

Barbara Gładysz, Wiktor Kołwzan, Jacek Mercik

Politechnika Wrocławska

PROGNOZOWANIE W ZARZĄDZANIU ENERGIA

W artykule przeprowadzono statystyczną analizę poziomu obciążenia systemu elektroenergetycznego. Wyróżniono istotne czynniki warunkujące wielkość zużycia energii, takie jak: sezonowość, cykliczność, struktura tygodniowa zapotrzebowania na energię oraz temperatura. Czynniki te przyjęto jako zmienne objaśniające w konstruowanych modelach ekonometrycznych. Obok wymienionych czynników uwzględniono także zużycie energii z poprzednich okresów. Modele były konstruowane dla obciążenia w poszczególnych godzinach na całej próbie i na jednorodnych podpróbach wyznaczonych metodami składowych głównych i analizy skupień. Modele te zastosowano w prognozowaniu. Porównano błędy prognoz dla różnych klas modeli ekonometrycznych oraz modeli ARIMA i Wintersa.

1. Wstęp

Rada Unii Europejskiej określiła w 1996 r. podstawowe zasady funkcjonowania rynku energii elektrycznej. W ślad za tym Polska wprowadziła odpowiednie uregulowania dla stworzenia rynku energii.

Na rynku energii elektrycznej pojawiła się konkurencja. Obecnie dystrybutorzy energii muszą nią tak zarządzać, aby zminimalizować koszty jej zakupu. Wysokość kosztów w praktyce zależy od właściwie podjętej decyzji o wysokości bieżącego popytu na energię. Zatem istotne jest dobranie efektywnych metod prognozowania.

W literaturze prognozowanie zapotrzebowania na energię elektryczną dzieli się na dwie zasadnicze grupy:

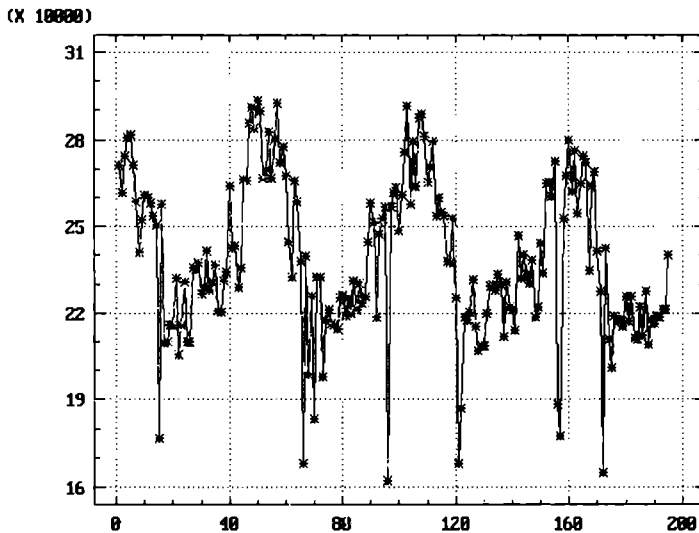
- prognozowanie długoterminowe,
- prognozowanie krótkoterminowe.

Zastosowane w artykule techniki prognozowania zapotrzebowania na energię elektryczną zaliczane są przez ekspertów z dziedziny energetyki do prognozowania krótkoterminowego. Ten rodzaj prognozowania jest istotnym elementem współczesnych systemów zarządzania energią, dokładne bowiem i dostarczone w porę informacje o przyszłym obciążeniu systemu odgrywają zasadniczą rolę w sterowaniu systemem i zasadniczo wpływają na poziom kosztów wytwarzania energii elektrycznej. Prognozowanie takie nazywane jest w obszarze energetyki krótkoterminową prognozą obciążenia i obejmuje horyzont czasowy od 1 godziny do 168 godzin.

Dane dotyczące systemów elektroenergetycznych zbierane są co godzinę. Zatem informacje o szeregach czasowych zapotrzebowania na energię są łatwo dostępne.

Przez pojęcie szeregu czasowego (stochastycznego) tradycyjnie rozumie się dowolny proces zachodzący w czasie o wartościach rzeczywistych. Szereg taki zawiera dwie składowe: systematyczną oraz składową przypadkową (losową). Składowa systematyczna w konkretnym procesie może występować w postaci stałego przeciętnego poziomu badanej zmiennej, trendu, składowej okresowej cyklicznej lub sezonowej.

Śledząc naturę popytu na energię, można zauważyć wyraźnie, że cechuje go cykliczność, sezonowość, losowość (rys. 1). Liczba i złożoność czynników determinujących poziom tego specyficznego procesu stochastycznego implikuje mnogość stosowanych modeli i technik prognozowania krótkoterminowego, takich jak modele autoregresji (ARIMA), metody wygładzania wykładniczego (Holta, Wintersa), metody sezonowej dekompozycji (CSD), wieloczynnikowe modele ekonometryczne, sieci neuronowe, metody rozmyte. Szczególnie cenione są techniki prognozowania, które uwzględniają czynniki meteorolo-



Rys. 1. Zapotrzebowanie na energię w poniedziałki o godzinie 12.00

Źródło: opracowanie własne.

giczne [Valenzuela, Mazumdar 2000]. Przegląd metod stosowanych w prognozowaniu popytu na energię można znaleźć w pracach [Malko 1995; Lotuffo, Minussi 1999].

W obszarze prognozowania długoterminowego wykorzystuje się metody statystyczne związane z modelami ekonometrycznymi. W charakterze czynników (zmiennych objaśniających) uwzględnia się wartość produktu krajowego brutto (PKB), udział przemysłu w tworzeniu PKB, zużycie energii pierwotnej oraz zatrudnienie w przemyśle [Malko 1995]. Podejmowane są też próby użycia sieci neuronowych do tworzenia prognoz w obszarze prognozowania długoterminowego. Stosowanie sieci w prognozowaniu sprawia, że procedury obliczeniowe wykazują większą elastyczność i zdolności adaptacyjne do zmieniających się warunków zewnętrznych.

W pracy zaprezentowano modele prognozowania krótkoterminowego ARIMA, Wintersa oraz ekonometryczne. Podobne podejście można znaleźć w pracach [Górecka, Kosyk, Szmit 2000; Szmit 2003], gdzie w modelach ekonometrycznych przyjęto jako zmienne objaśniające zmienne zero-jedynkowe [Szmit 2003]. W niniejszej pracy zaproponowano modele ekonometryczne ze zmiennymi objaśniającymi ciągłymi oraz z syntetycznymi zmiennymi dyskretnymi.

2. Wielowymiarowa analiza poziomu obciążenia systemu elektroenergetycznego

W celu wyróżnienia istotnych czynników mających wpływ na poziom obciążenia systemu elektroenergetycznego zastosowano analizę skupień oraz metodę składowych głównych [Morrison 1989]. Badano takie cechy, jak średnia i wariancja analizowanych szeregów czasowych, opierając się na danych z lat 1998-2001 (30 000 obserwacji).

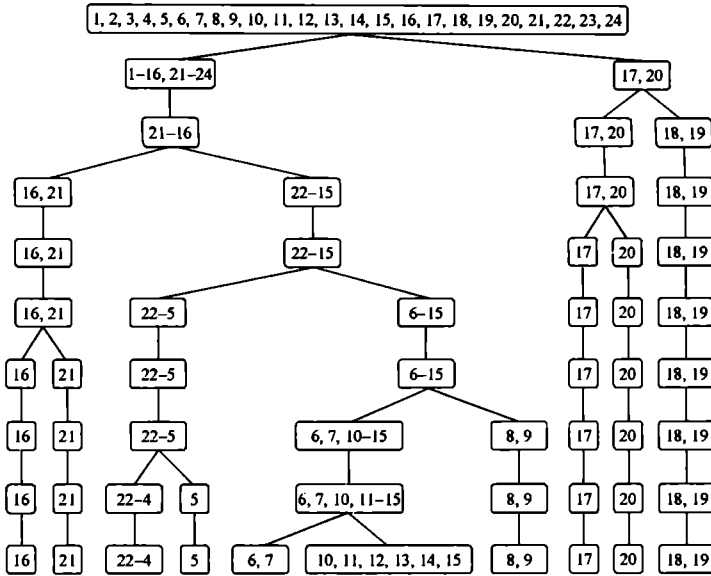
2.1. Analiza zmienności obciążenia

Celem metody analizy skupień jest wydzielenie spośród badanych obiektów grup jednorodnych ze względu na badane cechy. Matematycznie obiekty traktuje się jako punkty wielowymiarowej przestrzeni o składowych wyznaczonych przez wartości zaobserwowanych cech badanych obiektów. Problem polega na wyodrębnieniu obiektów podobnych (jednorodnych grup). Obiekty uważa się za podobne, gdy odległość między nimi jest mała. W analizie skupień odległość między obiektami rozumie się w bardzo szerokim sensie. Dla cech mierzalnych może być to odległość Euklidesowa i Mahalanobisa. Jako miarę odległości przyjmuje się również współczynnik korelacji. Aby uniezależnić wyniki grupowania od jednostek, w jakich zmienne są mierzone, odległość wartości zmiennych poddaje się standaryzacji.

Przeprowadzona metodą skupień klasyfikacja zmienności dziennego oraz miesięcznego zapotrzebowania na energię w poszczególnych godzinach pozwoliła wyróżnić następujące klasy godzin o zbliżonej wariancji:

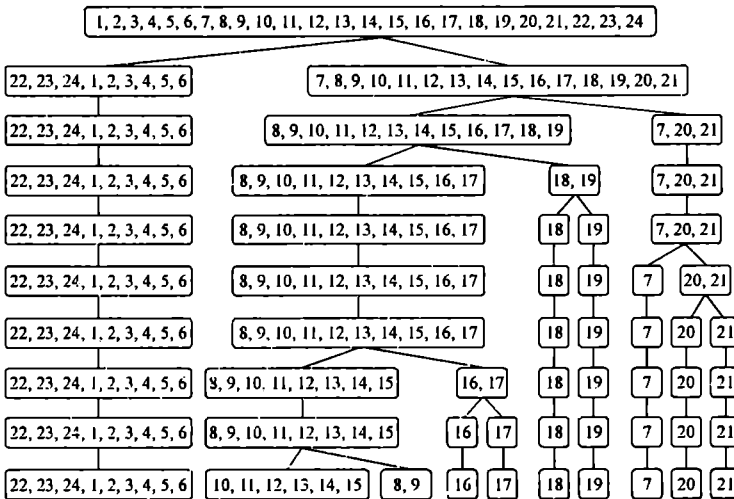
- godziny nocne,
- godziny pracy,

- godziny poranne i wieczorne.



Rys. 2. Klasy godzin dla dziennego zapotrzebowania

Źródło: opracowanie własne.



Rys. 3. Klasy godzin dla miesięcznego zapotrzebowania

Źródło: opracowanie własne.

Za miarę zmienności obciążenia systemu przyjęto wariancję. Na rys. 2 i 3 przedstawiono otrzymane grupy godzin dla różnej liczby klas skupień. W klasyfikacji zastosowano metodę średniej z metryką Euklidesową.

W obu klasyfikacjach najbardziej liczne grupy tworzą godziny nocne oraz godziny pracy. Natomiast godziny poranne oraz wieczorne tworzą pojedyncze klasy, gdyż charakteryzują się one największą i zróżnicowaną wariancją. Godziny nocne mają wariancję zbliżoną i małą. Godziny pracy mają pośrednią podobną zmienność.

2.2. Czynniki poziomu obciążenia

Dla wyróżnionych metodą skupień klas godzin przeprowadzono analizę składowych głównych średniego poziomu obciążenia elektroenergetycznego.

Metoda składowych głównych jest narzędziem, które umożliwia wyznaczenie czynników warunkujących badane zjawisko. Pozwala ona wykryć pewne wspólne czynniki wpływające na istnienie zależności między badanymi zmiennymi. Istota tej metody polega na tym, że obserwowalne zmienne można przedstawić w postaci liniowej funkcji mniejszej liczby ukrytych niezależnych zmiennych nieobserwowalnych, zwanych czynnikami [Aczel 2000]. Metoda ta nie wymaga spełnienia założenia o normalności rozkładu badanej populacji. Natomiast normalność wymagana jest wtedy, gdy bada się asymptotyczne własności współczynników składowych głównych, jak też w przypadku konstrukcji testów i przedziałów ufności dla tych współczynników.

Tabela 1. Analiza składowych głównych

Klasy godzin	Numer składowej	Procent wariancji	Skumulowany procent wariancji
Nocne	1	75,95	75,95
	2	20,33	96,28
	3	3,53	99,81
Popołudniowe i wieczorne	1	96,87	96,87
	2	1,60	98,47
	3	1,29	99,76
Robocze	1	97,01	97,01
	2	2,89	99,90

Źródło: opracowanie własne.

Z przeprowadzonej analizy można wnioskować, że na poziom obciążenia w godzinach nocnych mają wpływ trzy czynniki. Podobnie jest w przypadku godzin popołudniowych. Natomiast obciążenie w godzinach roboczych wyjaśniają dwa czynniki. Wartości wariancji poszczególnych składowych głównych dla wszystkich grup godzin przedstawiono w tab. 1.

3. Wieloczynnikowe modele prognozowania

Celem powyższej analizy było wyłonienie grup godzin o zbliżonej wariancji dla średniego dziennego i miesięcznego zapotrzebowania na energię. Nieco odmienna struktura dendrytów klasyfikacyjnych (rys. 2 i 3) pozwala wnioskować, iż na zmienność obciążenia elektroenergetycznego w poszczególnych godzinach wpływa zarówno dzień tygodnia, jak i miesiąc. Stąd modele na potrzeby prognozowania konstruowano odrębnie dla dni tygodnia, miesięcy oraz godzin. Ponadto w pracy [Gładysz, Kołwzan, Mercik 2004], stosując metodę składowych głównych pokazano, że obok pory dnia, dnia tygodnia oraz pory roku na poziom obciążenia wpływ mają takie czynniki, jak poziom obciążenia systemu w poprzednich okresach, czynniki meteorologiczne oraz dni świąteczne.

Uwzględniając sezonowość i cykliczność szeregów czasowych obciążenia oraz zależność bieżącego obciążenia od jego wartości w poprzednich okresach, jak również wpływ warunków klimatycznych, na potrzeby prognozowania przyjęto modele ARIMA, Wintersa oraz modele ekonometryczne. W modelowaniu ekonometrycznym jako potencjalne zmienne objaśniające przyjęto:

- OBCIĄŻENIE – poziom obciążenia systemu z poprzednich okresów,
- NIEDZIELA – zmienna binarna (1 – dla niedziel, 0 – dla pozostałych dni),
- DZIEŃ TYGODNIA zmienna dyskretna (1, 2, ..., 7 odpowiednio dla poniedziałku, ..., niedzieli),
- TYP DNIA – przyjmująca wartości 1, 2, 3, 4 w zależności od dnia tygodnia (sobota, niedziela, poniedziałek i pozostałe dni),
- ŚWIĘTO – zmienna binarna (1 – dla świąt, 0 – dla pozostałych dni),
- WEEKEND – zmienna binarna (1 – dla sobót i niedziel, 0 – dla dni powszednich),
- MIESIĄC – zmienna dyskretna,
- TEMPERATURA – temperatura.

Do poszczególnych modeli zmienne dobrano metodą krokową. Dane do budowy modeli pochodziły z blisko czterech lat (1369 obserwacji dziennych dla każdej godziny doby). Przy konstrukcji modeli odrzucono obserwacje z ostatniego miesiąca w celu wyznaczenia błędów *ex post* prognoz. Modele były konstruowane dla obciążenia w poszczególnych godzinach na całej próbie oraz na jednorodnych podpróbach, zgodnie z wcześniejszymi wynikami metody składowych głównych i analizy skupień. W tab. 2 przedstawiono wybrane modele oraz błędy prognoz *ex post* dla różnych horyzontów czasowych.

Tabela 2. Modele prognozowania, horyzont czasowy i błędy prognoz

Model	Liczba obserwacji	Zmienna objaśniana – obciążenie o godzinie	Zmienne objaśniające/ parametry modelu	Horyzont prognozy	Błąd <i>ex post</i> (%)
ARIMA	1339 obserwacji dziennych	12	ARIMA (1, 1, 1) (1, 0, 2) ₇ – bez stałej	48-godzinny	3,0
Model Wintersa	1339 obserwacji dziennych	12	Parametry (a, b, c) = (0,9; 0,15; 0,85)	48-godzinny	3,0
Model ekonometryczny	1339 obserwacji dziennych	12	OBCIĄŻENIE	kilkugodzinny	3,4
Model ekonometryczny	1339 obserwacji dziennych	12	OBCIĄŻENIE TEMPERATURA	kilkugodzinny	3,4
Model ekonometryczny	1339 obserwacji dziennych	12	OBCIĄŻENIE NIEDZIELA ŚWIĘTO TEMPERATURA	kilkugodzinny	1,7
Model ekonometryczny	1339 obserwacji dziennych	12	OBCIĄŻENIE TEMPERATURA	jednodniowy (co najmniej 24-godzinny)	3,3
Model ekonometryczny	1339 obserwacji dziennych	12	OBCIĄŻENIE NIEDZIELA ŚWIĘTO TYP DNIA TEMPERATURA	jednodniowy (co najmniej 24-godzinny)	2,8
Model ekonometryczny	90 obserwacji dziennych (wrzesień)	12	OBCIĄŻENIE WEEKEND TYP DNIA TEMPERATURA	kilkugodzinny	1,2
Model ekonometryczny	90 obserwacji dziennych (wrzesień)	12	OBCIĄŻENIE NIEDZIELA WEEKEND TYP DNIA TEMPERATURA	jednodniowy (co najmniej 24-godzinny)	2,2
Model ekonometryczny	195 obserwacji (poniedziałek)	12	OBCIĄŻENIE	kilkugodzinny	1,3
Model ekonometryczny	195 obserwacji (czwartek)	12	OBCIĄŻENIE	kilkugodzinny	0,7
Model ekonometryczny	195 obserwacji (środa)	12	OBCIĄŻENIE	kilkugodzinny	0,5
ARIMA	1339 obserwacji dziennych	1	ARIMA (3, 1, 1) (2, 0, 1) ₇ bez stałej	48-godzinny	2,1
Model Wintersa	1339 obserwacji dziennych	1	Parametry (a, b, c) = (0,85; 0,1; 0,8)	48-godzinny	2,1

Model ekonometryczny	1339 obserwacji dziennych	1	OBCIĄŻENIE NIEDZIELA ŚWIĘTO TYP DNIA MIESIĄC TEMPERATURA	Jednodniowy (co najmniej 24-godzinny)	1,9
Model ekonometryczny	90 obserwacji dziennych (wrzesień)	1	OBCIĄŻENIE DZIEŃ TYGODNIA TEMPERATURA	kilkugodzinny	0,9
Model ekonometryczny	90 obserwacji dziennych (wrzesień)	1	OBCIĄŻENIE TEMPERATURA	jednodniowy (co najmniej 24-godzinny)	1,9
Model ekonometryczny	90 obserwacji dziennych (wrzesień)	1	OBCIĄŻENIE NIEDZIELA TYP DNIA WEEKEND TEMPERATURA	jednodniowy (co najmniej 24-godzinny)	2,1
Model ekonometryczny	195 obserwacji (czwartek)	1	OBCIĄŻENIE	tygodniowy	1,7
Model ekonometryczny	195 obserwacji (środa-czwartek)	1	OBCIĄŻENIE	kilkugodzinny	0,7
Model ekonometryczny	195 obserwacji (czwartek)	4	OBCIĄŻENIE	tygodniowy	2,9
Model ekonometryczny	195 obserwacji (czwartek)	4	OBCIĄŻENIE	kilkugodzinny	0,4

Źródło: opracowanie własne.

Analizując tabelę, można zaobserwować, że w przypadku modeli zbudowanych na 1339 obserwacjach, w których zmienną objaśniającą jest poziom obciążenia systemu z poprzednich okresów, błędy prognoz są rzędu 3,5% dla godzin dziennych i 2% dla godzin nocnych. Uwzględnienie w tych modelach temperatury zasadniczo nie zmienia wielkości błędu. Błąd ten nie zmienia się również wraz ze zmianą długości horyzontu prognoz.

Uwzględnienie w modelach zmiennych objaśniających związanych ze strukturą obciążenia w skali tygodnia zmniejsza wielkość błędu dla godzin dziennych do 2,8%. W tym przypadku zmniejszenie horyzontu prognoz do kilku godzin powoduje dalszy spadek błędu do ok. 2%.

Należy podkreślić, że większym błędem obarczone są prognozy dla godzin dziennych niż dla godzin nocnych.

Konstruując modele oparte na szeregach czasowych dla wyróżnionych wcześniej wielowymiarową analizą statystyczną jednorodnych danych (np. wrzesień, czwartek), osiągnięto zmniejszenie błędu nawet poniżej 1% w przypadku horyzontu kilkugodzinnego. Wielkość błędu w tym przypadku zależy również od dnia tygodnia (największy błąd jest dla poniedziałku).

Prognozy wyznaczone sekwencyjnie na podstawie modeli ARIMA oraz Wintersa, szczegółowo zaprezentowane w pracy [Gładysz 2003], charakteryzują się błędem rzędu 3% dla godzin roboczych oraz 2% dla godzin nocnych.

4. Podsumowanie

W konstrukcji modeli prognostycznych popytu na energię starano się dobrać właściwy układ zmiennych objaśniających spośród licznego zbioru czynników warunkujących poziom obciążenia. Dobrane parametry dla modeli ARIMA, Wintersa oraz zmienne objaśniające dla modeli ekonometrycznych pozwoliły otrzymać zadowalające błędy *ex post* otrzymanych prognoz. Modele ekonometryczne skonstruowane dla wyróżnionych metodą skupień jednorodnych klas danych pozwoliły dodatkowo obniżyć wartości tych błędów. Zaobserwowano, że wielkość błędów prognoz zależy zarówno od długości horyzontu czasowego, jak i od jednorodności danych.

Literatura

- Aczel A.D., *Statystyka w zarządzaniu*, PWN, Warszawa 2000.
- Gładysz B., Kołwzan W., Mercik J., *The Analysis and Forecasting of the Electric-energy Demand Level*, Proceedings of 10-th International Multi-Conference on Advanced Computer Systems, October 22-24, Międzyzdroje 2003.
- Gładysz B., Kołwzan W., Mercik J., *Wieloczynnikowe modele prognozowania w elektroenergetyce w badania operacyjne i systemowe*, [w:] *Podjęmowanie decyzji. Podstawy metodyczne i zastosowania*, red. L. Kulikowski, J. Kacprzyk, J. Słowińska, Akademicka Oficyna Wydawnicza EXIT, Warszawa 2004.
- Górecka A., Kosyk A., Szmít M., *Demand for Electric Energy. Time Series Models*, 49-th International Conference, München 2000.
- Szmít A., *Analiza wybranych charakterystyk rozkładów błędów prognoz zapotrzebowania na energię elektryczną*, [w:] *Taksonomia 10. Klasyfikacja i analiza danych – teoria i zastosowania*, red. K. Jajuga, M. Waleśiak, AE, Wrocław 2003.
- Lotufo A.D.P., Minussi C.R., *Electric Power Systems Load Forecasting. A Survey*, Proceedings of IEEE Power Tech '99 Conference, Budapest 1999.
- Malko J., *Wybrane zagadnienia prognozowania w elektroenergetyce*, Oficyna Wydawnicza Politechniki Wrocławskiej, Wrocław 1995.
- Morrison D., *Wielowymiarowa analiza statystyczna*, PWN, Warszawa 1989.
- Valenzuela J., Mazumdar M., *Statistical Analysis of Electric Power Production Costs*, IIE Transactions, 32, 2000.

FORECASTING IN ENERGY SUPPLY MANAGEMENT

Summary

The analysis of capacity of electricity system was done for years 1998-2001 (30.000 observations). The significant factors were established. Among them there are: consumption of energy in previous periods, seasonality, cyclicity, atmospheric temperature, energy weekly demand structure. Those factors were used as explanation variables in econometric models. The delay variables were also used. The models were constructed for homogenous classes of data. Afterwards, models were used for forecasting. Prognostics errors for different classes were compared. Magnitude of errors depends on time horizon and homogeneity of data.