

Ewa Chodakowska, Joanicjusz Nazarko, Anna Worobiej
Politechnika Białostocka

MODELE MARIMA W PROGNOZOWANIU CEN ENERGII NA TOWAROWEJ GIELDZIE ENERGII SA*

1. Wstęp

W Polsce efektem zmian zachodzących w branży elektroenergetycznej jest liberalizacja handlu energią. Jedną z pierwszych wprowadzonych form rynkowego obrotu energią elektryczną była giełda energii.

W Polsce funkcjonuje jedna giełda energii elektrycznej – Towarowa Giełda Energii SA (TGE). Pierwszą uruchomioną na niej formą obrotu był rynek dnia następnego (RDN). Rynek ten prowadzony jest dzień przed dobą, w której nastąpi fizyczna dostawa energii i składa się z 24 notowań, na których cena ustalana jest systemem FIXING. Podstawą decyzji o liczbie ofert i cenie oferty zakupu energii składanej przez uczestników TGE na RDN powinna być m.in. prognoza jej cen, sporządzana z wyprzedzeniem dobowym.

Prognozowanie cen energii na giełdzie jest zagadnieniem nowym. Nie ma jeszcze zbyt wielu prac poświęconych tej tematyce. Przedstawiane są m.in. propozycje wykorzystania sztucznych sieci neuronowych, modeli ARIMA czy modeli GARCH do prognozowania cen energii [Conteras i in. 2003]. Autorzy w niniejszym artykule proponują zastosowanie wielowymiarowych modeli ARIMA (MARIMA – *multivariate autoregressive integrated moving average*) do analizy i prognozowania cen energii notowanych na TGE. W artykule przeprowadzono analizę kształtowania się cen energii elektrycznej na TGE oraz podjęto próbę wyznaczenia prognozy ceny energii z uwzględnieniem wolumenu obrotu.

2. Modele MARIMA dla ceny energii elektrycznej

Model MARIMA dla dwóch stacjonarnych szeregów czasowych Y i X składa się z funkcji przenoszenia TF (*transfer function*) i modelu szumu N_t (*noise model*) [DeLurgio 1998]. Można go zapisać równaniem [Liu, Hudak 2000]:

* Artykuł został opracowany w ramach pracy własnej W/WZ/03/02 finansowanej ze środków KBN.

$$Y_t = C + \frac{\omega(B)}{\delta(B)} X_t + N_t, \quad (1)$$

gdzie: $\delta(B) = 1 - \delta_1 B - \delta_2 B^2 - \dots - \delta_r B^r$,

$$\omega(B) = (\omega_0 - \omega_1 B - \omega_2 B^2 - \dots - \omega_s B^s) B^b,$$

B – operator przesunięcia wstecz,

Y_t – zmienna wyjściowa (objaśniana),

X_t – zmienna wejściowa (objaśniająca),

C – stała,

N_t – model szumu; przy czym jest to model $ARIMA(p,d,q)(P,D,Q)$ dla zmiennej Y , zawierający łączny efekt innych czynników oddziałujących na zmienną prognozowaną.

Model opisany wzorem (1) można też przedstawić w postaci zlinearyzowanej [Liu, Hudak 2000]:

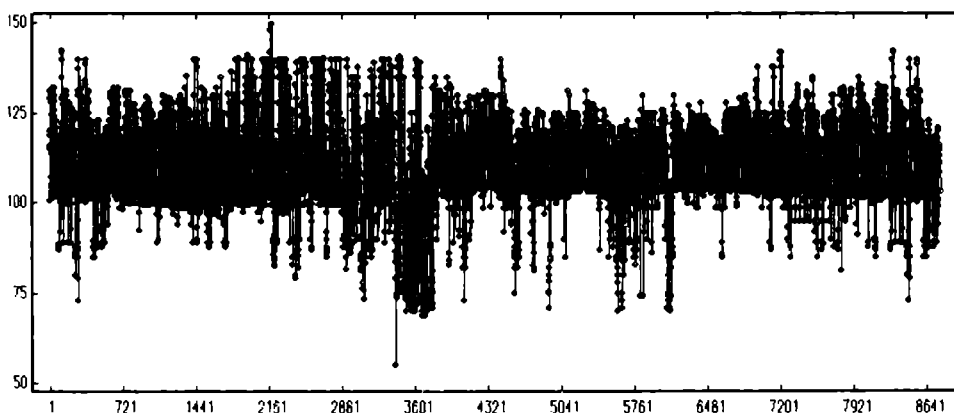
$$Y_t = C + v(B) X_t + N_t, \quad (2)$$

gdzie: $v(B) = v_0 + v_1 B + v_2 B^2 + \dots + v_k B^k$.

W artykule podjęto próbę opracowania modelu MARIMA dla ceny energii, w którym cena jest funkcją swoich poprzednich wartości i ilości energii sprzedawanej na TGE (wolumenu). Model opracowano za pomocą programu SCA Statistical System.

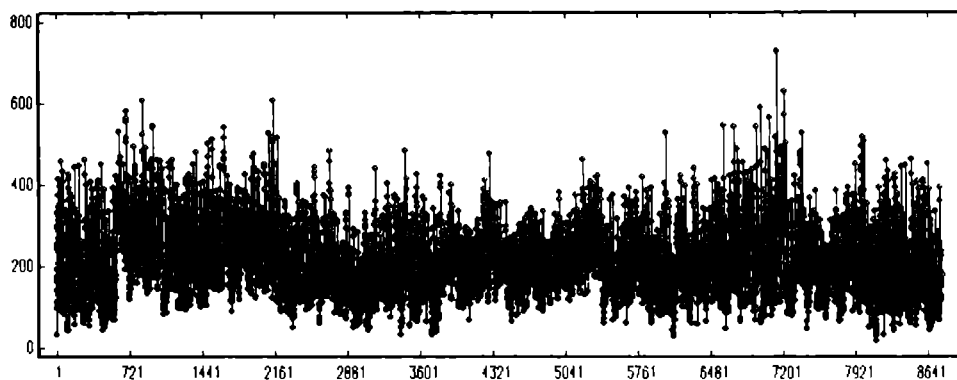
Do budowy modelu wybrano najnowsze dostępne dane: od 1 sierpnia 2003 r. do 31 lipca 2004 r. Analizą objęto pełny rok – sezon zarówno letni, jak i zimowy – czyli okresy o niskim i wysokim zapotrzebowaniu na energię. Zdaniem autorów, pozwala to na bardziej obiektywną ocenę jakości modelu.

Na rysunkach 1 i 2 pokazano kształtowanie się ceny oraz wolumenu w wybranym okresie.



Rys. 1. Kształtowanie się ceny energii w okresie od 1 sierpnia 2003 r. do 31 lipca 2004 r.

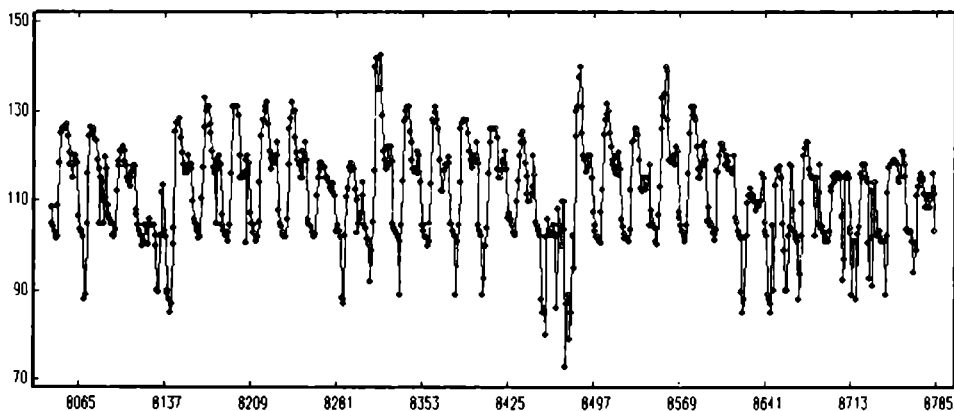
Źródło: opracowanie własne przy wykorzystaniu pakietu SCA.



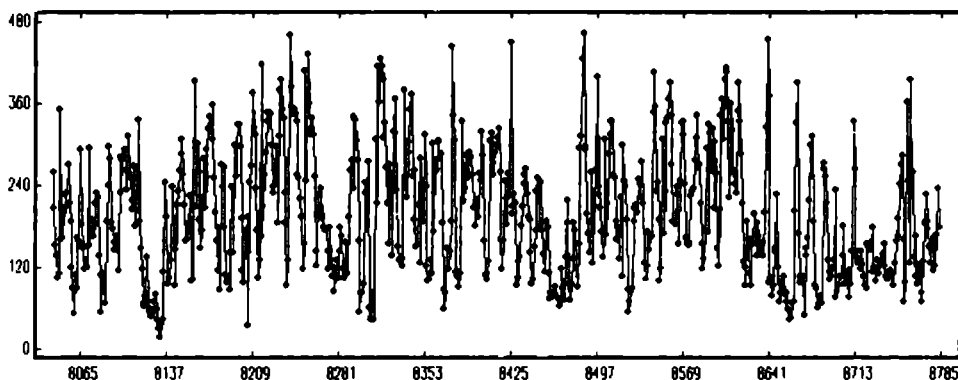
Rys. 2. Kształtowanie się wolumenu energii w okresie od 1 sierpnia 2003 r. do 31 lipca 2004 r.
Źródło: opracowanie własne przy wykorzystaniu pakietu SCA.

Analiza graficznej prezentacji danych pozwala zauważyć podobieństwa pomiędzy szeregami. Oba szeregi charakteryzują się wysoką częstotliwością oraz okresowością: dobową i tygodniową, przy czym wartości wolumenu wykazują znacznie większe wahania losowe.

Istniejącą okresowość lepiej ilustrują miesięczne wykresy kształtowania się cen energii i wolumenu obrotu (rys. 3 i 4).



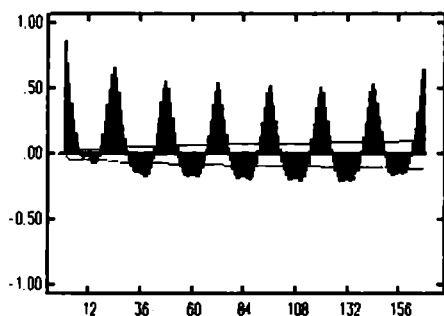
Rys. 3. Kształtowanie się ceny energii elektrycznej w lipcu 2004 r.
Źródło: opracowanie własne przy wykorzystaniu pakietu SCA.



Rys. 4. Kształtowanie się wolumenu obrotu w lipcu 2004 r.
Źródło: opracowanie własne przy wykorzystaniu pakietu SCA.

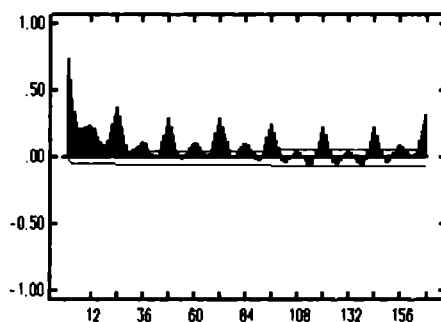
Na podstawie dokonanej wizualizacji danych można zauważyć, że najniższe ceny energii notowane są w weekendy. W każdej dobie energia jest najdroższa pomiędzy godz. 7 a 13 oraz 19 a 22.

Szeregi zarówno cen, jak i wolumenu są niestacjonarne, co potwierdzają wyliczone funkcje autokorelacji przedstawione na rys. 5 i 6. Metodyka budowy modelu MARIMA wymaga doprowadzenia szeregów do postaci stacjonarnej [DeLurgio 1998; Liu, Hudak 2000]. W celu eliminacji niestacjonarności zastosowano różnicowanie z opóźnieniem 1-godzinny oraz sezonowe z opóźnieniem 24-godzinny.



Rys. 5. ACF dla ceny energii

Źródło: opracowanie własne przy wykorzystaniu programu SCA.



Rys. 6. ACF dla wolumenu obrotu energią

Źródło: opracowanie własne przy wykorzystaniu programu SCA.

Rozważano także możliwość różnicowania z opóźnieniem 168, tak by usunąć sezonowość tygodniową. Jednak stwierdzono, że ta operacja nie wpływa istotnie na jakość ostatecznego modelu. Sezonowość dobową jest znacznie silniejsza niż tygodniowa (szczególnie w przypadku wolumenu), a wprowadzenie dodatkowo (lub zamiast różnicowania dobowego) różnicowania z opóźnieniem 168 utrudnia iden-

tyfikację i wydłuża proces estymacji parametrów ze względu na bardziej skomplikowaną postać modelu i jednocześnie nie wpływa na jego lepsze dopasowanie.

Po oddzielnym rozpatrzeniu stacjonarnych szeregów cen i wolumenu stwierdzono, że można by zbudować dla nich podobne sezonowe modele $ARIMA(1,1,1)(0,1,1)_{24}$. Mimo to, nie istnieje znaczna korelacja między nimi (współczynnik korelacji liniowej Pearsona wynosi 0,38). Spróbowano zatem opisać zależność pomiędzy szeregami, korzystając z modelu funkcji przenoszenia.

Na podstawie przesłanek, wynikających z teorii budowy modeli MARIMA, a także wypróbując różne modele do danych, uzyskano modeli kilka o podobnej jakości. Ostatecznie zdecydowano się na model zawierający najmniej parametrów:

$$(1-B)(1-B^{24})C_t = (0,002B^{24})(1-B)(1-B^{24})W_t + \frac{(1-0,88B^{24})(1+0,12B^{168})}{(1+0,13B)}a_t, \quad (3)$$

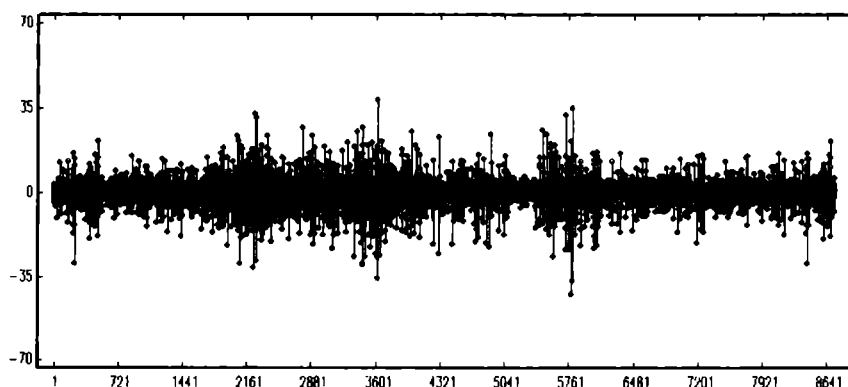
gdzie: C_t – cena w okresie t ,

W_t – wolumen w okresie t ,

a_t – zaburzenia losowe.

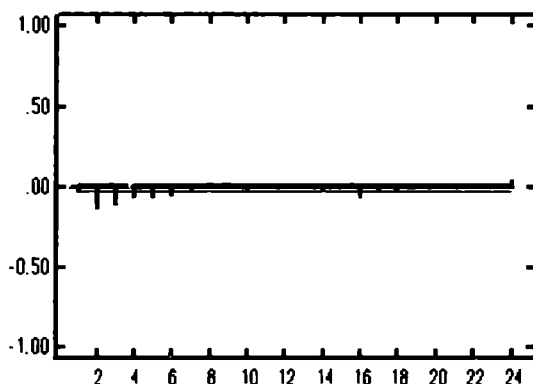
Diagnostyka zbudowanego modelu obejmuje wizualną ocenę jego reszt i sprawdzenie ich autokorelacji. Poprawny model powinien charakteryzować się brakiem autokorelacji reszt, a także brakiem korelacji pomiędzy resztami modelu a stacjonarną zmienną objaśniającą [Liu, Hudak 2000].

Reszty poprawnego modelu powinny oscylować wokół zera, mieć stałą wariancję w czasie oraz rozkład normalny [DeLurgio 1998]. Reszty przedstawionego modelu (rys. 7) oscylują wokół zera, lecz można zauważyć pewne grupowanie się w czasie ich wariancji. Sporządzony histogram reszt modelu nie wskazywał, aby reszty nie miały rozkładu normalnego. Korzystając z programu SCA, nie można zastosować testów statystycznych do weryfikacji hipotez o zgodności z rozkładem normalnym.



Rys. 7. Reszty modelu

Źródło: opracowanie własne przy wykorzystaniu programu SCA.



Rys. 8. ACF reszt modelu

Źródło: opracowanie własne przy wykorzystaniu programu SCA.

Przy badaniu funkcji autokorelacji można spostrzec, że dla kilku pierwszych opóźnień jej wartości okazały się statystycznie istotne. Oznacza to, że w szeregu reszt występuje autokorelacja. Jednak należy pamiętać, że bardzo duża liczba obserwacji w szeregu powoduje, że przy poziomie ufności 95% przedział ufności jest bardzo mały ($\pm 0,02$) (rys. 8). Podjęte próby wprowadzenia dodatkowych parametrów w modelu nie przynosiły istotnej poprawy jego jakości. Wobec tego zdecydowano się zaakceptować model (3) do dalszej analizy.

Sprawdzenie założenia o niezależności między stacjonarną zmienną objaśniającą a resztami modelu polega na obliczeniu wartości funkcji korelacji krzyżowej (CCF) między resztami jednowymiarowego modelu ARIMA dla zmiennej objaśniającej a resztami modelu MARIMA. W analizowanych danych funkcja CCF ma jedną znaczną wartość, wynoszącą 0,2 dla opóźnienia 0. Oznacza to, że w analizowanych danych jest wzajemna, nieopóźniona zależność pomiędzy ceną a wolumenem.

Zgodnie z założeniami, model nie powinien zawierać nieopóźnionej zmiennej objaśniającej. Po pierwsze dlatego, że właściwie modele MARIMA stosuje się, gdy zależność pomiędzy zmienną objaśnianą a objaśniającą jest jednokierunkowa. Warunek ten nie przekreśla całkowicie możliwości wypróbowania modeli MARIMA. Należy wówczas raczej użyć „zredukowanego” modelu MARIMA, czyli takiego, w którym nie występuje nieopóźniona zmienna objaśniająca [Liu, Hudak 2000]. Kolejnym powodem byłaby trudność wykorzystania takiego modelu w praktyce. Na TGE oferta cenowa jest składana wraz z wolumenem, czyli na okres, na który budowana jest prognoza cen, wolumen także nie jest znany. W takiej sytuacji opracowanie prognozy cen poprzedzać by musiała prognoza wolumenu.

Pomimo niespełnienia w 100% wszystkich założeń formalnych, model okazał się dobrze dopasowany do danych. Ocenę dopasowania przeprowadzono na podstawie analizy wartości najpopularniejszych miar trafności, takich jak: błąd średni (ME – *mean error*), pierwiastek średniego błędu kwadratowego (RMSE – *root mean square error*), średni

bezwzględny procentowy błąd (MAPE – *mean absolute percentage error*) oraz współczynnika determinacji R^2 (tab. 1).

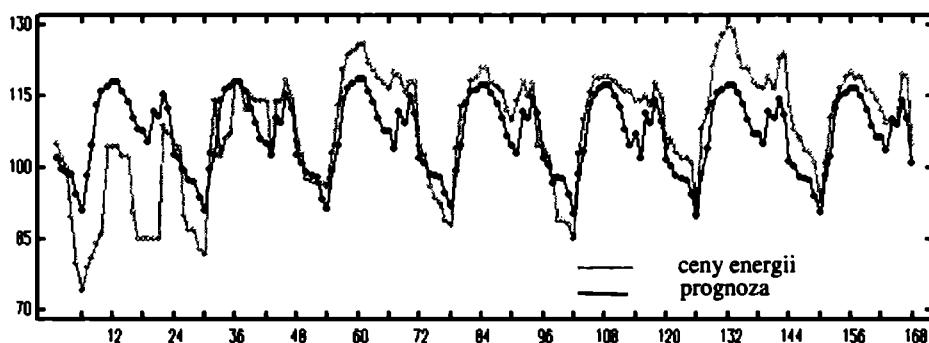
Tabela 1. Wartości błędów modelu

Miara	Wartość
ME	-0,004
RMSE	4,82
MAPE	2,77%
R^2	0,85

Źródło: opracowanie własne przy wykorzystaniu programu SCA.

Na podstawie oszacowanych błędów można stwierdzić, że wybrany model dobrze odzwierciedla zmiany wartości szeregu. Model ten ani systematycznie nie doszacowuje, ani nie przeszacowuje wartości prognoz, na co wskazuje wartość błędu średniego. Przeciętne odchylenia prognoz od wartości rzeczywistych wynoszą tylko 4,82 zł. Również miary procentowe wskazują na dobre dopasowanie modelu do danych rzeczywistych (MAPE wynosi tylko 2,7% wartości rzeczywistej, a – R^2 85%)

Korzystając z oszacowanego modelu, obliczono prognozę dla kolejnych dób pierwszego tygodnia sierpnia 2004 r. Uzyskane wyniki porównano z rzeczywistą wartością cen. Wyniki analizy przedstawiono wizualnie na rys. 9 oraz numerycznie w tab. 2.



Rys. 9. Ceny energii w sierpniu 2004 r. oraz prognoza wyznaczona za pomocą modelu MARIMA

Źródło: opracowanie własne przy wykorzystaniu programu SCA.

Tabela 2. Wartości błędów wyznaczonej prognozy

Błąd	Wartość
ME	1,08
RMSE	9,07
MAPE	6,97%

Źródło: opracowanie własne przy wykorzystaniu programu SCA.

Wartości błędów wyznaczonej prognozy są znacznie wyższe niż w ocenie modelu. Mimo że model był dobrze dopasowany, prognoza nie jest zadowalająca.

W badanym okresie model zwykle nie doszacowuje wartości prognozowanej zmiennej.

3. Podsumowanie

Modele MARIMA są dość uniwersalnym narzędziem modelowania i prognozowania. Przedstawiają zależność w postaci łatwej do interpretacji i pozwalają uwzględnić wpływ innych czynników na badaną zmienną. Jednak, jak dowiodła przeprowadzona analiza, opracowanie modelu nie jest proste. W praktyce spełnienie wszystkich założeń modelu może okazać się kłopotliwe.

W rozpatrywanym zagadnieniu prognozowania cen energii na TGE autorom nie udało się osiągnąć zadowalających wyników. Ich zdaniem, nie przekreśla to możliwości zastosowania modeli MARIMA w innej sytuacji prognostycznej.

Autorzy zamierzają kontynuować badania i sprawdzić modele MARIMA dla innych zestawów zmiennych objaśniających (np. pogody), mogących mieć wpływ na kształtowanie się ceny energii elektrycznej na giełdzie energii.

Literatura

Contreras J., Espinola R., Nogales F. J., Conejo A.J., *ARIMA Models to Predict Next-Day Electricity Prices*, „IEEE Transactions on Power Systems” sierpień 2003, vol. 18, nr 3, s. 1014-1020.

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

Halicka K., Nazarko J., Jurczuk A., *Zarządzanie portfelem zakupów na rynku dnia następnego*, „Zagadnienia Techniczno-Ekonomiczne. Kwartalnik AGH”, z. 1, t. 48, Kraków 2003, s. 111-123.

Liu L., Hudak G. B., *Forecasting and Time Series Analysis Using the SCA Statistical System*, Scientific Computing Associated Corp., Chicago 2000.

<http://www.polpx.pl>, 1 września 2004.

MARIMA MODELS FOR ENERGY PRICE FORECASTING AT THE POWER EXCHANGE IN POLAND

Summary

The paper discusses application of *Multivariate AutoRegressive Integrated Moving Average* models for energy prices forecasting at the Day-Ahead Market at The Power Exchange in Poland. The analysis of energy prices has been carried out and at the same time there has been made an attempt to find a relation between price and its volume. On the basis of available data, a MARIMA model has been proposed which describes price as a function of its previous values and energy volume.