

doi:10.15199/48.2021.01.36

Temperatura powietrza jako dodatkowa zmienna egzogeniczna w krótkoterminowym prognozowaniu modelem ANFIS obciążeń w podsystemie elektroenergetycznym

Streszczenie. Obciążenia elektryczne charakteryzują się zmiennością roczną, tygodniową i dobową. Ten typ krzywych, charakteryzujący się dobowym podobieństwem obciążeń, jest odpowiedni do prognozowania przy użyciu Adaptive Neuro-Fuzzy Inference System (model ANFIS). Badano wpływ doboru zmiennych na dokładność procesu prognozowania. Szczególną uwagę zwrócono na uwzględnienie dodatkowej zmiennej - temperatury powietrza.

Abstract. The electric loads are characterized by annual, weekly and daily variability. This type of curves, in the form of daily load similarity, is suitable for forecasting using the Adaptive Neuro-Fuzzy Inference System (ANFIS model). The influence of variable selection on the accuracy of the forecasting process has been tested. Special attention is paid to taking into consideration an additional variable – air temperature. (**Air temperature as an additional exogenous variable in short-term forecasting of loads in the power subsystem using ANFIS model.**)

Słowa kluczowe: krótkoterminowe prognozowanie obciążeń, model ANFIS, wybór zmiennych wejściowych.

Keywords: short-term load forecasting, ANFIS model, input variables selection.

Wstęp

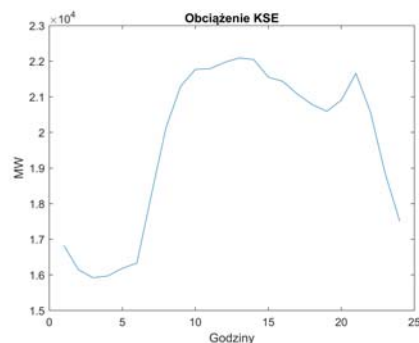
Prognozowanie obciążenia w horyzoncie krótkoterminowym wykorzystywane jest w eksploatacji oraz planowaniu systemu i podsystemów elektroenergetycznych. Na zderegulowanym rynku prognozy są potrzebne w celu planowania i realizacji operacji na rynku energii. Tego typu prognozy są potrzebne również na poziomie operatorów systemu dystrybucyjnego (OSD) i ich części czyli spółek dystrybucyjnych. Przebiegi obciążenia na poziomie spółek dystrybucyjnych lub rejonów energetycznych charakteryzują się sporymi rozrzutami, związanymi z niestabilnym zachowaniem głównych odbiorców, stąd wynikają spore wyzwania stojące przed modelami prognozowania dla regionów w obrębie spółek dystrybucyjnych.

W artykule podjęto trud wykorzystania modelu prognostycznego minimalizującego średni absolutny błąd prognozy, ale podstawowym problemem jest wybór zmiennych w zaproponowanym modelu i analiza wpływu zbioru zmiennych egzogenicznych na dokładność modelu. Czynniki, od których zależy obciążenie, można podzielić na grupę określającą rytm dnia i tygodnia, grupę określającą typ odbiorcy (bytowo-komunalni, przemysłowi itp) i grupę określającą warunki pogodowe [1, 2]. Biorąc powyższe pod uwagę można założyć, że w doborze postaci modelu prognostycznego spore znaczenie ma zmienność dobową, tygodniową i roczną o charakterze cyklicznym [3, 4]. Wykorzystuje się tę cechę w prognozowaniu krótkoterminowym. Z czynników pogodowych w polskiej strefie klimatycznej w zasadzie tylko temperatura wykazuje znaczący wpływ na przebieg obciążenia w systemie i podsystemach elektroenergetycznych [2], chociaż czynniki pogodowe wykorzystywane są w wielu ogólnie stosowanych modelach [5, 6]. Kluczowym zagadnieniem w krótkoterminowym prognozowaniu staje się dobór wielkości wejściowych modelu [7, 8]. Artykuł dotyczy tego zagadnienia i wpływu zmiennych, a dokładniej dodatkowej zmiennej w postaci temperatury otoczenia, na dokładność procesu prognozowania modelem wykorzystującym cykliczne zmiany zmiennych wejściowych.

Czynniki kształtujące zmienność obciążeń

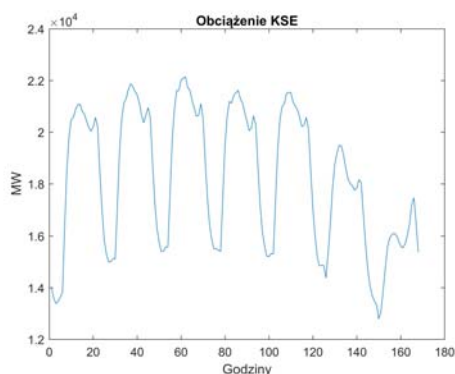
Jedną z głównych przyczyn zmienności obciążenia jest rytm życia ludności, zależny od pór roku, tradycji i zwyczajów oraz charakteru wykonywanej pracy kształtującej rytm dnia roboczego i świątecznego.

Obciążenie systemu kształtują zarówno czynniki nieprzypadkowe, jak i losowe [4]. Do czynników nieprzypadkowych zaliczamy położenie geograficzne rozpatrywanego obszaru (związany z nim klimat), zmiany momentów wschodu i zachodu słońca oraz uwarunkowania systemu elektroenergetycznego (struktura odbiorów, poziom rozwoju technologicznego, długość dnia pracy oraz system zmian roboczych) [1, 2, 4]. Czynniki losowe zależą od zmian pogody, charakteryzowanej przez temperaturę, opady, zachmurzenie, czy prędkość wiatru, od awarii urządzeń odbiorczych i wytwórczych itp. Przebieg obciążenia odbioru można rozpatrywać jako funkcję losową [1, 2].



Rys.1. Dobowy przebieg zmienności obciążeń dla KSE dla 21.08.2019 r. (Źródło: opracowanie własne na podstawie PSE-Operator S.A. <http://www.pse-operator.pl>)

O przebiegu obciążenia w ciągu doby (Rys.1) decyduje rytm życia i pracy. Od godziny 23 do 6 rano mamy do czynienia z wyraźną doliną nocną. Większość osób pracujących rozpoczyna swą działalność zawodową między godz. 6-tą a 10-tą, co powoduje duży przyrost mocy. Zmniejszenie obciążenia występuje od ok. godz.14-tej do 18-tej. Jest to spowodowane wyłączeniem części pracujących odbiorników energii elektrycznej w zakładach przemysłowych, firmach usługowych i handlu. W polskim systemie elektroenergetycznym występuje wyraźny szczyt wieczorny. Dotychczas był on spowodowany zwiększonym korzystaniem z oświetlenia i urządzeń domowych. Wraz ze zmianami technologicznymi następują ewolucyjne zmiany w dobowym przebiegu obciążenia, również np. związane z coraz popularniejszym korzystaniem z wentylacji i klimatyzacji.



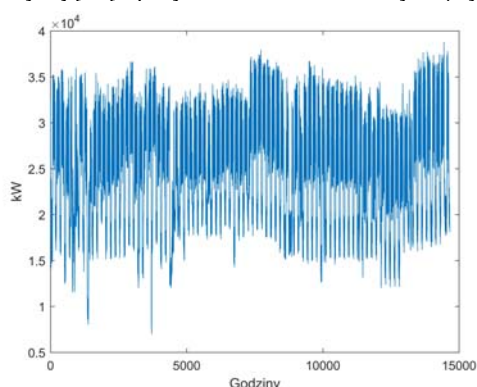
Rys. 2. Tygodniowa zmienność obciążeń Polskiego Systemu Elektroenergetycznego od poniedziałku do niedzieli dla okresu 5-11.08.2019 r. (Źródło: opracowanie własne na podstawie PSE-Operator S.A. <http://www.pse-operator.pl>)

Wyraźne zmniejszenie obciążenie w soboty i niedziele widoczne na Rys. 2 w stosunku do obciążenia podczas dnia roboczego to konsekwencja pięciodniowego tygodnia pracy w Polsce.

Prawidłowości w dobowej i tygodniowej zmienności obciążenia wykorzystuje się w procesie prognozowania. Stabilne zmiany w kolejnych tygodniach roku powodują, że nowoczesne metody prognozowania obciążeń w polskim systemie elektroenergetycznym osiągają średnie procentowe błędy absolutne poniżej 2 %. Natomiast zgoła odmienne, napotykające na spore trudności jest prognozowanie na poziomie spółek dystrybucyjnych, obejmujących niewielkie regiony. Obciążenie spółek charakteryzuje się o wiele większą niestabilnością zachowań odbiorców, a przebiegi obciążenia mogą znacznie różnić się od kształtów krajowych krzywych obciążenia. W artykule zostały wykorzystane dane o obciążeniach podsystemu elektroenergetycznego charakteryzującego się trudnym do przewidzenia przebiegiem obciążenia o sporych odchyleniach losowych.

Zbiór danych egzogenicznych i endogenicznej dla okresu historycznego

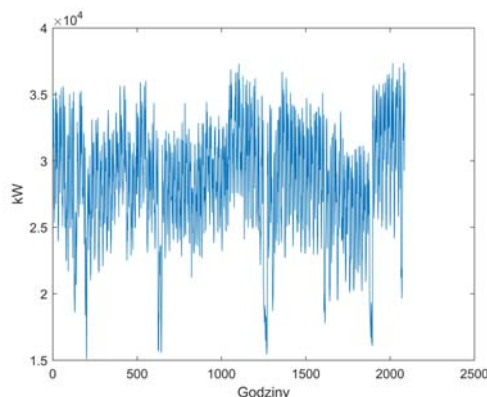
Dane statystyczne wykorzystane w publikacji dotyczą godzinowych obciążeń z dwuletniego okresu w wybranym podsystemie elektroenergetycznym. Przebiegi obciążeń charakteryzują się sporymi wahaniami losowymi (Rys.3).



Rys.3. Przebieg obciążenia wybranego podsystemu elektroenergetycznego – szereg czasowy chronologiczny, dane z okresu 22 miesięcy, czyli ok. 2 lat (Źródło: opracowanie własne)

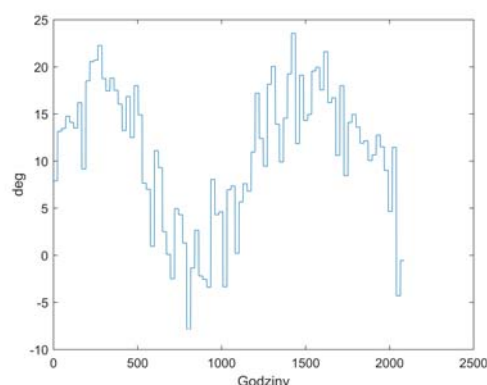
Badania wykonano zarówno dla pełnego szeregu chronologicznego, jak i ograniczonego do wybranego dnia tygodnia. Analizując zmienność tygodniową (Rys.2) można wyróżnić wyraźnie mniejsze zapotrzebowanie na energię elektryczną w niedziele i święta. Dni robocze tygodnia wykazują spore podobieństwo. Do dalszych analiz wybrano

środe, jako typowy dzień roboczy, tworząc zbiór o liczności siedmiokrotnie mniejszej od wyjściowego szeregu chronologicznego. Nowy zbiór, zbudowany tylko z 24-godzinnych szeregów kolejnych śród zaprezentowano na Rys 4.



Rys.4. Przebieg obciążenia wybranego podsystemu elektroenergetycznego w kolejne środy – dane z okresu ok. 2 lat (Źródło: opracowanie własne)

Dotychczasowe badania w polskim systemie elektroenergetycznym [1, 2] wskazują, że na przebieg obciążenia mają wpływ w pewnym zakresie warunki pogodowe, które najlepiej charakteryzuje temperatura otoczenia. Dla przebiegu obciążenia z Rys.4 zbudowano przebieg średnich temperatur dobowych odpowiadających kolejnym środom (Rys.5).



Rys.5. Przebieg średniej dobowej temperatury (°C) na obszarze podsystemu elektroenergetycznego w kolejne środy – dane z okresu ok. 2 lat (Źródło: opracowanie własne)

Można zauważyć ujemną korelację pomiędzy obciążeniem podsystemu elektroenergetycznego a temperaturą. W okresach obniżenia temperatury obserwuje się wyższe zapotrzebowanie na energię elektryczną. Czy ten wpływ jest na tyle istotny, aby do modelu prognostycznego wprowadzić dodatkową zmienną w postaci temperatury otoczenia, mają odpowiedzieć badanie zaprezentowane w artykule.

Model prognostyczny ANFIS

Podstawy teoretyczne adaptacyjnego, neuronowo-rozmytego systemu wnioskowania ANFIS opisano w [9,10]. ANFIS, który jest techniką uczenia danych, wykorzystuje zarówno sieci neuronowe, jak i logikę rozmytą. W systemie ANFIS sztuczne sieci neuronowe są wykorzystywane do określenia parametrów rozmytego systemu wnioskowania. W konstrukcji rozmytego system wnioskowania wykorzystano pakiet obliczeń inżynierskich Matlab®, MathWorks, Inc. Wynikiem obliczeń, wykonanych w pakiecie Matlab są prezentowane m.in. rysunki 6 i 7.

Modelowanie w systemie ANFIS, wykorzystujące zbiory rozmyte budowane na podstawie danych wejściowych, jest podobne do wielu innych technik identyfikacji. W zbiorze danych wejściowych wyodrębnia się dwie części. Pierwsza część jest wykorzystywana w procesie trenowania (uczenia) modelu ANFIS, którego podstawowym celem jest wyznaczenie parametrów funkcji przynależności. Natomiast w procesie weryfikacji modelu wykorzystywana jest druga część zbioru wejściowego. Wykorzystanie modelu ANFIS w prognozowaniu w elektroenergetyce opisano w wielu publikacjach, np. [11-14].

Procedury pakietu Matlab wykorzystano do zbudowania i weryfikacji modelu ANFIS, który umożliwia predykcję. W budowie modelu predykcji dobowych obciążeń w podsystemie elektroenergetycznym kluczowe znaczenie ma estymacja funkcji przynależności. Zadaniem funkcji przynależności w bloku fuzyfikacji – rozmywania jest wyznaczenie stopnia przynależności wartości ostrych do zbiorów rozmytych. Model rozmyty jest modelem logicznym. Konstruując reguły „jeśli-to”, określające zależności między zmiennymi, użyto metody Sugeno. W procesie defuzyfikacji, określając ostrą wartość wyjściową, wykorzystano metodę ważonej średniej arytmetycznej.

W procesie konstrukcji systemu ANFIS założono, że znane są wartości dobowych obciążeń $X(t)$ dla $t=1, \dots, T$ oraz średnie dobowe temperatury $\theta(t)$ dla $t=1, \dots, T$ dla wybranego dnia tygodnia (środy) w okresie ok. 2 lat. W wyniku prób wybrano kilku-wymiarowy wektor wejściowy danych uczących $w(t)=[\theta(t) \dots X(t-48) X(t-24) X(t)]$, a dane wyjściowe zbioru uczącego odpowiadają trajektorii predykcji $s(t)=X(t+24)$. Tygodniowe przesunięcie w danych wejściowych uzasadnia cykl zmienności miesięcznej każdej z analizowanych zmiennych.

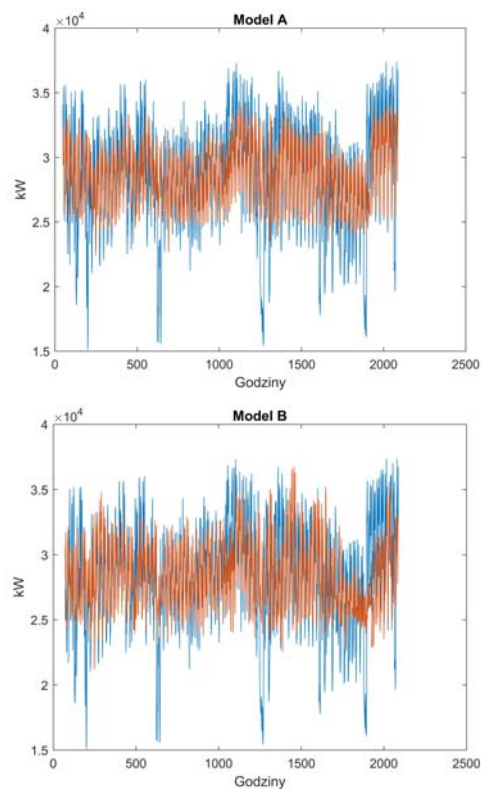
Baza danych zawiera historyczne dane z 611 dni, po 24 pomiary godzinowych obciążeń dla każdego dnia, co daje w sumie 14 664 rekordów. Dodatkową zmienną są obserwacje średnich dobowych temperatur dla powyższych 611 dni. Podzbiór danych, użyty w modelach prognostycznych, został ograniczony do danych dotyczących tylko śród, czyli wyselekcjonowano 87 dni po 24 obserwacje godzinowych obciążeń, co dało w sumie 2088 obserwacji. Dla każdego z 87 dni dodatkową zmienną jest średniodobowa temperatura otoczenia.

Jeśli rozpatruje się model o dwóch zmiennych wejściowych, to dla każdego t w przedziale $3 \div 87$, uzyskano 84 zestawy wartości zbioru danych wejściowo/wyjściowego. Pierwszą część podzbioru, czyli 42 zestawy wartości, użyto w procesie uczenia, a drugą część, również 42 zestawy wartości, w procesie weryfikacji modelu predykcji. W podobny sposób przygotowywane są podzbiory danych wejściowo/wyjściowych w przypadku większej liczby zmiennych wejściowych (egzogenicznych).

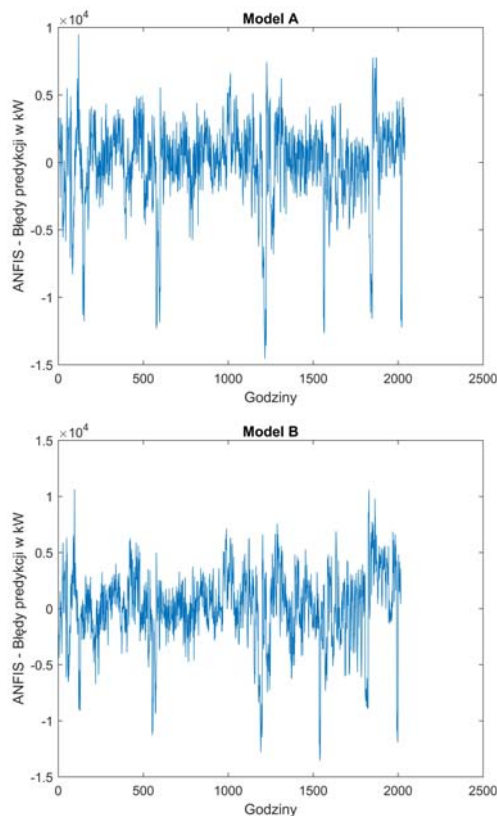
Wynikiem estymacji i weryfikacji systemu ANFIS jest konstrukcja modelu predykcji. Szeregi chronologiczne zmiennej wyjściowej $X(t)$, obciążeń w środy, charakteryzują się sinusoidalnym przebiegiem z trendem (rys.4). Średnie dobowe temperatury w środy charakteryzują się sinusoidalnym przebiegiem (rys.5).

Wyniki *ex post* prognozy obciążeń dla dwóch modeli przedstawiono na Rys.6 i 7 oraz w tabeli 1. Model A zbudowano dla dwóch przesuniętych zmiennych egzogenicznych, czyli wektor wejściowy można opisać jako $w(t)=[X(t-24) X(t)]$, a dane wyjściowe zbioru odpowiadają trajektorii $s(t)=X(t+24)$. Natomiast model B zbudowano dla dwóch przesuniętych zmiennych egzogenicznych i dodatkowej zmiennej – temperatury otoczenia, czyli wektor wejściowy stanowi $w(t)=[\theta(t) X(t-24) X(t)]$, a dane wyjściowe zbioru odpowiadają trajektorii $s(t)=X(t+24)$. Porównanie rzeczywistych przebiegów obciążenia w środy z predykcją

ex post dla dwóch modeli przedstawiono na Rys. 6, a bezwzględne błędy predykcji *ex post* obciążeń na Rys. 7. Średnie absolutne błędy procentowe MAPE zamieszczono w Tabeli 1.



Rys. 6. Porównanie rzeczywistych przebiegów obciążenia w środy (kolor niebieski) z predykcją *ex post* (kolor czerwony) dla dwóch modeli A i B



Rys.7. Bezwzględne błędy predykcji *ex post* obciążeń dla dwóch modeli A i B

Tabela 1 prezentuje błędy MAPE predykcji ex post obciążeń podsystemu elektroenergetycznego dla wybranego dnia tygodnia (śród).

Tabela 1. Błędy MAPE predykcji ex post obciążeń podsystemu elektroenergetycznego

Model	Zmienna endogeniczna i zmienne egzogeniczne	Pojemność integralna nośników informacji H_m	MAPE [%]
A	$X(t+24) X(t) X(t-24)$	0.378	8.30
B	$X(t+24) X(t) X(t-24) \theta(t)$	0.364	8.95

Jedną z wielu opisywanych w literaturze metod doboru składowych do modeli ekonometrycznych jest metoda pojemności integralnych informacji Hellwiga [7, 8]. W Tabeli 1 umieszczono wskaźnik, pojemność integralną nośników informacji H_m , charakteryzujący zbiór zmiennych egzogenicznych. Im większa wartość wskaźnika H_m , tym lepiej zbiór zmiennych egzogenicznych objaśnia zmiany zmiennej endogenicznej.

Pewnym oszacowaniem jakości modeli prognostycznych ANFIS jest porównanie ich z modelem naiwnym. W przypadku prognozy dobowego obciążenia w środę, prognoza naiwna polega na przyjęciu jako prognozy realizacji obciążeń z tygodnia poprzedzającego prognozę. Błąd dla prognozy naiwnej z całego zakresu danych, czyli 2064 obserwacji, wynosi MAPE = 9.49 %. Zarówno model A, jak i model B uzyskały mniejsze błędy MAPE.

Model A charakteryzuje się niższym błędem MAPE, a jednocześnie większą wartością pojemności integralnej nośników informacji H_m w porównaniu do modelu B. Wprowadzenie do modelu B dodatkowej zmiennej egzogenicznej w postaci temperatury otoczenia nie zwiększa wartości informacyjnej, jaką niesie ze sobą podzbiór zmiennych egzogenicznych w kontekście objaśniania zmian zmiennej endogenicznej, a jednocześnie nie prowadzi do obniżenia błędu MAPE w stosunku do modelu A.

Wnioski i uwagi

W prognozowaniu w elektroenergetyce, również w krótkoterminowych metodach prognozowania obciążeń systemu i jego podsystemów, tendencją ostatnich kilku dziesięcioleci są próby rozwijania metod opartych na sztucznej inteligencji [15-18]. Zdając się na powyższe metody zatracamy możliwość wykrywania związków przyczynowo-skutkowych w sposób jawny. Większość algorytmów uczenia maszynowego przeciętni użytkownicy traktują jako czarną skrzynkę, a badacz ma niewielki wpływ na ostateczną postać modelu. To raczej przygotowana przez badacza baza danych i jej zawartość informacyjna wpływa na końcowy rezultat.

Model ANFIS jest przydatnym narzędziem do prognozowania na podstawie szeregów czasowych charakteryzujących się okresową zmiennością. Potwierdza to porównanie wyników prezentowanych modeli z prognozą naiwną.

W artykule wykazano, że odpowiedni wybór zestawu zmiennych egzogenicznych ma istotny wpływ na dokładność modelu. Metoda Hellwiga wykorzystująca pojemności integralne nośników informacji do wyboru optymalnego zestawu zmiennych egzogenicznych pozwala jednocześnie budować modele odzwierciedlające powiązania pomiędzy zmiennymi objaśniającymi i zmienną objaśnianą.

W artykule przebadano wpływ uwzględnienia dodatkowej zmiennej egzogenicznej w postaci temperatury otoczenia na dokładność prognoz ex post obciążeń. Dla

danych przyjętych do budowy modelu ANFIS temperatura otoczenia jako dodatkowa zmienna wejściowa nie poprawia dokładności prognoz. Należy podkreślić, że jest to tylko badanie konkretnego przypadku, dla danych podsystemu elektroenergetycznego o strukturze obciążeń z przewagą sektora przemysłowego. Trudno wyciągać ogólne wnioski na podstawie badania przypadku, więc warto kontynuować badania dla szerokiego spektrum danych.

Autor: dr hab. inż. Janusz Sowiński, profesor uczelni, Politechnika Częstochowska, Wydział Elektryczny, Katedra Elektroenergetyki, Al. Armii Krajowej 17, 42-200 Częstochowa, E-mail: janusz.sowinski@pcz.pl.

REFERENCES

- [1] Dobrzańska I., Dąsał K., Łyp J., Popławski T., Sowiński J., *Prognozowanie w elektroenergetyce. Zagadnienia wybrane.*, WPCz, Częstochowa (2002)
- [2] Popławski T., Dąsał K., Łyp J., Sowiński J., *Wybrane zagadnienia prognozowania długoterminowego w systemach elektroenergetycznych*, WPCz, Częstochowa (2011)
- [3] Weron R., *Modeling and forecasting electricity loads and prices: a statistical approach*, John Wiley & Sons Ltd., England (2006)
- [4] Zalewski W., Statistical Analysis of the Load Variability in Distribution Network, *Economy and Management*, 4 (2010), 203-213
- [5] Mirasgedis S., Sarafidis Y., Georgopoulou E., Lalas D. P., Moschovits M., Karagiannis F., Papakonstantinou D., Models for Mid-Term Electricity Demand Forecasting Incorporating Weather Influences, *Energy*, 31 (2006), 208-227
- [6] Taylor W.J., Buizza R., Using weather ensemble predictions in electricity demand forecasting, *International Journal of Forecasting*, 19 (2003)
- [7] Zeliaś A., *Teoria prognozy*, PWE, Wrocław, (1997)
- [8] Cassotti M., Grisoni F., Variable selection methods: an introduction, University of Milano-Bicocca, Available online: in Tutorial 6 www.moleculardescriptors.eu (accessed on 15.04.2020)
- [9] Zadeh L.A., Fuzzy sets, *Information and Control*, 8 (1965), 338.
- [10] Jang J.-R., ANFIS: adaptive-network-based fuzzy inference system, *IEEE Transactions on Systems, Man, and Cybernetics*, 23 (1993), No. 3, 665-685.
- [11] Acakpovi A., Ternor A.T., Asabere N.Y., Adjei P., Iddrisu A.-S., Time series prediction of electricity demand using Adaptive Neuro-Fuzzy Inference Systems, *Mathematical Problems in Engineering*, 2020 (2020)
- [12] Wei D., Wang J., Ni K., Tang G., Research and Application of a Novel Hybrid Model Based on a Deep Neural Network Combined with Fuzzy Time Series for Energy Forecasting, *Energies*, 12 (2019), 3588
- [13] Jin X.B., Wang J.F., Zhang H.Y., Cao L.H., ANFIS Model for Time Series Prediction, *Applied Mechanics and Materials*, 385-386 (2013), 1411-1414
- [14] Sowiński J., Szydłowski M., Forecast of Electricity Supply Using Adaptive Neuro-Fuzzy Inference System, In Proceedings of 18th International Scientific Conference on Electric Power Engineering, Rusek S., Gono R., Eds., IEEE New York, USA (2017)
- [15] Prado F., Minutolo M.C., Kristjanpoller W., Forecasting based on an ensemble Autoregressive Moving Average - Adaptive neuro - Fuzzy inference system - Neural network - Genetic Algorithm Framework, *Energy*, 197 (2020)
- [16] Dudek G., Short-Term Load Forecasting Based on Kernel Conditional Density Estimation, *Przegląd Elektrotechniczny*, 86 (2010), No. 8, 164-167
- [17] Santra A.S., Lin J.-L., Integrating Long Short-Term Memory and Genetic Algorithm for Short-Term Load Forecasting, *Energies*, 12 (2019), 2040
- [18] Park R.-J., Song K.-B., Kwon B.-S., Short-Term Load Forecasting Algorithm Using a Similar Day Selection Method Based on Reinforcement Learning, *Energies*, 13 (2020), 2640