

doi:10.15199/48.2015.10.41

Ocena wpływu wyodrębnienia składowej stałej na jakość prognozy obciążeń elektroenergetycznych

Streszczenie. W pracy przedstawione zostało nowe podejście do prognozowania zapotrzebowania na energię elektryczną w małym systemie elektroenergetycznym w Polsce. W rozwiązaniu problemu wykorzystano dekompozycję obciążeń 24-godzinnych na składową stałą oraz zmienną. Główny wysiłek badawczy został skierowany na predykcję wektora 24-elementowego składowej zmiennej.

Abstract. This paper presents new approach to the issue of power load forecasting in a small system in Poland. The load is decomposed into two components: the mean value of the day and the variable part. The main problem is prediction of 24-hour load pattern of the variable term and this paper is concerned mainly on this task. **The new approach to the issue of power load forecasting in a small system in Poland**

Słowa kluczowe: prognozowanie obciążeń, składowa zmienna obciążenia, eksploracja danych, sieci neuronowe.

Keywords: power load forecasting, variable component, data mining, neural networks.

Wstęp

Praca przedstawia porównanie trzech metod neuronowych prognozowania obciążeń elektrycznych w małym systemie elektroenergetycznym w Polsce, wykorzystując w rozwiązaniu dekompozycję 24-godzinnego obciążenia na składową stałą dla danego dnia i składową zmienną dla każdej godziny doby.

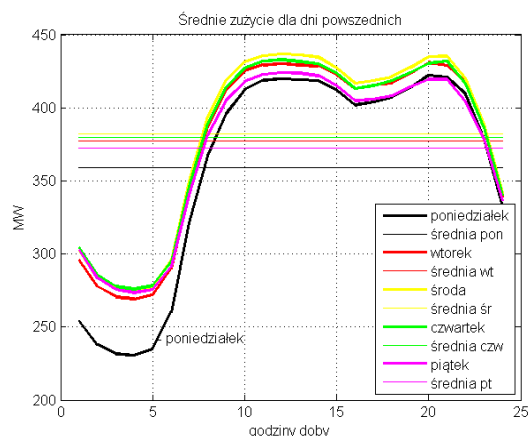
Dane wejściowe poddane eksperymentom w środowisku MATLAB stanowią wartości poboru mocy elektrycznej w małym systemie elektroenergetycznym dla każdej godziny spośród wszystkich dni z przedziału od 2 stycznia 2004 r. do 31 grudnia 2005 r. (wyrażone w megawatach). Łączna liczba dni poddanych analizie wynosi 729. Dla każdego dnia zastosowano następujące cechy diagnostyczne (atrzybuty wejściowe dla predyktora): rzeczywista moc elektryczna (24 atrybuty – jedna wielkość dla każdej godziny doby), typ dnia (jeden atrybut zakodowany binarnie: 0 – dzień roboczy, 1 – sobota lub niedziela), pora roku (2 atrybuty zakodowane binarnie: 01 – wiosna, 00 – lato, 10 – jesień, 11 – zima), dane temperatury (24 atrybuty – po jednym dla każdej godziny doby), dane prędkości wiatru (24 atrybuty – jeden dla każdej godziny doby). Bazą wyjściową do obliczeń jest więc macierz danych o wymiarach 729×75. Do uczenia wykorzystano 500 spośród dostępnych 729 zestawów danych. Pozostałe losowo wybrane 229 zestawów danych zostały przeznaczone do testowania.

Charakterystyka dobowych obciążeń

Analizowane obciążenia charakteryzują się dużą zmiennością dobową. Jest to związane ze zróżnicowanym rozkładem obciążenia w ciągu dnia. Największe zużycie energii elektrycznej występuje w godzinach 11-14 oraz w godzinach wieczornych (tzw. szczyt wieczorny w godzinach 19-21). Najniższe zużycie notuje się dla godzin nocnych pomiędzy godzinami 24 a 5 rano. Natomiast największe zróżnicowanie obciążenia występuje dla godzin porannych między godziną 5 a 9 (wzrost o 50% w stosunku do najniższego zróżnicowania).

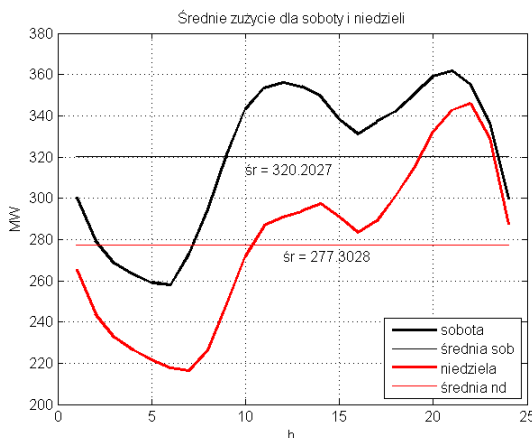
Średnie obciążenia systemu w latach 2004–2005 dla dni powszednich zostały zobrazowane na rys. 1. Spośród dostępnego zbioru danych wyodrębniona została składowa stała, będąca wartością średnią w badanym przedziale dobowym. Predykcja składowej stałej może opierać się na analizie średniej wartości obciążenia z dnia poprzedniego w przypadku dnia powszedniego. Reguła ta nie dotyczy poniedziałków, w wypadku których pod uwagę brane jest

średnie zużycie z piątków. Ma to na celu zachowanie ciągłości w strukturze danych zarówno dni powszednich jak i weekendowych. Przy zastosowaniu tej prognozy naiwnej wartość średnią obciążenia dla sobót i niedziel określa się na podstawie odpowiadających im wartości sprzed tygodnia.



Rys.1. Typowy wykres zużycia energii elektrycznej w kolejnych godzinach doby w dni powszednie.

Na rys. 2 przedstawiono typowe przebiegi obciążeń 24-godzinnych systemu w dni weekendowe z zaznaczeniem wartości średniego obciążenia, oddzielnie dla soboty i niedzieli.



Rys.2. Typowy wykres zużycia energii elektrycznej w kolejnych godzinach doby w soboty i niedziele

Charakterystyka zmienności obciążenia dla grupy dni powszednich znacznie odbiega od charakterystyki zmienności dla grupy dni weekendowych.

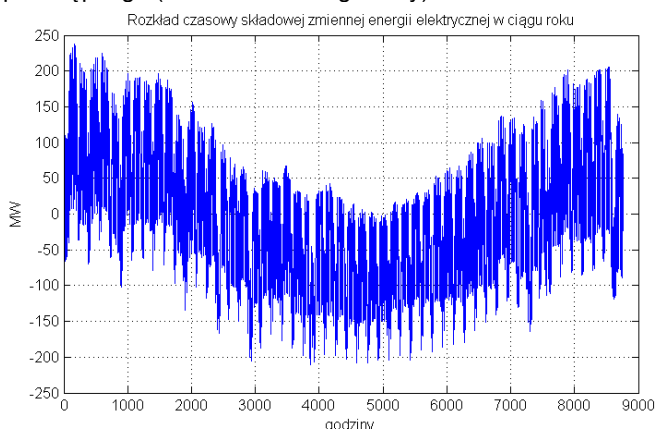
W tab. 1 zestawione zostały wartości wyrażające zawartość składowej zmiennej w obciążeniu dobowym poszczególnych dni tygodnia.

Tabela 1. Wartości zawartości składowej zmiennej w całkowitym obciążeniu oraz stosunku składowej zmiennej do średniej obciążenia w poszczególnych dniach tygodnia.

Dzień tygodnia	$\frac{ P_{zm} }{ P_{rzecz} }$ [%]	$\frac{ P_{zm} }{ P_{sred} }$ [%]
poniedziałek	20,92	17,90
wtorek	16,23	14,61
środa	15,89	14,35
czwartek	15,39	13,97
piątek	14,79	13,48
sobota	10,55	10,06
niedziela	12,70	12,18

Największą zmiennością charakteryzuje się obciążenie z poniedziałku, zwłaszcza z godzin porannych. Ma to związek z przejściem pomiędzy modelem weekendowym a powszednim oraz dynamicznym wzrostem poboru mocy w poniedziałek w godzinach porannych tj. od 5 do 9, co ilustruje rys. 1.

Na rys. 3 zilustrowano wyodrębnioną z całego obciążenia składową zmienną w ciągu jednego roku przestępnego (stanowi to 8784 godziny).



Rys.3. Wykres składowej zmiennej energii elektrycznej w badanym roku.

Zawartość składowej zmiennej w obciążeniu rzeczywistym w grupie dni powszednich zawiera się w przedziale (14,79÷20,92) %. Największą zmiennością charakteryzują się poniedziałki, co również widoczne jest na rys. 1. Spowodowane jest to przejściem z modelu obciążenia weekendowego na model powszedni. Zawartość składowej zmiennej w grupie dni weekendowych zawiera się w przedziale (10,55÷12,70) %.

Dobowy pobór mocy w dniach weekendowych cechuje się znacznie mniejszą amplitudą wahań godzinnych niż w dni robocze.

Przygotowanie danych do predykcji

Z całkowitego obciążenia 24-godzinnego, którego rozkład przedstawiony został na rys. 1 i 2 wyodrębniona została składowa zmienna P_{zm} oraz składowa stała P_{sred} dla każdego dnia. Dla uzyskania pełnego rozwiązania predykcji należy przewidzieć wartości zarówno P_{zm} jak i P_{sred} . W pracy przewidziano dwa rozwiązania predykcji składowej stałej dla danego dnia: jako znana już wartość z dnia poprzedniego (dla dnia powszedniego)

$$(1) P_{sred(d_p)} = P_{sred(d_p-1)}$$

lub wartość obciążenia odpowiedniego dnia weekendowego sprzed tygodnia (w przypadku dni weekendowych)

$$(2) P_{sred(d_w)} = P_{sred(d_w-7)}$$

Dla dni powszednich d przyjęty został średni pobór mocy z dnia poprzedniego $d-1$ przy czym w przypadku poniedziałku przyjęto obciążenie z ostatniego piątku.

Dla dni weekendowych (sobota oraz niedziela) d przyjęty został średni pobór mocy z poprzedniej soboty oraz niedzieli oznaczony we wzorze (2) wskaźnikiem $d-7$.

Drugą metodę prognozowania wartości średniej dla danego dnia oparto na zastosowaniu sieci neuronowej MLP.

Znacznie bardziej złożone jest prognozowanie 24-składników składowej zmiennej obciążenia P_{zm} . Składowa ta podlega w pracy prognozie za pomocą sztucznych sieci neuronowych MLP, RBF i SVM. Badania zostały przeprowadzone w środowisku programistycznym MATLAB. Każda sieć dokonuje prognozy 24-godzinnej z wyprzedzeniem dobowym. Wynik predykcji porównany z wartością rzeczywistą obciążenia godzinowego dla składowej zmiennej oraz stałej pozwala określić 2 rodzaje błędów bezwzględnych dla każdego rozwiązania prognozy:

- dla składowej zmiennej:

$$(3) \Delta P_{zm(d)} = P_{zm\ rzecz(d)} - P_{zm\ pred(d)}$$

- dla składowej stałej:

$$(4) \Delta P_{sred(d_p)} = P_{sred\ rzecz(d_p)} - P_{sred(d_p-1)}$$

dla d_p oznaczającego dzień powszedni oraz

$$(5) \Delta P_{sred(d_w)} = P_{sred\ rzecz(d_w)} - P_{sred(d_w-7)}$$

gdzie d_w oznacza dzień weekendowy. Stąd całkowity błąd bezwzględny predykcji jest równy:

$$(6) \Delta P_{cal(d)} = \Delta P_{zm(d)} + \Delta P_{sred(d)}$$

Podobnie można wprowadzić miarę względną błędów predykcji dla obu składowych oraz błąd względny całkowity. W przypadku predykcji składowej stałej i zmiennej obciążenia definiuje się

- błąd MAPE dla składowej zmiennej:

$$(7) MAPE_{zm} = mean \left(\frac{|\Delta P_{zm(d)}(h)|}{P_{zm\ rzecz(d)}(h)} \right)$$

- błąd MAPE dla składowej stałej:

$$(8) MAPE_{sred} = mean \left(\frac{|\Delta P_{sred(d)}|}{P_{sred\ rzecz(d)}|} \right)$$

W pierwszej kolejności zastosowano metodę naiwną predykcji składowej stałej przyjmując wartość obciążenia średniego z dnia poprzedniego (dla dni roboczych) lub sprzed tygodnia dla dni weekendowych. Średni błąd predykcji składowej stałej określony w ten sposób dla całego zbioru danych estymowany według zależności (8) wyniósł 5,69%. Jest to wartość dalece odbiegająca od oczekiwań. W celu poprawienia jakości prognozy, posłużono się predykcją za pomocą sieci neuronowej MLP, posiadającej 8 sygnałów wejściowych. Przy predykcji wartości średniej dla dnia d , odpowiadają one kolejno za następujące atrybuty: 1 – typ dnia, 2 i 3 – pora roku zakodowana binarnie, 4 – wartość składowej stałej dla dnia d , 5 – wartość średnia dla dnia $d-1$, 6 – wartość średnia dla dnia $d-2$, 7 - wartość średnia dla dnia $d-3$, 8 – wartość średnia dla dnia $d-7$. Liczba neuronów sigmoidalnych w warstwie ukrytej, została dobrana w wyniku kilku wstępnych prób i wyniosła 6. Błąd MAPE testowania, obliczony wg

zależności (9) wyniósł 1,33%, co stanowi znaczną poprawę jakości predykcji w porównaniu do predykcji naiwnej.

$$(9) \quad MAPE_{sredn} = mean \left(\frac{|P_{sred(d)} - P_{progsred(d)}|}{|P_{sred(d)}|} \right)$$

Po określeniu wartości prognozowanych składowej stałej dla poszczególnych dni i składowej zmiennej dla poszczególnych godzin definiuje się całkowity błąd względny MAPE

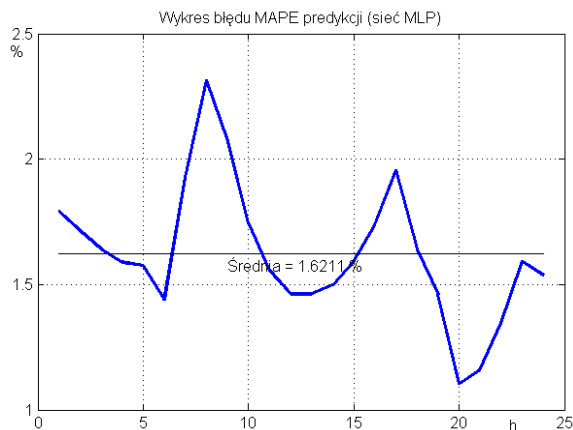
$$(10) \quad MAPE_{total} = mean \left(\frac{|P_{rzecz(d)} - P_{prognoz(d)}|}{|P_{rzecz(d)}|} \right)$$

Błąd $MAPE_{total}$ jest miarą jakości predykcji pełnego obciążenia. Będzie on porównany przy zastosowaniu różnych neuronowych modeli predykcyjnych w dalszych badaniach.

Wyniki predykcji przy użyciu sieci MLP

W przypadku prognozowania z użyciem sieci MLP warstwa wejściowa posiada 75 sygnałów odpowiadających atrybutom (podobnie w sieci RBF). Sieć ta w warstwie ukrytej zawierała 5 neuronów sigmoidalnych, których liczba została ustalona w wyniku wstępnych prób symulacyjnych. W warstwie wyjściowej sieć posiada 24 neurony liniowe. W uczeniu wykorzystano algorytm Levenberga–Marquardta, ponieważ metoda ta z uwagi na wysoką efektywność, pozwala uzyskać rozwiązanie bliskie minimum globalnemu przy najmniejszej liczbie cykli uczących. Zastosowano adaptację typu wsadowego po zaprezentowaniu sieci wszystkich danych uczących.

Na rys. 4 przedstawiono wykres błędu $MAPE_{total}$ testowania dla każdej godziny doby w sieci MLP przy wykorzystaniu predykcji składowej zmiennej obciążen z wyprzedzeniem 24-godzinnym i zastosowaniu sieci MLP do predykcji składowej stałej.



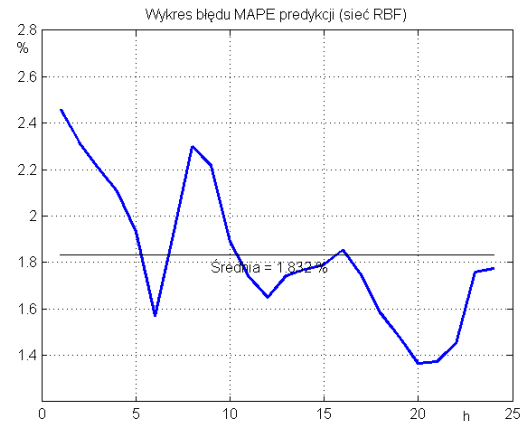
Rys.4. Wykres błędu $MAPE_{total}$ przy wykorzystaniu predykcji składowej zmiennej w sieci MLP.

Najmniejszy błąd testowania występuje dla godziny 20 i jest równy 1,10%. Natomiast największy błąd występuje dla godziny 8 i jest równy 2,31%. Średni dobowy błąd testowania dla tej sieci wynosi 1,62%. Umieszczenie maksimum lokalnego w godzinie 8 wiąże się ze stosunkowo dużym wzrostem obciążenia związanego ze szczytem porannym. Natomiast minimum lokalne w godzinie 20 wiąże się ze spadkiem wartości obciążenia oraz niskimi wahaniami obciążenia w godzinach nocnych.

Sieć RBF

Model sieci RBF zawiera jedną warstwę ukrytą z neuronami typu gaussowskiego [2,8] (odpowiada to jądro gaussowskiemu w sieci SVM) i warstwę wyjściową liniową

zawierającą 24 neurony. Odpowiadają one prognozie obciążenia dla danej godziny doby. Doborowi podlegają centra funkcji gaussowskich, ich szerokość oraz wagi liniowych neuronów wyjściowych [2]. Dobór położen centrów wykonany został przy użyciu algorytmu samoorganizacji K-means. Natomiast adaptacja wag neuronów wyjściowych została przeprowadzona przy użyciu algorytmu uczącego opartego na metodzie dekompozycji macierzy Grama według wartości osobliwych (SVD).

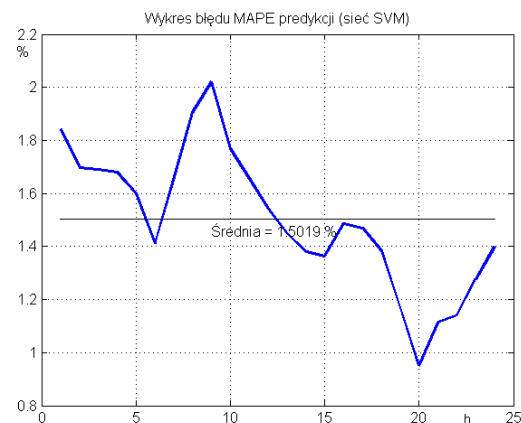


Rys.2. Wykres błędu $MAPE_{total}$ przy wykorzystaniu predykcji składowej zmiennej w sieci RBF.

Na rys. 2 przedstawiono wykres błędu $MAPE_{total}$ testowania sieci RBF przy wykorzystaniu predykcji składowej zmiennej obciążen z wyprzedzeniem 24-godzinnym. Najmniejszy błąd testowania występuje również dla godziny 20 i jest równy 1,36%. Natomiast największy błąd występuje dla godziny 1 i wynosi 2,46%. Średni dobowy błąd testowania dla tej sieci wynosi 1,83%.

Sieć SVM

Prognozowanie przeprowadzono stosując 24 sieci Support Vector Machine (SVM). Każda z nich została zastosowana do predykcji jednej godziny doby.



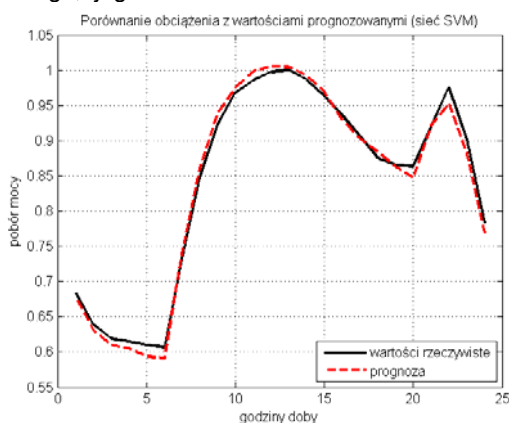
Rys.5. Wykres błędu $MAPE_{total}$ przy zastosowaniu predykcji składowej zmiennej w sieci SVM.

Sieć użyta w badaniach pracowała w trybie regresji [8] przy z góry zadanej wartości marginesu tolerancji $\epsilon = 0,03$, wybranego na podstawie kilku wstępnych prób. Współczynnik regularyzacji C, mający wpływ na złożoność sieci został ustawiony na wartość 1000.

Na rys. 5 przedstawiono wykres błędu $MAPE_{total}$ testowania sieci SVM przy wykorzystaniu predykcji składowej zmiennej obciążen z wyprzedzeniem 24-godzinnym. Najmniejszy błąd testowania występuje również dla godziny 20 i jest równy 0,95%. Natomiast największy błąd występuje dla godziny 9 i wynosi 2,02%. Średni dobowy błąd testowania dla tej sieci wynosi 1,50%.

Porównując przeprowadzone eksperymenty numeryczne, spośród przebadanych sieci najlepszymi właściwościami prognostycznymi cechuje się sieć SVM. Średni dobowy błąd $MAPE_{total}$ wyniósł tu 1,50%.

Na rys. 6 przedstawiono wykres znormalizowanego obciążenia dobowego podczas jednego dnia wraz z wartościami prognozowanymi za pomocą sieci SVM. Wykonana za jej pomocą prognoza wykazuje najwyższą zgodność z wartościami rzeczywistymi mierzoną miarą $MAPE_{total}$. Najlepsze dopasowanie prognozy oraz wartości rzeczywistych można zaobserwować dla godzin maksimum lokalnych, które przypadają dla szczytu porannego obciążenia, w tym przypadku godz. 6-9 oraz dla szczytu wieczornego, tj. godz. 19-21.



Rys.6. Wykres obciążenia dobowego wraz z wartościami prognozowanymi za pomocą sieci SVM.

Wnioski końcowe

W pracy dokonano badań jakości prognozy obciążeń 24-godzinnych systemu przy rozbiciu zadania na predykcję składowej stałej i zmiennej obciążenia elektroenergetycznego. Badania przeprowadzone w środowisku MATLAB wykazały znaczący wzrost jakości predykcji po wyodrębnieniu z całości obciążenia składowej zmiennej. Najlepsze właściwości predykcyjne mierzone błędem $MAPE_{total}$ zostały uzyskane przy wykorzystaniu sieci SVM. Średni błąd dobowy $MAPE_{total}$ wyniósł 1,50%. Stanowi to względny spadek wartości błędu o ponad 40% w

stosunku do badań prognostycznych bez podziału na wartość zmienną i składową stałą [2]. Pozostałe sieci neuronowe (MLP i RBF) pozwoliły otrzymać wyniki nieznacznie odbiegające od najlepszego. Wyniki te kształtowały się na poziomie $MAPE_{total} = 1,83\%$ dla sieci RBF i $MAPE_{total} = 1,62\%$ dla sieci MLP.

LITERATURA

- [1] T. Ciechulski: *Data clusterization in application to the analysis of load distribution in small power system*, 5th AFCEA Student Conference & 10th International Electronic and Telecommunication Conference of Students and Young Scientists SECON 2013.
- [2] T. Ciechulski, S. Osowski, *Badanie jakości predykcji obciążeń elektroenergetycznych za pomocą sieci neuronowych SVM, RBF i MLP*, Przegląd Elektrotechniczny, R. 90, Nr 8/2014, pp. 148-151.
- [3] T. Ciechulski, *Prognozowanie obciążeń dla systemu dostarczania energii z uwzględnieniem warunków pogodowych*, Logistyka, Nr 6/2014, pp. 2759-2768.
- [4] N. Kandil, R. Wamkeue, M. Saad, S. Georges, *An efficient approach for short term load forecasting using artificial neural networks*, Electrical Power and Energy Systems, vol. 28, 2006, pp. 525-530.
- [5] Z. Kowalski: *Podstawy prognozowania elektroenergetycznego*, Łódź 1980.
- [6] P. Mandal, T. Senjyu, N. Urasaki, T. Funabashi, *A neural network based several hours ahead electric load forecasting using similar days approach*, Electrical Power and Energy Systems, vol. 28, 2006, pp. 367-373.
- [7] *Matlab manual, user's guide*, MathWorks, Natick, 2002.
- [8] S. Osowski: *Metody i narzędzia eksploracji danych*, BTC 2013.

Autorzy: mgr inż. Tomasz Ciechulski, Wojskowa Akademia Techniczna, Instytut Systemów Elektronicznych, Zakład Systemów Informacyjno-Pomiarowych, ul. Kaliskiego 2, 00-908 Warszawa, E-mail: tciechulski@wat.edu.pl;
 prof. dr hab. inż. Stanisław Osowski, Politechnika Warszawska, Instytut Elektrotechniki Teoretycznej i Systemów Informacyjno-Pomiarowych, ul. Koszykowa 75, 00-661 Warszawa oraz Wojskowa Akademia Techniczna, Instytut Systemów Elektronicznych, Zakład Systemów Informacyjno-Pomiarowych, ul. Kaliskiego 2, 00-908 Warszawa
 E-mail: sto@iem.pw.edu.pl.