczny*, Wydawnictwo Naukowe PWN, Warszawa 1995.
5. Cieślak M. (red.), *Prognozowanie gospodarcze. Metody i zastosowanie*, Wydawnictwo Naukowe PWN, Warszawa 2011.
6. Clemen R.T., Combining forecasts: A review and annotated bibliography, *International Journal of Forecasting* nr 5, North-Holland 1989, s. 559-583.
7. Chodakowska E., Komuda M., Metody predykcji ekonomicznej na podstawie szeregu czasowego z wahaniami sezonowym, rozdz.: L. Kiełtyka, J. Nazarko (red. nauk.), *Metody i procesy usprawniania zarządzania przedsiębiorstwem: wybrane zagadnienia* Wydawnictwo Menadżerskie PTM, Warszawa 2006, s. 28-37.
8. DeLurgio S.A., *Forecasting principles and applications*, Irwin/McGraw-Hill, Boston 1998.
9. Dittman P., *Prognozowanie w przedsiębiorstwie. Metody i ich zastosowanie*, Wolters Kluwer, Kraków 2009.
10. Dittman P., *Metody prognozowania sprzedaży w przedsiębiorstwie*, Wydawnictwo Akademii Ekonomicznej im. Oskara Langego we Wrocławiu, Wrocław 2000.
11. Dym H., McKean H.P., *Fourier Series and Integrals (Probability and Mathematical Statistics)*, ACADEMIC PressINC, 1985.
12. Elliott G., Averaging and the optimal combination of forecasts, *Working Paper University of California*, San Diego 2011, s. 1-30.
13. Featherstone K., James D., Assessment of NZIER Forecast Accuracy, *New Zealand Institute of Economic Research, Research Working Paper 94/9*, Wellington 1994.

14. Fourier J.B.J., *Théorie analytique de la chaleur (The Analytical Theory of Heat)*, Firmin Didot, Paris 1822.
15. Graefe A., Armstrong J.S., Jones Jr. R.J., Cuzan A.G., Combining forecasts: An application to elections, *International Journal of Forecasting*, volume 30, Issue 1, 2014, s. 43–54.
16. Gruszczynski M., Podgórska M., *Ekometria*, Szkoła Główna Handlowa, Warszawa 1996.
17. Halicka K., Godlewska J., *Prognozowanie średniego miesięcznego kursu kupna USD*, *Ekonomia i Zarządzanie*, t. 5, nr 1, Oficyna Wydawnicza Politechniki Białostockiej, Białystok 2013, s. 19-29.
18. Halicka K., Wieńkowski C., *Wykorzystanie metod wygładzania wykładniczego do prognozowania kursu sprzedaży EUR*, *Ekonomia i Zarządzanie*, t. 5, nr 2, Oficyna Wydawnicza Politechniki Białostockiej, Białystok 2013, s. 70-80.
19. Hilas C.S., Goudos S.K., Sahalos J.N., Seasonal decomposition and forecasting of telecommunication data: A comparative case study, *Technological Forecasting & Social Change* 73, 2006, s. 495–509.
20. Larose D.T., *Metody i modele eksploracji danych*, Wydawnictwo Naukowe PWN, Warszawa 2016.
21. Maciąg A., Pietroń R., Kukła S., *Prognozowanie i symulacje w przedsiębiorstwie*, Polskie Wydawnictwo Ekonomiczne, Warszawa 2013.
22. Nazarko J. (red.), *Prognozowanie w zarządzaniu przedsiębiorstwem. Cz. I. Wprowadzenie do metodyki prognozowania*, Wydawnictwo Politechniki Białostockiej, Białystok 2004.
23. Nazarko J. (red.), *Prognozowanie w zarządzaniu przedsiębiorstwem. Cz. II. Prognozowanie na podstawie szeregów czasowych*, Wydawnictwo Politechniki Białostockiej, Białystok 2004.
24. Nazarko J. (red.), *Prognozowanie w zarządzaniu przedsiębiorstwem. Cz. III. Prognozowanie na podstawie modeli adaptacyjnych*, Wydawnictwo Politechniki Białostockiej, Białystok 2005.
25. Nowak E., *Zarys metod ekonometrii. Zbiór zadań*, Wydawnictwo Naukowe PWN, Warszawa 2002.
26. Orwat A., Przykład zastosowania metody odpornej w modelowaniu finansowych szeregów czasowych. *Zeszyty Naukowe Szkoły Głównej Gospodarstwa Wiejskiego. Ekonomika i Organizacja Gospodarki Żywnościowej* 60, 2006, s. 279-288.

27. Ostasiewicz S., Wykorzystanie metod dyskryminacyjnych w prognozowaniu dyskretnym, Prace Naukowe Akademii Ekonomicznej we Wrocławiu nr 232, seria: Monografie i opracowania nr 14, 1983.
28. Pickover C.A., *The Math Book: From Pythagoras to the 57th Dimension, 250 Milestones in the History of Mathematics*, Sterling Publishing Company, Inc., 2009.
29. Radzikowska B. (red.), *Metody prognozowania: zbiór zadań*, Wydawnictwo Akademii Ekonomicznej im. Oskara Langego we Wrocławiu, Wrocław 2000.
30. Sobczyk M., *Prognozowanie*, Placet, Warszawa 2008.
31. Stanisław T., *Funkcje jednej zmiennej w badaniach ekonomicznych*, Wydawnictwo Naukowe PWN, Warszawa 1993.
32. Welfe A., *Ekonometria*, Wydawnictwo PWE, Warszawa 1995.
33. Yang Y., Combining forecasting procedures: some theoretical results, *Econometric Theory*, vol. 1, 2004, s. 176-222.
34. Zeliaś A., *Teoria prognozy*, PWE, Warszawa 1997.
35. Zeliaś A., Pawełek B., Wanat S., *Prognozowanie ekonomiczne: teoria, przykłady, zadania*, Wydawnictwo Naukowe PWN, Warszawa 2004.
36. Zong-Chang Y., Modeling and forecasting monthly movement of annual average solar insolation based on the least-squares Fourier-model, *Energy Conversion and Management*, vol. 81, May 2014, s. 201-210.

Oddawany do rąk Czytelników podręcznik „Prognozowanie w zarządzaniu przedsiębiorstwem” przekazuje wiedzę, zarówno podstawową, jak i zaawansowaną, z zakresu metod i narzędzi prognozowania gospodarczego. Podręcznik składa się z dziewięciu odrębnych części, które obejmują:

- wprowadzenie do metodyki prognozowania,
- podstawy prognozowania szeregów czasowych,
- zastosowanie modeli adaptacyjnych,
- zastosowanie modeli trendu,
- wykorzystanie modeli ARIMA do prognozowania gospodarczego,
- prognozowanie na podstawie modeli ekonometrycznych,
- zastosowanie w prognozowaniu metod sztucznej inteligencji,
- prognozowanie z zastosowaniem metod heurystycznych,
- wykorzystanie różnych metod prognozowania w zarządzaniu przedsiębiorstwem.

Każda część kończy się pytaniami i zadaniami, których samodzielne rozwiązanie pomoże Czytelnikowi ugruntować zdobytą wiedzę.

Autorami IV części podręcznika są pracownicy Wydziału Inżynierii Zarządzania Politechniki Białostockiej: prof. dr hab. inż. Joanicjusz Nazarko, dr hab. inż. Katarzyna Halicka, dr inż. Ewa Chodakowska i dr inż. Arkadiusz Jurczuk.

W podręczniku skoncentrowano się na metodach i technikach prognozowania, które są szczególnie przydatne w zarządzaniu przedsiębiorstwem. Autorzy, dokonując wyboru materiału i jego zakresu merytorycznego, oparli się na swoim kilkuletnim doświadczeniu w nauczaniu prognozowania na studiach magisterskich, podyplomowych i MBA oraz na doświadczeniach zdobytych w uczelniach zagranicznych. Dzięki stałej współpracy z gospodarką zawarli w swojej pracy wiedzę nie tylko teoretyczną, lecz także praktyczną.

Na podkreślenie zasługuje fakt, że podręcznik nie tylko dostarcza podstawowej wiedzy teoretycznej, ale – poprzez przystępnie prowadzony wykład i ilustrację przykładami empirycznymi wszystkich omawianych zagadnień – umożliwia Czytelnikowi bezpośrednie zastosowanie opisywanych metod do rozwiązywania zadań z praktyki gospodarczej.

prof. dr hab. Dorota Witkowska
Katedra Finansów i Strategii Przedsiębiorstwa
Uniwersytet Łódzki

(...) recenzowany podręcznik to wartościowa publikacja o charakterze dydaktycznym, która może służyć jako podręcznik podstawowego kursu prognozowania dla studentów politechnik i innych szkół wyższych. Przystępny język podręcznika, ścisłość i przejrzystość wywodów wspartych licznymi przykładami obliczeń oraz wykresami sprawia, że może on być polecony także pracownikom instytucji i organizacji zajmujących się problematyką modelowania i prognozowania na podstawie szeregów czasowych.

Z przyjemnością rekomenduję podręcznik „Prognozowanie w zarządzaniu przedsiębiorstwem. Część IV. Prognozowanie na podstawie modeli trendu” jako wartościową publikację z zakresu modelowania i prognozowania szeregów czasowych.

dr hab. Stanisław Łobjeiko, prof. SGH
Kolegium Nauk o Przedsiębiorstwie
Szkoła Główna Handlowa