

Joanna Chraołowska\*, Joanicjusz Nazarko\*\*

## MODELE ARIMA W PROGNOZOWANIU SPRZEDAŻY\*\*\*

*W artykule przedstawiono metodykę budowy modeli ARIMA oraz ich wykorzystanie do prognozowania jednowymiarowych szeregów czasowych. Wykorzystano jedno z ogólnie stosowanych podejść zaproponowane przez Boxa i Jenkinsa. Opisano i przedyskutowano kolejne etapy tworzenia modelu na przykładzie danych dotyczących przedsiębiorstwa handlowego typu cash & carry oraz przedsiębiorstwa produkcyjnego.*

### ARIMA MODELS IN FORECASTING SALES GROSSINCOME

*The paper presents construction methodology of ARIMA models and their application in one-dimensional time series forecasting. The Box and Jenkins approach, being one of the widely used, has been employed.*

*Consecutive phases of the model constructing have been described and discussed on the basis of a cash & carry type of trade as well as productive enterprise.*

## 1. WPROWADZENIE

Wśród wielu prognoz szczególną rolę w zarządzaniu przedsiębiorstwem odgrywa prognoza sprzedaży. Jest ona punktem wyjścia do przewidywania większości pozycji kształtujących poziom wyniku finansowego w firmie. Prognoza sprzedaży jest w szczególności podstawą do tworzenia planu biznesowego, alokacji środków i określenia wielkości zapasów. Wykorzystywana jest również do określenia przyszłych potrzeb finansowych przedsiębiorstwa.

Konstrukcja prognozy sprzedaży, tak jak i innych rodzajów prognoz, jest procesem wieloetapowym. Przebiega według pewnego, ogólnego schematu postępowania prognostycznego. Do prognozowania sprzedaży stosuje się zarówno metody ilościowe, jak i heurystyczne. Zastosowanie konkretnych wynika ze specyfiki rynku i przyjętych przesłanek prognostycznych. Podstawowymi kryteriami, które brane są pod uwagę przy wyborze metody, jest oczekiwana trafność prognozy, koszt jej opracowania i wdrożenia oraz łatwość zastosowania [1].

---

\* Politechnika Białostocka; joanna@piasta.pl " Politechnika Białostocka; jnazarko@cksr.ac.bialystok.pl ♦♦ Artykuł powstał w ramach prac badawczych nr W/WZ/7/03

W artykule przedstawiono metodykę budowy oraz przykłady wykorzystania modeli ARIMA (*Autoregressive Integrated Moving Average*) do prognozowania jednowymiarowych szeregów czasowych. Wyznaczono prognozą przychodów ze sprzedaży na przykładzie przedsiębiorstwa handlowego typu cash & carry oraz przedsiębiorstwa produkcyjnego.

Wykorzystano jedno z ogólnie stosowanych podejść do budowy modeli ARIMA, zaproponowane przez Boxa i Jenkinsa.

Narzędziem informatycznym wykorzystanym przez autorów, w procesie opracowania modelu i wyznaczania prognoz był moduł *Szeregi czasowe i prognozowanie* programu STATISTICA PL wersji 6.0.

## 2. MODELE ARIMA W ANALIZIE SZEREGÓW CZASOWYCH

Modele ARIMA, czyli scałkowane procesy autoregresji i średniej ruchomej, są bardzo ogólną klasą modeli szeregów czasowych. Ich budowa oparta jest na zjawisku autokorelacji. Mogą być one wykorzystywane do modelowania stacjonarnych szeregów czasowych lub szeregów niestacjonarnych sprowadzalnych do stacjonarnych.

Wyróżnia się trzy podstawowe rodzaje modeli tej klasy:

- 1) modele autoregresji (AR),
- 2) modele średniej ruchomej (MA),
- 3) modele mieszane autoregresji i średniej ruchomej (ARMA).

Symbol I użyty w nazwie modelu wskazuje, że szereg czasowy poddany został operacji różnicowania.

Do zapisu postaci modelu ARIMA używa się uniwersalnej notacji ARIMA  $(p, d, q) \times (P, D, Q)_s$ , literami  $p, P$  - oznacza się rząd autoregresji;  $d, I$  - rząd różnicowania;  $q, Q$  - rząd średniej ruchomej, odpowiednio dla składowej niesezonowej i sezonowej.

Model ARIMA można zapisać również w postaci operatorowej:

$$\phi(B)\Phi(B^S)(1-B^S)^D Y_t = c + \theta(B)\Theta(B^S)e_t, \quad t = 1, 2, \dots, n \quad (1)$$

przy czym:

$$\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p \quad (1.1)$$

$$\Phi(B) = 1 - \Phi_1 B^S - \Phi_2 B^{2S} - \dots - \Phi_p B^{pS} \quad (1.2)$$

$$\theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q \quad (1.3)$$

$$\Theta(B^S) = 1 - \Theta_1 B^S - \Theta_2 B^{2S} - \dots - \Theta_Q B^{QS} \quad (1.4)$$

gdzie:

- $\phi, \Phi$  – parametry modelu autoregresji,
- $B$  – operator przesunięcia wstecz,
- $s$  – liczba sezonów w okresie,
- $c$  – stała,
- $\theta, \Theta$  – parametry modelu średniej ruchomej,
- $e_t$  – składnik losowy (reszta modelu) modelu w okresie lub momencie  $t$ .

Proces budowy modelu można ująć w trzy zasadnicze fazy [5]:

1. identyfikację,
2. estymację i testy,
3. zastosowanie.

W fazie pierwszej następuje wstępna identyfikacja cech analizowanego szeregu czasowego. Podejmowana jest decyzja o potrzebie transformacji danych w celu stabilizacji wariancji i/lub o różnicowaniu szeregu w celu stabilizacji średniej. Dla szeregu stacjonarnego dokonuje się identyfikacji jego potencjalnych modeli poprzez badanie funkcji autokorelacji (ACF) i autokorelacji cząstkowej (PACF).

W fazie drugiej szacowane są parametry wybranych modeli. Ostateczny wybór modelu następuje zazwyczaj na podstawie analizy kilku kryteriów, tj.: istotności parametrów modelu, błędu średniokwadratowego, kryterium informacyjnego (np.: AIC - *Akaike's Information Criterion*, BIC - *Bayesian Information Criterion* lub FPE - *Final Prediction Error*).

Wybrany model poddawany jest następnie sprawdzeniu diagnostycznemu. Podstawą diagnostyki jest analiza własności szeregu reszt modelu. Jeżeli reszty modelu są procesem białego szumu, nie występują znaczące wartości funkcji ACF lub PACF szeregu reszt modelu, model może zostać zastosowany do prognozowania. W przeciwnym wypadku należy wrócić do fazy drugiej i wybrać inny model lub do fazy pierwszej, dokonując ponownej identyfikacji modelu.

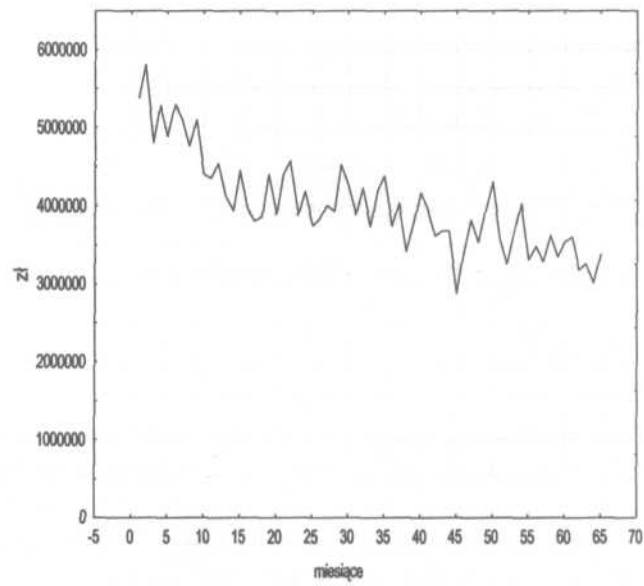
W fazie trzeciej następuje wykorzystanie modelu do sporządzenia prognozy, czyli jego implementacja.

### 3. PRZYKŁADY ZASTOSOWAŃ MODELI ARIMA W PROGNOZOWANIU PRZYCHODÓW ZE SPRZEDAŻY

#### 3.1. CHARAKTERYSTYKA OBIEKTÓW PROGNOZY

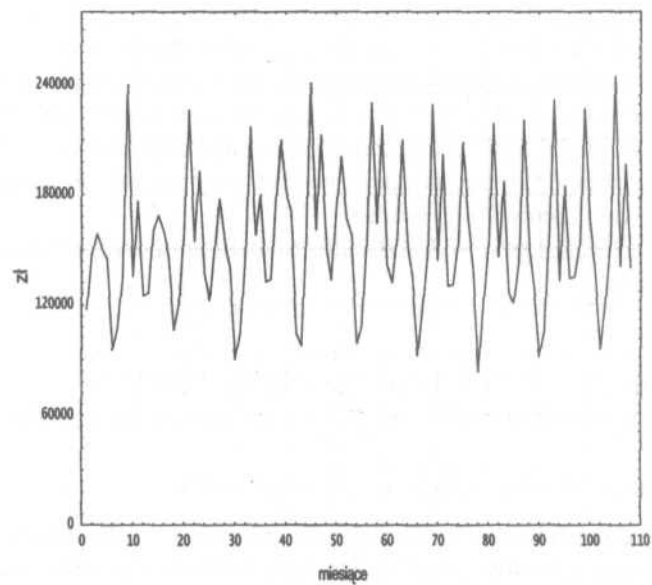
W artykule dokonano analizy szeregów czasowych przychodów ze sprzedaży dla dwóch firm: przedsiębiorstwa handlowego typu cash & carry oraz przedsiębiorstwa produkcyjnego.

Miesięczne przychody ze sprzedaży produktów w badanych przedsiębiorstwach przedstawiono w formie wykresów na rysunkach 1 i 2.



Rys. 1. Miesięczne przychody ze sprzedaży w przedsiębiorstwie handlowym w okresie: styczeń 1998 – czerwiec 2003

Źródło: dane wewnętrzne przedsiębiorstwa



Rys. 2. Miesięczne przychody ze sprzedaży w przedsiębiorstwie produkcyjnym w okresie: styczeń 1994 – grudzień 2002

Źródło: dane wewnętrzne przedsiębiorstwa

Przedsiębiorstwo handlowe jest filią sieci hurtowni spożywczych. Charakteryzuje je znaczne zróżnicowanie asortymentowe (ok. 7000 artykułów) i duża liczba odbiorców, którymi są głównie podmioty gospodarcze (ok. 5000). Ogólny popyt na artykuły oferowane przez firmę nie podlega znacznym wahaniom.

Przedsiębiorstwo produkcyjne jest jednym z zakładów włókienniczych, wytwarzających rocznie około 1000 jednostek zróżnicowanego wyrobu. Klientami firmy są głównie hurtownie i sklepy patronackie. Sprzedaż produktów wytwarzanych w przedsiębiorstwie podlega dosyć silnym wahaniom sezonowym, szczególnie w miesiącach letnich. Firma odnotowuje wówczas znacząco niższe przychody ze sprzedaży oferowanych wyrobów.

Celem budowanej przez autorów prognozy było stworzenie podstaw do określenia miesięcznych i kwartalnych prognoz przychodów ze sprzedaży dla każdej z firm. Ze względu na dostępne dane do wyznaczenia prognozy zaproponowano użycie modeli szeregów czasowych.

### 3.2. WYZNACZENIE PROGNOZY ZE SPRZEDAŻY NA PRZYKŁADZIE PRZEDSIĘBIORSTWA HANDLOWEGO TYPU CASH & CARRY

Dostępny szereg czasowy liczy 68 miesięcznych obserwacji, z tego do budowy modelu przyjęto 65 obserwacji, natomiast trzy ostatnie pozostawiono w celu porównania z prognozami uzyskanymi w efekcie zastosowania modelu.

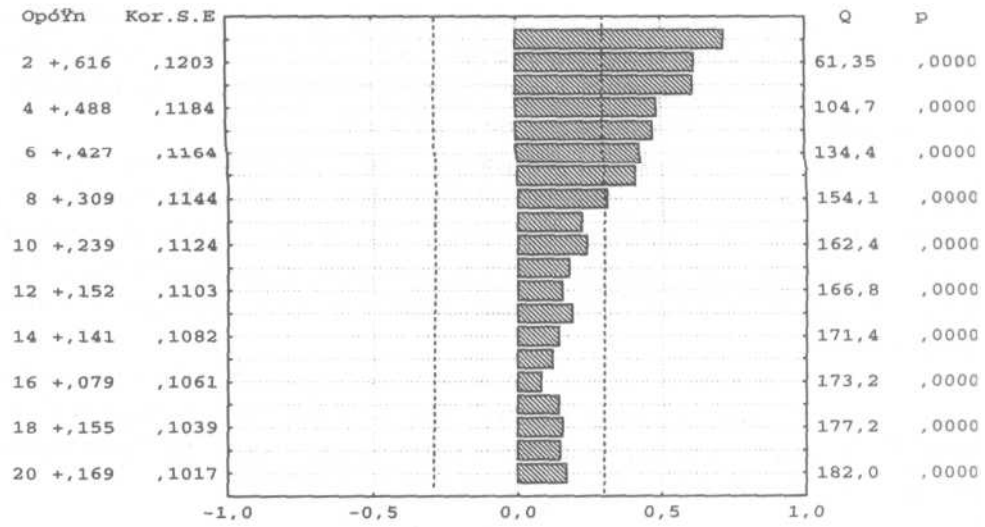
Ze wstępnej oceny wykresu (rys. 1) oraz z charakteru zjawiska można wnioskować, iż w szeregu występuje trend oraz wahania. Zauważalne są spadki i wzrosty przychodów ze sprzedaży na przestrzeni kolejnych miesięcy.

Sporządzony wykres szeregu sugeruje, iż analizowany szereg jest niestacjonarny. Powyższe przypuszczenie zostało zweryfikowane za pomocą analizy korelogramu, która jest najczęściej stosowaną metodą identyfikacji niestacjonarności szeregu. Wyznaczone dla analizowanego szeregu wartości współczynników autokorelacji przedstawiono na rysunku 3.

Wartości współczynników autokorelacji maleją powoli. Dla przyjętego poziomu istotności  $\alpha = 0,05$  statystycznie istotnie różne od zera okazały się współczynniki autokorelacji dla opóźnień od 1 do 7 miesięcy. Potwierdza to wcześniejsze przypuszczenie o niestacjonarności szeregu czasowego miesięcznych przychodów ze sprzedaży.

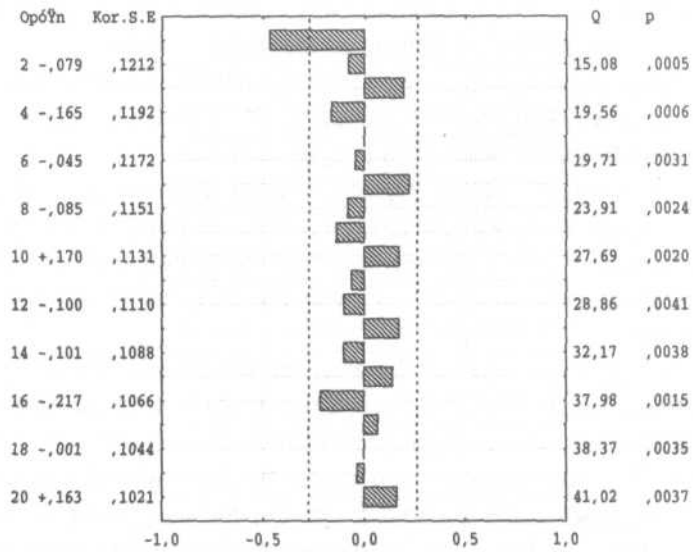
W celu eliminacji niestacjonarności szereg poddano operacji różnicowania. Wyznaczone dla szeregu jednokrotnie zróżnicowanego wartości współczynników autokorelacji i autokorelacji cząstkowej przedstawiono na rysunkach 4 i 5.

Wykresy współczynników autokorelacji (rys. 4, 5) wskazują, iż szereg uzyskany w wyniku różnicowania jest szeregiem stacjonarnym.



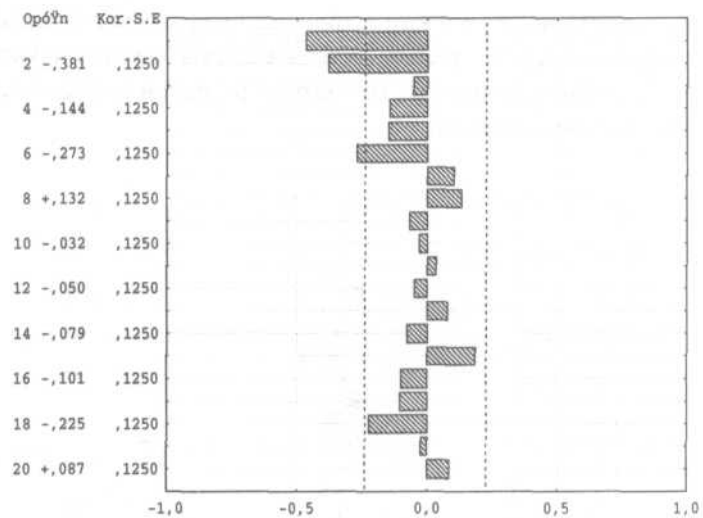
Rys. 3. Współczynniki autokorelacji szeregu czasowego miesięcznych przychodów ze sprzedaży przedsiębiorstwa handlowego

Źródło: obliczenia własne wykonane za pomocą programu Statistica PL



Rys. 4. ACF szeregu po operacji różnicowania

Źródło: obliczenia własne wykonane za pomocą programu Statistica PL



**Rys. 5.** PACF szeregu po operacji różnicowania Źródło: obliczenia własne wykonane za pomocą programu Statistica PL

Pojedyncza znacząca wartość funkcji ACF oraz wykładnicze zanikanie wartości współczynników autokorelacji cząstkowej sugerują, iż modelem adekwatnym do budowy prognoz będzie model ARIMA (0,1,1).

Model ma postać

$$(1 - B) \hat{y}_t = (1 - 0,726) \hat{\varepsilon}_t + \hat{\varepsilon}_t \quad (2)$$

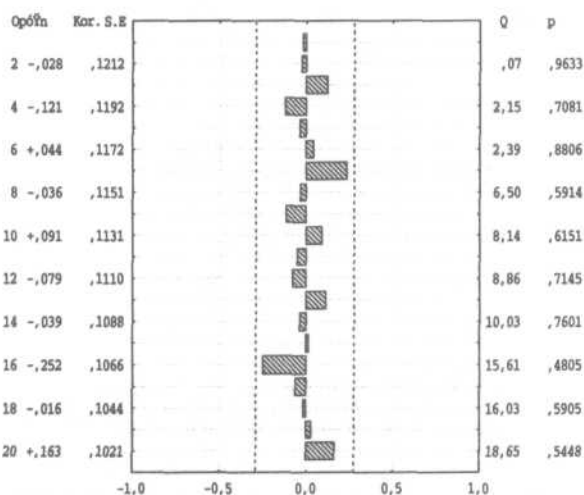
Zwykle dany szereg czasowy można opisać różnymi modelami ARIMA. Autorzy przeprowadzili analizę alternatywnych modeli do MA (1). Jako kryterium wyboru najlepszego modelu przyjęto istotność parametrów modelu oraz najmniejszą wartość średniego bezwzględnego błędu procentowego (MAPE).

Na podstawie przeprowadzonej analizy do budowy prognozy przychodów ze sprzedaży przyjęty został model ARIMA (0,1,1), charakteryzujący się najlepszymi wskaźnikami oceny.

Estymacji parametrów modelu dokonano za pomocą przybliżonej metody największej wiarygodności według McLeoda i Salesa [4]. Stałą  $c$ , oszacowana na poziomie (-33 189,73), można interpretować jako współczynnik kierunkowy trendu szeregu niezróżnicowanego. Oszacowane parametry modelu spełniają warunki zapewniające stabilność modelu.

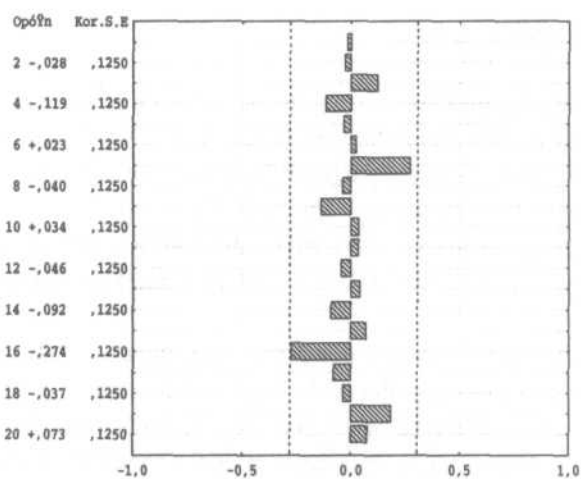
Model (2) jest równoważny modelowi prostego wygładzania wykładniczego, ze stałą wygładzania  $\alpha$  równą (1 - 0,726).

Estymowany model poddano sprawdzeniu diagnostycznemu. Przedstawione na rysunkach 6 i 7 wykresy przebiegu funkcji ACF i PACF reszt modelu wskazują na brak autokorelacji reszt, co pozwala przyjąć, że reszty są procesem białego szumu i nie są istotne w procesie budowy prognozy.



Rys. 6. Wykres funkcji ACF reszt modelu ARIMA (0,1,1)

Źródło: obliczenia własne



Rys. 7. Wykres funkcji PACF reszt modelu ARIMA (0,1,1)

Źródło: obliczenia własne



Opracowany model uznano jako adekwatny do wyznaczenia prognoz przychodów ze sprzedaży na kolejne miesiące. Rezygnując z zapisu operatorowego oraz po przeindeksowaniu uzyskuje się postać dogodną do obliczania prognoz

$$y_{t+1}^* = y_t - 0,726(y_t - y_t^*) - 33189,73 \quad (3)$$

gdzie  $y_{t+1}^*$  - prognoza wartości zmiennej prognozowanej na okres  $t+1$ .

Korzystając z modelu (3) wyznaczono prognozy na jeden miesiąc oraz dwa i trzy miesiące. Uzyskane wyniki porównano z rzeczywistą wartością przychodów. Wyniki obliczeń zestawiono w tabeli 1.

Tabela 1  
Prognoza przychodów ze sprzedaży wyznaczona modelem ARIMA (0,1,1)

Nr kolejnego miesiąca	Dane obserwowane	Prognoza przychodów ze sprzedaży		Błąd procentowy
66	3 567 258,12	$y_{65+1}^*$	3 414 065,73	4,3%
67	3 359 277,46	$y_{65+2}^*$	3 447 255,46	-2,6%
68	3 652 917,99	$y_{65+3}^*$	3 480 445,19	4,7%

Źródło: obliczenia własne

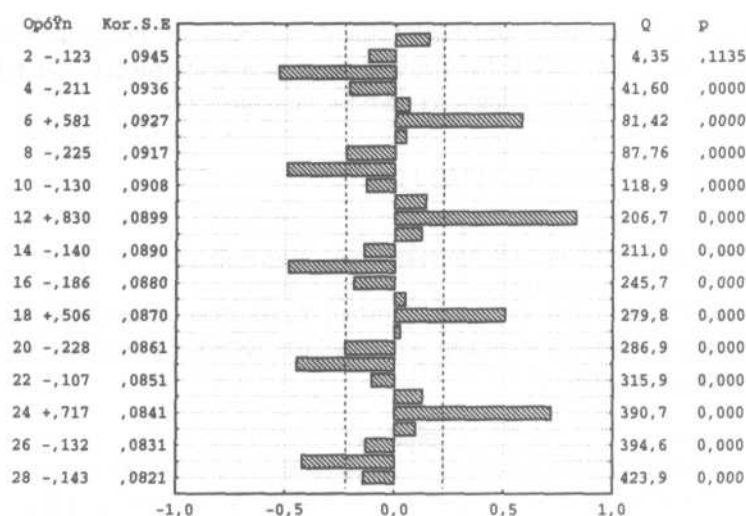
Otrzymane wyniki są zadowalające z praktycznego punktu widzenia. Względny błąd procentowy nie przekroczył 5%.

### 3.3. WYZNACZENIE PROGNOZY ZE SPRZEDAŻY NA PRZYKŁADZIE PRZEDSIĘBIORSTWA PRODUKCYJNEGO

Dostępny szereg czasowy przychodów ze sprzedaży liczy 111 miesięcznych obserwacji, z tego do budowy modelu przyjęto 108 obserwacji, trzy ostatnie pozostawiono w celu porównania z prognozami uzyskanymi w efekcie zastosowania modelu.

Ze wstępnej oceny wykresu (rys. 2) oraz z charakteru zjawiska można wnioskować, iż w szeregu występuje stały poziom oraz wahania. Zauważalne są spadki i wzrosty przychodów ze sprzedaży. Wahania te mają skłonność do powtarzania się w określonych odstępach.

Sporządzony wykres szeregu (rys. 2) sugeruje, iż analizowany szereg jest niestacjonarny. Powyższe przypuszczenie zostało zweryfikowane za pomocą analizy korelogramu. Wyznaczone dla analizowanego szeregu wartości współczynników autokorelacji przedstawiono na rysunku 8.



Rys. 8. Współczynniki autokorelacji szeregu czasowego miesięcznych przychodów ze sprzedaży przedsiębiorstwa produkcyjnego

Źródło: obliczenia własne wykonane za pomocą programu Statistica PL

Przebieg funkcji ACF wyraźnie wskazuje na obecność składowej sezonowej. Najwyższe wartości ACF obserwowane są dla opóźnień będących wielokrotnością 12 miesięcy.

W celu eliminacji niestacjonarności szereg poddano operacji różnicowania sezonowego, o okresie 12. Wyznaczone dla szeregu zróżnicowanego wartości współczynników autokorelacji i autokorelacji cząstkowej przedstawiono na rysunkach 9 i 10.

Przebieg funkcji ACF (rys. 9) wykazuje wykładnicze zanikanie początkowych wartości współczynników autokorelacji, co jest charakterystyczne dla modelu autoregresji, na wykresie PACF (rys. 10) widoczna jest jedna istotna wartość współczynnika. Można zauważyć również znaczące wartości ACF dla opóźnienia 12, oraz widoczne na wykresie PACF, zanikanie wartości współczynników autokorelacji cząstkowej dla opóźnień będących wielokrotnością 12.

Przy takim obrazie przebiegu funkcji ACF i PACF można zaproponować model  $ARIMA(1,0,0)(1,1,0)_{12}$

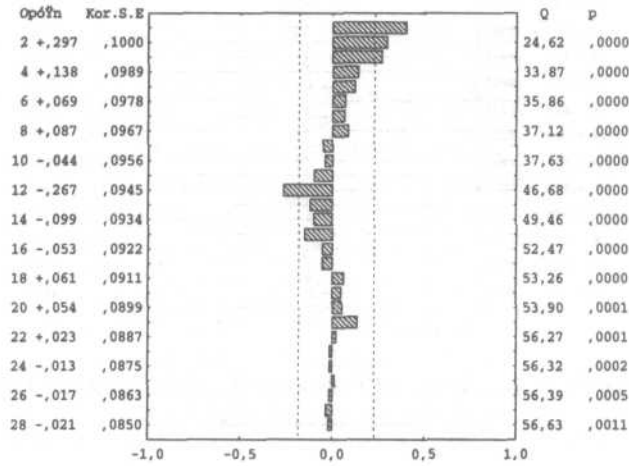
Model ma postać

$$(1 - 0,409 B)(1 + 0,284 B^{12})(1 - B^{12})Y_t = e_t \quad (4)$$

Autorzy przeprowadzili analizę alternatywnych modeli. Jako kryterium wyboru najlepszego modelu przyjęto: istotność parametrów modelu oraz najmniejszą wartość średniego bezwzględnego błędu procentowego (MAPE).

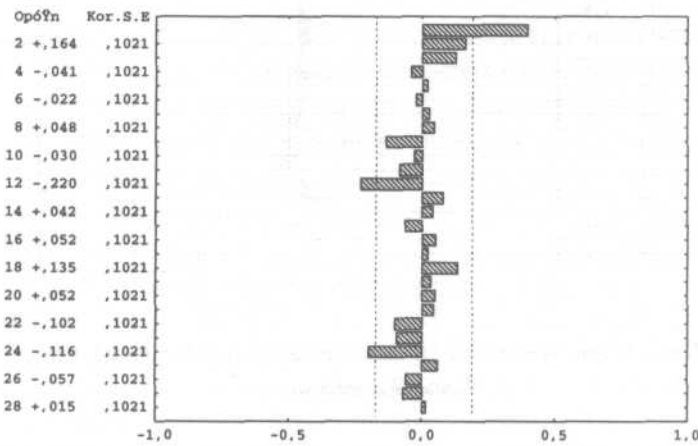
Na podstawie przeprowadzonej analizy do budowy prognozy przychodów ze sprzedaży przyjęty został model wybrany pierwotnie, charakteryzujący się najlepszymi wskaźnikami oceny.

Oszacowane parametry modelu spełniają warunki zapewniające stabilność modelu. Estymowany model poddano sprawdzeniu diagnostycznemu.



Rys. 9. ACF szeregu po operacji różnicowania

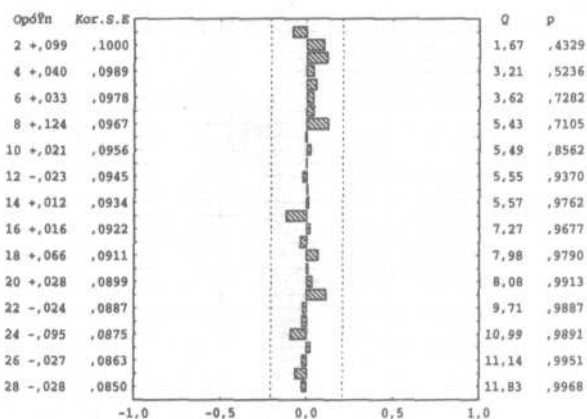
Źródło: obliczenia własne wykonane za pomocą programu Statistica PL



Rys. 10. PACF szeregu po operacji różnicowania

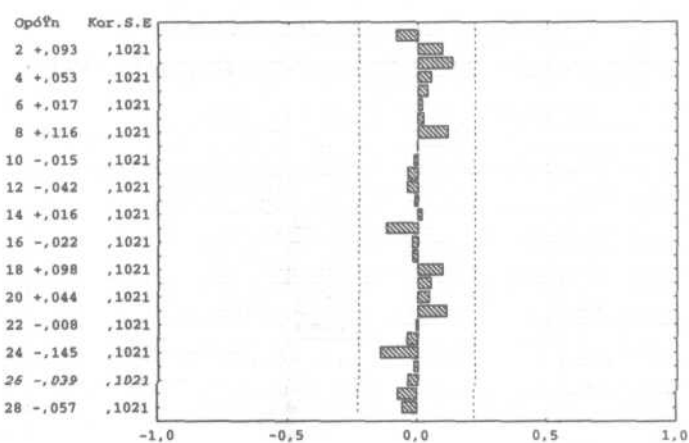
Źródło: obliczenia własne wykonane za pomocą programu Statistica PL

Przedstawione na rysunkach 11 i 12 wykresy przebiegu funkcji ACF i PACF reszt modelu wskazują na brak autokorelacji reszt, co pozwala przyjąć, że reszty są procesem białego szumu i nie są istotne w procesie budowy prognozy.



Rys. 11. Wykres funkcji ACF reszt modelu ARIMA (1,0,0)(1,1,0)<sub>12</sub>

Źródło: obliczenia własne



Rys. 12. Wykres funkcji PACF reszt modelu ARIMA (1,0,0)(1,1,0)<sub>12</sub>

Źródło: obliczenia własne

Opracowany model uznano za odpowiedni do wyznaczenia prognoz przychodów ze sprzedaży na kolejne miesiące. Rezygnując z zapisu operatorowego uzyskuje się model zapisany w postaci jawnej

$$y_{t+1}^* = \phi_1 y_t + (1 + \Phi_1) y_{t-1} - (\phi_1 \Phi_1 + \Phi_1) y_{t-2} - \phi_1 y_{t-3} + \phi_1 \Phi_1 y_{t-4} \quad (5)$$

Korzystając z modelu (5) wyznaczono prognozy na jeden miesiąc oraz dwa i trzy miesiące. Uzyskane wyniki porównano z rzeczywistą wartością przychodów. Wyniki obliczeń zestawiono w tabeli 2.

**Tabela 2**

Prognoza przychodów ze sprzedaży wyznaczona modelem ARIMA (1,0,0)(1,1,0)<sub>2</sub>

Nr kolejnego miesiąca	Dane obserwowane	Prognoza przychodów ze sprzedaży		Błąd procentowy
109	138 244,88	$y_{108+1}^*$	134 685,3583	2,57%
110	134 455,80	$y_{108+2}^*$	134 702,1522	-0,18%
111	131 771,79	$y_{108+3}^*$	126 543,4297	3,97%

Źródło: obliczenia własne

Względny procentowy błąd prognozy nie przekracza 5%. Można uznać, iż uzyskana dokładność jest zadawalająca.

#### 4. WNIOSKI

Modele klasy ARIMA są wykorzystywane w wielu sytuacjach prognostycznych. Opracowanie modeli jest dosyć pracochłonne i wymaga specjalistycznej wiedzy. Wykorzystując modele tej klasy do prognozowania szeregów czasowych, można spotkać się z wieloma problemami, m.in. dotyczącymi określenia rzędu różnicowania i rodzaju modelu. Zastosowanie modeli ARIMA nie gwarantuje w każdym przypadku lepszych wyników w porównaniu z otrzymanymi innymi, prostszymi metodami. Jest to przyczyną, iż w prognozowaniu gospodarczym, zwłaszcza w skali przedsiębiorstwa, stosowane są one dosyć rzadko [1]. Mają one jednak tę zaletę, iż wskazują na wewnętrzną strukturę szeregu i objaśniają mechanizm jego generowania.

Proponowane podejście ilościowe wzbogaca informację decyzyjną i poszerza metodologię prognozowania stosowaną obecnie w obu przedsiębiorstwach. Dotychczas firmy szacowały przyszłe wartości przychodów ze sprzedaży, opierając się na opinii bezpośrednich sprzedawców, a w przypadku firmy produkcyjnej specjalistów z działu logistyki i produkcji. Formułowane opinie były jednak niejednokrotnie oparte na fałszywych przesłankach prognostycznych, co wpływało negatywnie na trafność tak wyznaczonych prognoz.

Wydaje się, iż najbardziej efektywnym podejściem byłoby połączenie w procesie prognostycznym metod jakościowych oraz ilościowych.

### Literatura

- [1] Cieślak M.: *Prognozowanie gospodarcze*. Warszawa, Wydawnictwo Naukowe PWN 2001
- [2] Dittman R.: *Metody prognozowania sprzedaży w przedsiębiorstwie*. Wrocław, Wydawnictwo Akademii Ekonomicznej im. Oskara Langego 2000
- [3] Delurgio S.A.: *Forecasting Principles and Applications*. Boston, Irwin/McGraw-Hill 1998
- [4] Luszniwicz A., Słaby T.: *Statystyka z pakietem komputerowym STATISTICA™PL Teoria i zastosowania*. Warszawa, Wydawnictwo C.H. Beck 2001 [5] Makridakis S., Wheelwright S., Hyndman R.: *Forecasting. Methods and Applications*. New York, John Wiley & Sons, Inc. 1998 [6] Nowak E. (red.): *Prognozowanie gospodarcze. Metody, modele, zastosowanie, przykłady*. Warszawa, Agencja Wydawnicza PLACET 1998 [7] Rutkowski K. (red.): *Logistyka dystrybucji*. Warszawa, Difin 2001