

WPLYW WSTĘPNEGO PRZETWARZANIA DANYCH NA JAKOŚĆ KRÓTKOTERMINOWYCH PROGNOZ ZAPOTRZEBOWANIA NA ENERGIĘ ELEKTRYCZNĄ

Krzysztof Nęcka

*Katedra Energetyki i Automatykacji Procesów Rolniczych
Uniwersytet Rolniczy w Krakowie*

Streszczenie. Celem pracy była analiza wpływu różnych metod wstępnego przetwarzania danych wejściowych, takich jak np. średnia ruchoma, wyrównywanie wykładnicze, filtr 4253H, na jakość prognoz godzinowego zapotrzebowania na energię elektryczną opracowanych metodami regresyjnymi. Cel pracy zrealizowano na podstawie badań własnych wykonanych w rozdzielni nN, zlokalizowanej na terenie nowoczesnej ubojni drobiu w południowej części Małopolski. Wykonane analizy skupień metodą *k*-średnich i metodą EM pokazały, że ze względu na podobieństwo przebiegu godzinowego zapotrzebowania na energię elektryczną optymalny będzie podział dni tygodnia na 3 skupienia, tj. dni robocze, dni poprzedzające dzień wolny od pracy oraz dni wolne od pracy, i budowa trzech niezależnych modeli. W zastosowaniach praktycznych najważniejszym parametrem oceny modeli jest sumaryczna wartość rzeczywistej ilości energii bilansującej ΔESR . Dla większości budowanych modeli na bazie zmiennych przekształconych zaobserwowano zmniejszenie wartości wskaźnika ΔESR względem modeli budowanych w oparciu o zmienną egzogeniczną nieprzekształconą. Największe, ponad 6% zmniejszenie wartości analizowanego wskaźnika uzyskano w modelu III dla zmiennej wejściowej wygładzonej oknem Daniela o rozpiętości 5. Ze względu na najniższą wartość sumarycznej ilości energii bilansującej w zastosowaniach praktycznych powinny być jednak preferowane modele budowane na bazie szeregu czasowego godzinowego zużycia energii elektrycznej dla całego zakładu wygładzonego filtrem 4253H.

Słowa kluczowe: model regresyjny, prognoza krótkoterminowa, przetwarzanie wstępne, zużycie energii elektrycznej, zasada TPA.

Wprowadzenie

Proces liberalizacji rynku energii elektrycznej w Polsce rozpoczął się 4 września 1998 roku. W tym dniu 21 największych odbiorców końcowych (zużywających rocznie powyżej 500 GWh energii) uzyskało prawo do zmiany sprzedawcy energii elektrycznej. W kolejnych latach rynek otwierał się na coraz to mniejszych odbiorców. Od 1 lipca 2004 roku

wszystkie firmy mogą swobodnie wybierać sprzedawcę energii elektrycznej, który może być jej wytwórcą lub spółką obrotu (Opis zasady TPA ...2012). Odbiorca, decydujący się na korzystanie z przysługującego mu prawa udziału w rynku dobowo-godzinowym, może liczyć na obniżenie ceny jednostkowej zakupu energii elektrycznej od kilku do kilkunastu procent względem ceny taryfowej (Koncepcja wspólnego zakupu energii; Pozycja konsumenta na rynku energii elektrycznej). Zmuszony jest on jednak do opracowywania grafików zapotrzebowania na energię elektryczną w poszczególnych godzinach doby handlowej i ich dostarczania w dobie $n-2$ lub $n-1$ do lokalnego Operatora Systemu Dystrybucyjnego. Jakość opracowanego grafiku przekłada się na wynik finansowy przedsiębiorstwa. Jeżeli bowiem w dowolnej godzinie doby handlowej odbiorca zużył inną ilość energii niż zamówił w dobie $n+1$ musi dokonać rozliczenia różnicy bilansowej i zamiast oczekiwanych zysków mogą pojawić się koszty (Koszty bilansowania...). Niedobór energii lub jej nadwyżki dla poszczególnych godzin należy kupić lub sprzedać na rynku bilansującym, gdzie ceny mogą być znacząco odmienne od tych zapisanych w kontrakcie, gdyż są one kształtowane przez rynek.

Rozwój konkurencyjnego rynku energii elektrycznej oraz ciągły wzrost kosztów energii zmuszają jego aktywnych uczestników do ciągłego poszukiwania nowych oraz doskonalenia znanych już narzędzi, pozwalających na prognozowanie zapotrzebowania na energię elektryczną z odpowiednim wyprzedzeniem czasowym. Predykcja godzinowego zapotrzebowania na energię jest zadaniem bardzo trudnym ze względu na dużą liczbę szybkozmennych czynników, wpływających na wartość zapotrzebowania na energię elektryczną. W literaturze jest opisanych wiele modeli i technik prognozowania krótkoterminowego (Gładysz, 2009; Łyp, 2007; Małopolski i Trojanowska, 2009a, 2009b; Nęcka, 2011a, 2011b; Piotrowski, 2007; Trojanowska, 2002; Trojanowska i Małopolski, 2009), jednak nie udało się opracować jednej uniwersalnej procedury, gwarantującej osiągnięcie optymalnego wyniku finansowego dla dowolnego przedsiębiorstwa. Istnieje zatem konieczność ciągłego poszukiwania efektywnych metod prognozowania godzinowego zapotrzebowania na energię elektryczną dla odbiorców końcowych.

Cel i zakres pracy oraz opis metody badań

Celem pracy była analiza wpływu różnych metod wstępnego przetwarzania danych wejściowych na jakość prognoz godzinowego zapotrzebowania na energię elektryczną opracowanych metodami regresyjnymi.

Cel pracy zrealizowano na podstawie badań własnych wykonanych w rozdzielni nN, zlokalizowanej na terenie nowoczesnej ubojni drobiu w południowej części Małopolski. Badania polegały na ciągłym pomiarze zużycia energii elektrycznej w dwóch niezależnych obwodach zasilających obiekt (linia produkcyjna, chłodnia), które następnie były rejestrowane w 15-minutowych interwałach czasu. Do pomiarów, które realizowano od grudnia 2011 r. do września 2012 r., wykorzystano przenośny analizator CIR-e3 oraz AS-3 Plus. W wyniku przeprowadzonych badań zgromadzono obszerną bazę danych, złożoną z ok. 26 tysięcy rekordów, zawierających informacje dotyczące zużycia energii elektrycznej w poszczególnych okresach czasu. Zgromadzone informacje podzielono na dwa zbiory: uczący i testowy. Do pierwszego z nich zakwalifikowano wszystkie pomiary wykonane do

końca lipca 2012 roku. Najnowsze obserwacje wykorzystano do testowania budowanych modeli prognostycznych. W pierwszym etapie badań ustalono listę potencjalnych zmiennych, objaśniających godzinowe zapotrzebowanie na energię elektryczną, a następnie sprawdzano, jak różne metody wstępnego przetwarzania danych wpłyną na jakość opracowywanych modeli.

Modele prognostyczne opracowywano czterema metodami:

Metoda 1 – budowa modeli prognostycznych w oparciu o godzinowe zużycie energii elektrycznej w całym zakładzie;

Metoda 2 – budowa odrębnych modeli prognostycznych w oparciu o godzinowe zużycie energii elektrycznej w obwodzie linii produkcyjnej i chłodni, a następnie wyznaczanie zapotrzebowania na energię dla całego zakładu;

Metoda 3 – budowa modeli prognostycznych godzinowego zapotrzebowania na energię elektryczną w oparciu o 15-minutowe zużycie energii elektrycznej w całym zakładzie;

Metoda 4 – budowa odrębnych modeli prognostycznych godzinowego zapotrzebowania na energię elektryczną w oparciu o 15-minutowe zużycie energii elektrycznej w obwodzie linii produkcyjnej i chłodni, a następnie wyznaczanie zapotrzebowania na energię dla całego zakładu.

Jakość budowanych modeli oceniano na podstawie:

- średniego względnego błędu prognozy (*MAPE* – *Mean Absolute Percentage Error*):

$$MAPE = \frac{1}{n} \cdot \sum_{t=1}^n \frac{|E_t - E_t^*|}{E_t} \cdot 100\% \quad (1)$$

- sumarycznych wartości rzeczywistej ilości energii bilansującej (ΔESR):

$$\Delta ESR_t = \sum_{t=1}^n |E_t^* - E_t| \quad (2)$$

gdzie:

E_t – rzeczywista ilość zużytej energii elektrycznej w godzinie t ,

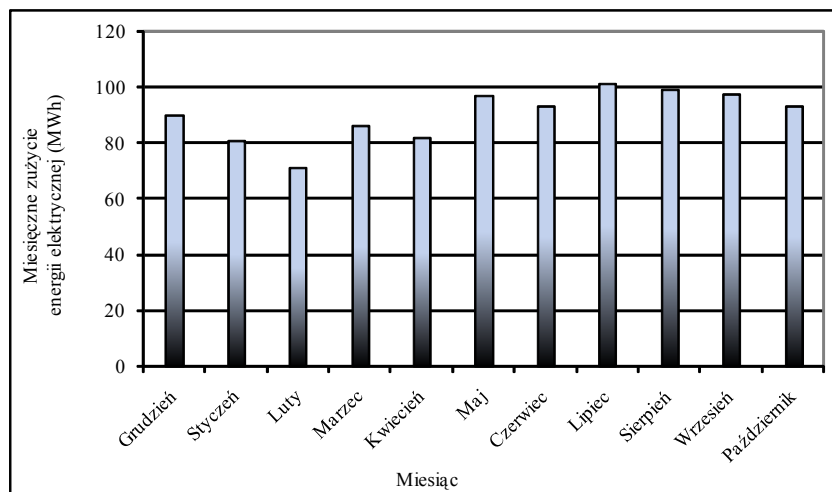
E_t^* – prognozowane zużycie energii elektrycznej w godzinie t ,

n – numer ostatniej obserwacji zmiennej prognozowanej.

Wyniki badań

Celem budowanych modeli regresyjnych było wyjaśnienie zmian zachodzących w godzinowym zapotrzebowaniu na energię elektryczną dla badanego obiektu na podstawie dostępnych zmiennych objaśniających. Modele te umożliwiają nie tylko prognozowanie poziomu badanego zjawiska, ale również ocenę wpływu zmiennych objaśniających na zmienną objaśnianą.

W pierwszym etapie badań dokonano analizy zmienności obciążeń elektroenergetycznych dla całego zakładu w celu ustalenia potencjalnej listy zmiennych objaśniających.



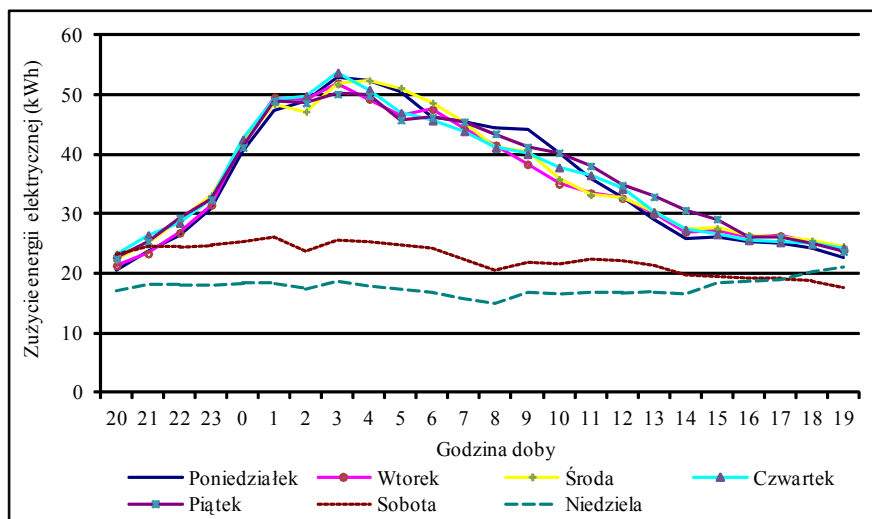
Rysunek 1. Miesięczne zużycie energii elektrycznej
Figure 1. Monthly consumption of electric energy

Na rysunku 1 jest przedstawiona miesięczna zmienność zużycia energii elektrycznej w badanym obiekcie. Ponieważ badania nie były prowadzone we wszystkich miesiącach roku, a wyznaczony współczynnik zmienności analizowanych wskaźników był na niewysokim poziomie wynoszącym 11%, nie zdecydowano się na wprowadzenie zmiennej objaśniającej w postaci zakodowanej nazwy miesiąca do budowanego modelu.

Analizując zużycie energii dla krótszych okresów czasu, zauważono występowanie wyraźnych wahań o cyklu tygodniowym, jak i dobowym. Rysunek drugi obrazuje zmiany zapotrzebowania na energię w poszczególnych dniach tygodnia.

W dni robocze widać wyraźny wzrost dobowego zużycia energii elektrycznej względem dni wolnych od pracy od 40 do 60%. Zmienia się również kształt profilu obciążenia. Dla dni, w których odbywała się produkcja, wzrost obciążenia obserwowano już od godziny 20:00 dnia poprzedniego. W czasie prowadzenia uboju maksymalne zużycie energii przypadło pomiędzy godziną 3:00 a 5:00 i było ono powodowane łączną pracą urządzeń linii uboju oraz schładzaniem do temperatury przechowalniczej gotowego produktu. Następnie rejestrowano powolny spadek godzinowego zapotrzebowania na energię do poziomu ok. 25 kWh, który był osiąganym około godziny 14:00. W kolejnym okresie trwającym do godziny 20:00 nie obserwowano znaczących zmian w zapotrzebowaniu na energię elektryczną. Głównymi odbiornikami pracującymi w tym czasie były urządzenia wykorzystywane do utrzymania wymaganej temperatury w chłodniach.

Dla dni poprzedzających dzień wolny od pracy godzinowe zużycie energii nieznacznie się zmieniało, zmniejszając swoją wartość od 25 do 18 kWh. Przez większość czasu trwania doby świątecznej zużycie energii utrzymywało się na poziomie ok. 18 kWh i było związane z pracą urządzeń chłodni. Od godziny 20:00 obserwowano wzrost zapotrzebowania na energię elektryczną ze względu na rozpoczynającą się produkcję dnia roboczego.

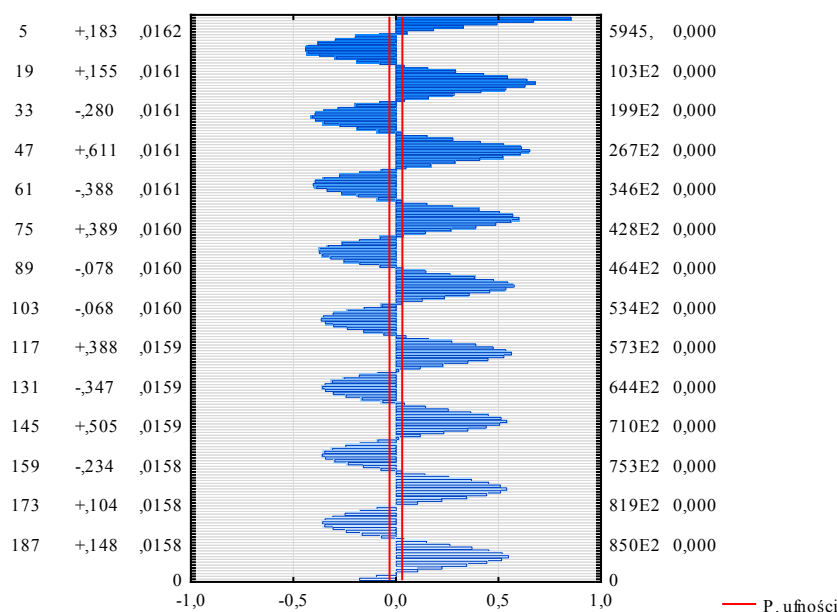


Rysunek 2. Dobowa zmienność obciążeń
Figure 2. Daily variability of loads

W celu ustalenia charakterystycznych dni tygodnia wykonano dodatkowo analizę skupień metodą k -średnich i metodą EM w programie *Statistica 10* z wykorzystaniem v -krotnego sprawdzianu krzyżowego. Wykonane analizy pokazały również, że ze względu na podobieństwo przebiegu godzinowego zapotrzebowania na energię elektryczną optymalny będzie podział dni tygodnia na 3 skupienia, tj. dni robocze, dni poprzedzające dzień wolny od pracy oraz dni wolne od pracy. Następnie posiadaną bazę danych podzielono na trzy skupienia i w ramach każdego z nich szukano zmiennej najlepiej wyjaśniającej zmienność zapotrzebowania na energię elektryczną w poszczególnych godzinach doby. Na podstawie analizy współczynników korelacji i opracowanego wykresu autokorelacji (rys. 3) ustalono, że spośród dostępnych zmiennych najlepiej zmienność zapotrzebowania na energię w grupie dni roboczych wyjaśnia zużycie energii opóźnione o 1, 2 i 24 godziny oraz średnia dobowa wartość temperatury zewnętrznej.

Niestety, ze względu na specyfikę zamawiania energii na rynku, tak małe opóźnienia nie mogą być w zastosowaniach praktycznych wykorzystane. Do dalszych analiz przyjęto więc zużycie energii opóźnione o 24 godziny, średnią dobową wartość temperatury zewnętrznej, dla których współczynniki korelacji były statystycznie istotne i wynosiły odpowiednio 0,66 i 0,24. Dla pozostałych dwóch skupień najlepszymi predyktorami okazały się zużycie energii elektrycznej w poprzednim analogicznym typie dnia oraz również temperatura zewnętrzna.

Jako pierwsze zbudowano trzy odrębne modele godzinowego zapotrzebowania na energię elektryczną dla całego zakładu, dla charakterystycznych dni tygodnia, bez wstępnego przetwarzania zmiennych objaśniających. Wartości wskaźników oceniających ich jakość zestawiono na rysunkach 4 i 5.

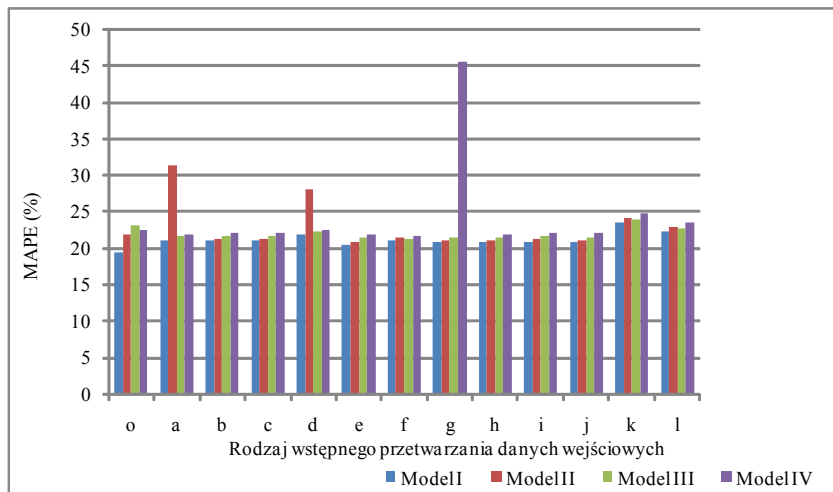


Rysunek 3. Wykres autokorelacji godzinowego zużycia energii elektrycznych w dniach roboczych dla k z przedziału $(0, 200>$

Figure 3. Graph of autocorrelation of an hour consumption of electric energy on working days for k from the range $(0, 200>$

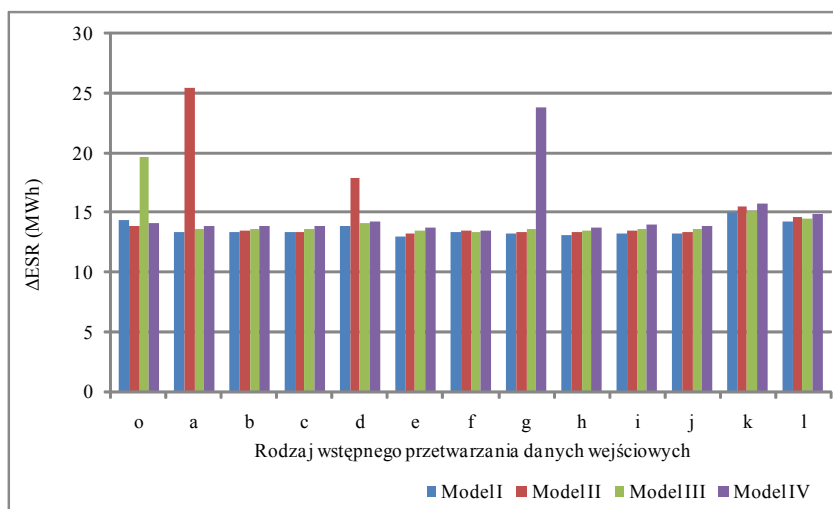
W dalszej części badań w programie *Statistica 10* dokonano następujących przekształceń opóźnionego godzinowego zużycia energii elektrycznej: o) brak przekształcenia; a) metoda średniej ruchomej dla $k=2$; b) metoda średniej ruchomej ważonej dla $k=2$ (wagi odpowiednio 0,6 i 0,4); c) metoda mediany ruchomej dla $k=2$; d) proste wyrównywanie wykładnicze; e) filtr 4253H; okno wygładzania f) Daniela; g) Tukeya; h) Hamminga; i) Parzena; j) Bertletta o rozpiętości 5; k) logarytmowanie (\ln); l) pierwiastkowanie dla stałej $C = 2$). Parametry charakteryzujące jakość modeli zbudowanych w oparciu o przekształconą zmienną wejściową przedstawia rysunek 4 i 5.

W kolejnym etapie badań budowano modele prognostyczne odrębnie dla obwodu zasilającego linię produkcyjną i urządzenia chłodni w oparciu o przekształcone w analogiczny sposób dane wejściowe. Następnie na podstawie uzyskanych prognoz wyznaczano łączne zapotrzebowanie na energię dla całego zakładu w poszczególnych godzinach doby. Wielkości charakteryzujące dobroć dopasowania opracowanych modeli zestawiono na rysunkach 4 i 5.



Rysunek 4. Wartości błędu MAPE zbioru testowego dla poszczególnych modeli prognostycznych i metod wstępnego przetwarzania danych

Figure 4. MAPE error values of a test set for particular prognostic models and methods of initial data processing



Rysunek 5. Suma energii bilansującej zbioru testowego dla poszczególnych modeli prognostycznych i metod wstępnego przetwarzania danych

Figure 5. The sum of balancing energy of a test set for particular prognostic models and methods of initial data processing

Wnioski

1. Wykonane obliczenia pokazały, że wstępne przetwarzanie danych wejściowych dla modeli regresyjnych budowanych metodą I, II oraz IV powodowało przeciętny wzrost średniego, względnego błędu prognozy (MAPE) od 1 do 2%. Wyjątek stanowiły modele budowane metodą III, dla której zaobserwowano przeciętne zmniejszenie błędu prognozy o 1%.
2. Zaobserwowano również, że budowa modeli prognostycznych w oparciu o 15-minutowe zużycie energii elektrycznej oraz budowa odrębnych modeli dla poszczególnych obwodów zasilających powodowała wzrost błędu prognoz godzinowych dla całego zakładu w porównaniu do metody I.
3. W zastosowaniach praktycznych ważniejszym parametrem oceny modeli jest sumaryczna wartość rzeczywistej ilości energii bilansującej Δ ESR. Dla większości budowanych modeli na bazie zmiennych przekształconych zaobserwowano zmniejszenie wartości wskaźnika Δ ESR względem modeli budowanych w oparciu o zmienną egzogeniczną nieprzekształconą. Największe, ponad 6% zmniejszenie wartości analizowanego wskaźnika uzyskano w modelu III dla zmiennej wejściowej wygładzonej oknem Daniela o rozpiętości 5.
4. Ze względu na najniższą wartość sumarycznej ilości energii bilansującej w zastosowaniach praktycznych powinny być jednak preferowane modele budowane na bazie szeregu czasowego godzinowego zużycia energii elektrycznej dla całego zakładu wygładzonego filtrem 4253H.

Literatura

- Gładysz, B. (2009). Metoda określania wielkości kontraktów na energię elektryczną. *Badania Operacyjne i Decyzyjne*, 3, 19-26.
- Lyp, J. (2007). Prognozy krótkoterminowe obciążeń małych odbiorców energii elektrycznej. *Polityka Energetyczna*, 10. Zeszyt specjalny, 2, 277-287.
- Małopolski, J.; Trojanowska, M. (2009a). Modele rozmyte zapotrzebowania na moc dla potrzeb krótkoterminowego prognozowania zużycia energii elektrycznej na wsi. Część I. Algorytmy wyznaczania modeli rozmytych. *Inżynieria Rolnicza*, 5(114), 177-183.
- Małopolski, J.; Trojanowska, M. (2009b). Modele rozmyte zapotrzebowania na moc dla potrzeb krótkoterminowego prognozowania zużycia energii elektrycznej na wsi. Część II. Opracowanie predykcyjnych modeli relacyjnych. *Inżynieria Rolnicza*, 5(114), 185-191.
- Nęcka, K. (2011a). Wykorzystanie modeli regresyjnych do prognozowania godzinowego zapotrzebowania na energię elektryczną w zakładzie przemysłu rolno-spożywczego. *Technika Rolnicza Ogrodnicza Leśna*, 4, 13-15.
- Nęcka, K. (2011b). Use of data mining techniques for predicting electric energy demand. *TEKA Komisji Motoryzacji i Energetyki Rolnictwa*, Vol. XIX, 230-236.
- Piotrowski, P. (2007). Prognozowanie krótkoterminowe godzinowych obciążeń w spółce dystrybucyjnej z wykorzystaniem sieci neuronowych – analiza wpływu doboru i przetworzenia danych na jakość prognoz. *Przegląd Elektrotechniczny*, 7-8(83), 40-43.
- Trojanowska, M. (2002). Krótkoterminowe prognozy sprzedaży energii elektrycznej odbiorcom w gminach wiejskich. *Inżynieria Rolnicza*, 3(36), 293-300.

Trojanowska, M.; Małopolski, J. (2009). Wykorzystanie modeli Takagi–Sugeno do krótkoterminowego prognozowania zapotrzebowania na energię elektryczną odbiorców wiejskich. *Inżynieria Rolnicza*, 1(110), 325-330.

Koncepcja wspólnego zakupu energii elektrycznej na wolnym rynku w oparciu o zasadę Grupy Bilansującej dla organów i instytucji samorządowych oraz biznesu. Kraków, Keriva, (2012). Pozyskano z: <http://www.keriva.pl/jst3.php>.

Koszty bilansowania – zmora klienta. Wirtualny Nowy Przemysł, (2012). Pozyskano z: http://www.wnp.pl/wiadomosci/koszty-bilansowania-zmora-klienta,-3359_1_0_0_0_1.html

Opis zasady TPA, podstawy prawne, wymagania. Warszawa, CEZPolska, (2012). Pozyskano z: <http://www.cezpolska.pl/pl/produkty/oferta-dla-odbiorcow-koncowych/opis-zasady-tpa-podstawy-prawne-wymagania.html>.

Pozycja konsumenta na rynku energii elektrycznej. Warszawa-Wrocław Urząd Ochrony Konkurencji i Konsumentów, (2011). Pozyskano z: www.uokik.gov.pl/download.php?plik=10178

DATA PRE-PROCESSING EFFECT ON THE QUALITY OF SHORT-TERM FORECAST OF ELECTRICITY DEMAND

Abstract. The objective of this study was to analyse the influence of different methods of pre-processing of the input data, such as moving average, exponential smoothing, filter 4253H on the quality of forecasts of hourly demand for electricity developed with regression methods. The objective of the study was carried out on the basis of own research carried out in the nN switchboard, located on the territory of a modern poultry slaughterhouse in the southern part of Małopolska region. The cluster analysis carried out with k-means and the EM method has shown that due to the similarity of the course of hourly demand for electricity division of weekdays into three days of cluster that is, working days, days preceding the days off, days off and construction of three independent models will be optimal. The total value of the actual amount of balancing energy ΔESR is the most important parameter of the models assessment in the practical applications. For majority of models constructed on the basis of the transformed variables, the decrease in the rate ΔESR towards models constructed based on exogenous not transformed variable was reported. The largest over 6% reduction in the value of the analysed indicator was obtained in model III for the input variable smoothed with 5th span Daniel window. Due to the lowest value of the total amount of balancing energy in practical applications, models built on the basis of a time series of hourly electricity consumption for the entire plant smoothed filter 4253H should be preferred.

Key words: regression model, short-term forecast, pre-processing, electricity consumption, TPA principle

Adres do korespondencji:

Krzysztof Nęcka; e-mail: krzysztof.necka@ur.krakow.pl
Katedra Energetyki i Automatykacji Procesów Rolniczych
Uniwersytet Rolniczy w Krakowie
ul. Balicka 116B
30-149 Kraków