

## Prognozowanie zapotrzebowania mocy w KSE z horyzontem dobowym przy zastosowaniu zespołu sieci neuronowych

**Streszczenie.** Artykuł przedstawia neuronową metodę prognozowania 24-godzinnego zapotrzebowania na moc elektryczną w Krajowym Systemie Elektroenergetycznym w Polsce z wyprzedzeniem dobowym. W predykcji zastosowano zespół sieci neuronowych złożony z RBF, MLP i SVM. Wskazania każdego członka zespołu podlegają integracji tworząc końcową prognozę zapotrzebowania na moc dla kolejnych 24 godzin następnego dnia. Metoda została przetestowana na danych rzeczywistych Krajowego Systemu Elektroenergetycznego wykazując znaczną przewagę nad pojedynczymi rozwiązaniami neuronowymi.

**Abstract.** The paper shows the neural method applied for 24-hour load forecasting for the next day in National Power System in Poland. Three neural networks: RBF, MLP and SVM arranged in an ensemble were used for predicting hourly load demands in the power system. Their individual predictions have been integrated into final forecast. The method was tested on the real data of Polish Power System, proving its effectiveness and significant improvement in accuracy. **The neural method applied for 24-hour load forecasting for the next day in National Power System in Poland**

**Słowa kluczowe:** prognozowanie obciążeń, krajowy system energetyczny, zespół sieci neuronowych, integracja zespołu.

**Keywords:** load forecasting, national electric power system, ensemble of neural networks, integration of ensemble.

### Statystyczna analiza danych zapotrzebowania na moc w KSE

Krajowy System Energetyczny (KSE) obejmuje wszystkie urządzenia wytwórcze i odbiorcze energii elektrycznej w Polsce. Moc zainstalowana w KSE w dniu 31.12.2017 wynosiła 43421 MW, natomiast moc osiągalna – 43332 MW [7]. Ponieważ wpływ pojedynczych zakładów przemysłowych, zwłaszcza w momentach włączania i wyłączania poboru mocy, na pracę całego systemu jest stosunkowo niewielki, system odznacza się znacznie mniejszymi wahaniami godzinowymi zapotrzebowania na moc dostarczaną, niż Lokalne Systemy Energetyczne obejmujące swoim zasięgiem pojedyncze miasta. Praca dotyczy prognozowania zapotrzebowania dobowego w Krajowym Systemie Energetycznym (KSE) i porównania z małym regionem elektroenergetycznym.

Prognozowanie krótkoterminowe zapotrzebowania na moc jest tematem wielu prac, zarówno krajowych jak i zagranicznych [2,3,4,8,9]. Prezentowane były różnorodne podejścia do rozwiązania tego problemu, prowadzące do różnych poziomów błędu. W ostatnich latach najwięcej prac dotyczy zastosowania sztucznych sieci neuronowych. Stosowane są różne rodzaje sieci: perceptron wielowarstwowy MLP, sieć RBF, SVM, sieci rekurencyjne, itp. Prezentowane dokładności w sensie błędu bezwzględnego średniego MAPE (*Mean Absolute Percentage Error*) przyjmują wartości w przedziale 2%÷3% [1,2,8,9]. Na poziom błędu duży wpływ ma sposób generacji danych wejściowych, zwłaszcza metoda normalizacji oraz zastosowane rozwiązanie predyktora.

W pracy zostanie pokazane, że prognozy indywidualne wielu predyktorów pracujących w zespole po ich integracji uzyskują znacznie wyższy poziom dokładności. Zespół zaproponowany w pracy składa się z trzech rodzajów sieci: perceptronu wielowarstwowego (MLP), sieci wektorów nośnych (SVM) i sieci radialnej (RBF), przy czym każdy z nich generuje prognozę profilu 24-godzinnego obciążenia na dzień następny z określonym poziomem błędu. Ich połączenie w zespół pozwala na znaczne obniżenie tego poziomu.

Dane zapotrzebowania godzinne zostały zebrane z lat 2014-2017. Zbiór danych zapotrzebowania na moc obejmuje zatem 1461 dni, co stanowi łącznie 35064 godzin. W celu dokonania wstępnej analizy danych, poszczególne zbiory danych dobowych zostały podzielone według dni

powszednich i weekendowych. Następnie przeanalizowano wielkości statystyczne oddzielnie dla dni powszednich i weekendowych oraz różnych pór roku. Podstawowe parametry statystyczne, takie jak: wartość minimalna, maksymalna, średnia, mediana, odchylenie standardowe i zakres zmian dla dnia powszedniego, dnia weekendowego oraz uśrednionego dnia spośród wszystkich dni (z całego 4-letniego zakresu czasu) poddanych badaniu zostały przedstawione w tabeli 1.

Tabela 1. Zestawienie parametrów statystycznych mocy w KSE w latach 2014-2017 dla uśrednionego dnia powszedniego i weekendowego oraz uśrednionego dnia spośród wszystkich dni, przedstawione w MW

Parametr	Dzień powszedni	Dzień weekendowy	Wszystkie dni
	MW	MW	MW
Wartość minimalna	15130,7	14080,9	14868,6
Wartość maksymalna	21710,1	18521,1	20797,7
Średnia	19410,6	16642,6	18618,7
Mediana	20993,2	17247,9	19992,1
Odchylenie standardowe	2552,1	1566,8	2252,2
Zakres zmian	6579,4	4440,2	5929,1

Na podstawie tabeli 1 można stwierdzić, że średnia wartość zapotrzebowania dobowego dla typowego dnia powszedniego jest o 16,6% większa niż dla typowego dnia weekendowego. Z kolei wartość maksymalna jest analogicznie większa o 17,2%. Wartość odchylenia standardowego jest większa o 62,9%, a zakres zmian większy o 48,2%, co wynika z faktu większej zmienności dobowej w dni powszednie, związanej z działalnością zakładów przemysłowych o wysokim zapotrzebowaniu na energię elektryczną. Powyższe wysokie wskaźniki (62,9% oraz 48,2%), świadczące o zróżnicowaniu charakterystyk obydwu typów stwarzają konieczność rozróżnienia danych na powszednie i weekendowe w modelu predykcji dla sieci neuronowych.

W tabeli 2 przedstawiono wartości wybranych parametrów statystycznych odpowiadających podziałowi danych na 4 pory roku. Dotyczą one obu rodzajów dni (robocze i weekendowe). Dane pokazują istotne zróżnicowanie wartości poszczególnych parametrów

odpowiadające różnym porom roku. Największe wartości występują w okresie jesieni i zimy. Dotyczy to zwłaszcza wielkości zakresu zmian i odchylenia standardowego. Oznacza to konieczność włączenia pory roku do cech diagnostycznych podawanych na wejście predyktorów.

Tabela 2. Zestawienie parametrów statystycznych w KSE w latach 2014-2017 dla 4 pór roku

Parametr	Wiosna	Lato	Jesień	Zima
	MW	MW	MW	MW
Wartość minimalna	10494,0	11397,2	11199,5	11871,5
Wartość maksymalna	22975,9	23411,5	25984,1	26046,1
Średnia	17635,3	17858,8	19430,4	19561,1
Mediana	17729,9	18085,9	19616,5	19788,0
Odchylenie standardowe	2758,1	2709,3	3213,2	3048,5
Zakres zmian	12481,9	12014,3	14784,6	14174,6

Istotnym czynnikiem warunkującym jakość prognozy jest sposób normalizacji danych, pozwalający na zbalansowanie wpływu zróżnicowanych poziomów wartości zmiennych odpowiadających danym różnego typu (obciążenie w MW, kody typu dnia i pory roku binarne). W pracy analizie poddane zostaną 2 sposoby normalizacji [10]:

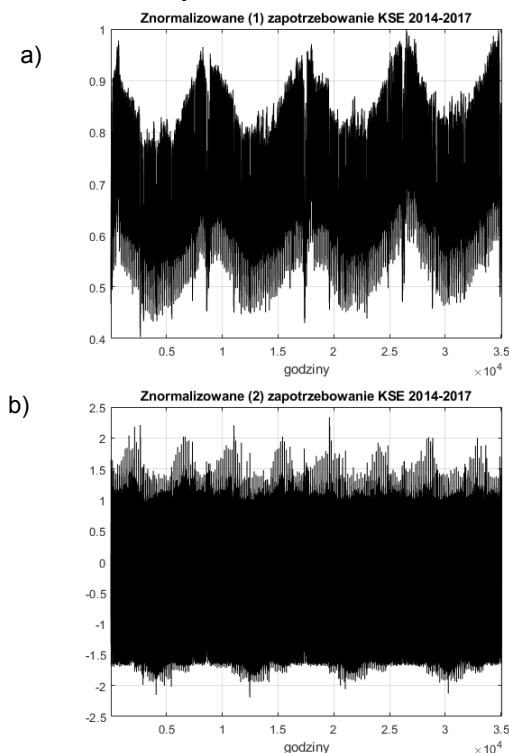
– według wartości maksymalnej:

$$(1) \quad P_{n1}(d, h) = \frac{P(d, h)}{\max(P)}$$

– według odchylenia standardowego i wartości średniej

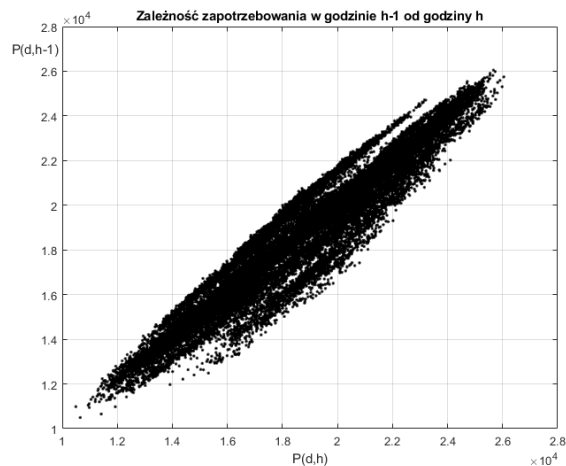
$$(2) \quad P_{n2}(d, h) = \frac{P(d, h) - \text{mean}(P(d))}{\text{std}(P(d))}$$

gdzie  $P_n(d, h)$  stanowi obciążenie znormalizowane w dniu  $d$  o godzinie  $h$ ,  $\text{mean}(P(d))$  – wartość średnia obciążenia w dniu  $d$ ,  $\text{std}(P(d))$  – odchylenie standardowe obciążeń w dniu  $d$ . Przykład graficzny ilustrujący różnice profilu obciążeń w wyniku obu sposobów normalizacji danych przedstawiono na rysunkach 1a i 1b.



Rys. 1. Obciążenie w KSE w latach 2014-2017 znormalizowane odpowiednio: a) według wartości maksymalnej, b) według wartości średniej i odchylenia standardowego

W zależności od rodzaju zastosowanej normalizacji ogólny kształt profilu obciążenia jest różny, istotne różnice obserwuje się również w zakresie wartości. Dodatkowo, w przypadku normalizacji (2), będącej standaryzacją, z wykresu usuwany jest trend wzrostowy obciążenia, dobrze widoczny na rysunku 1a. Skalę trudności predykcji bardzo dobrze ilustruje wykres zależności zapotrzebowania na moc w godzinie  $h-1$  od zapotrzebowania w godzinie następnej, czyli  $h$ . Im bardziej wykres ten jest rozproszony w relacji do diagonalnej, tym większy poziom złożoności procesu predykcji. Wykres taki dla danych KSE z lat 2014-2017 przedstawiony jest na rysunku 2.



Rys. 2. Zależność zapotrzebowania mocy KSE dla danych rzeczywistych w godzinie  $h-1$  od zapotrzebowania w godzinie  $h$ , w dniu  $d$  (dane w MW)

Istotną rolę w predykcji odgrywa korelacja między obciążeniami godzinnymi kolejnych dni tygodnia. Jej analiza pozwala na podjęcie decyzji o włączeniu do zbioru atrybutów wejściowych określonych danych jako cech diagnostycznych (atrybutów wejściowych predyktora).

Tabela 3. Współczynniki korelacji poszczególnych dni – dane oryginalne znormalizowane według wzoru (1)

$\epsilon_1$	Pon	Wto	Śro	Czw	Pią	Sob	Nie
Pon	1	0,944	0,870	0,850	0,859	0,865	0,858
Wto	0,944	1	0,876	0,874	0,870	0,870	0,833
Śro	0,870	0,876	1	0,930	0,915	0,873	0,796
Czw	0,850	0,874	0,930	1	0,960	0,901	0,801
Pią	0,859	0,870	0,915	0,960	1	0,926	0,808
Sob	0,865	0,870	0,873	0,901	0,926	1	0,898
Nie	0,858	0,833	0,796	0,801	0,808	0,898	1

Tabela 4. Współczynniki korelacji poszczególnych dni – dane znormalizowane według wzoru (2)

$\epsilon_2$	Pon	Wto	Śro	Czw	Pią	Sob	Nie
Pon	1	0,983	0,976	0,970	0,971	0,927	0,875
Wto	0,983	1	0,980	0,979	0,979	0,944	0,863
Śro	0,976	0,980	1	0,979	0,981	0,948	0,859
Czw	0,970	0,979	0,979	1	0,982	0,948	0,859
Pią	0,971	0,979	0,981	0,982	1	0,956	0,855
Sob	0,927	0,944	0,948	0,948	0,956	1	0,885
Nie	0,875	0,863	0,859	0,859	0,855	0,885	1

W tabelach 3 i 4 zestawione zostały współczynniki korelacji obciążeń godzinnych poszczególnych dni tygodnia

przy dwu różnych metodach normalizacji. Widoczne jest wysokie skorelowanie obciążeń różnych dni tygodnia. Na uwagę zwraca fakt, że największa korelacja  $\varepsilon_{pon\_wt\_1} = 0,944$  oraz  $\varepsilon_{pon\_wt\_2} = 0,983$  występuje pomiędzy poniedziałkiem a wtorkiem. Z kolei najmniejszą korelację odnotowuje się między niedzielą a pozostałymi dniami tygodnia (tabela 3), choć wartości współczynnika korelacji nawet w tym przypadku są wysokie. To spostrzeżenie obowiązuje niezależnie od sposobu normalizacji, przy czym normalizacja (2) wykazuje nieco wyższą korelację danych.

### Model predykcji

Predykcja dotyczyć będzie prognozowania obciążeń 24-godzinnych na dzień następny przy uwzględnieniu obciążeń z dnia poprzedniego, typu dnia i pory roku. W efekcie dane uczące grupujące sygnały wejściowe sieci prognostycznej dla dnia  $d$  oznaczone w postaci wektora  $\mathbf{x}(d)$  i towarzyszące im wielkości wyjściowe  $\mathbf{dest}(d)$  zostały zgrupowane w postaci macierzy  $\mathbf{X}$ . W macierzy  $\mathbf{X}$  każdy wiersz reprezentuje określony dzień. Odpowiadające im wiersze macierzy  $\mathbf{dest}$  reprezentują prognozowane 24-godzinne obciążenia rzeczywiste dnia następnego. Typowy wiersz macierzy  $\mathbf{X}$  dla dnia  $d$  ma postać:

$$\mathbf{X}(d)=[P_n(d,1), P_n(d,2), \dots, P_n(d,24), \text{typ\_dnia}, \text{pora\_roku}],$$

gdzie  $P_n(d,h)$  oznacza znormalizowane obciążenie w godzinie  $h$ -tej dnia  $d$ . Typ dnia zakodowano binarnie w postaci: 0 – dzień weekendowy, 1 – dzień roboczy, natomiast porę roku w postaci 2 bitów: 00 – lato, 01 – wiosna, 10 – jesień, 11 – zima.

Odpowiadający im wektor  $\mathbf{dest}$  reprezentuje rzeczywiste obciążenie 24-godzinne dnia następnego:

$$\mathbf{dest}(d)=[P(d+1,1), P(d+1,2), \dots, P(d+1,24)].$$

Do uczenia sieci przeznaczono 3 kolejne lata: 2014, 2015 oraz 2016, natomiast do testowania – rok 2017. Każdy sposób normalizacji ma swoje cechy charakterystyczne, wymagające innych ustawień parametrów sieci prognostycznych.

W przypadku normalizacji (1) w wyniku wielu prób wstępnych przeprowadzonych przy użyciu programu Matlab dobrano optymalne struktury sieci przedstawione poniżej:

- sieć MLP 27-12-6-24 o 24 neuronach wyjściowych liniowych, każdy odpowiedzialny za prognozę na kolejną godzinę doby. Neurony sigmoidalne w warstwach ukrytych, algorytm uczący Levenberga-Marquardta.
- 24 sieci RBF (po jednej dla każdej godziny doby): algorytm uczący OLS Grama-Schmidta. Parametr szerokości funkcji gaussowskiej  $\sigma = 0,8$ .
- 24 sieci SVM (każda odpowiedzialna za prognozę jednej godziny doby) z jądrem gaussowskim  $K(\mathbf{x}_i, \mathbf{x}) = \exp(-\gamma \|\mathbf{x} - \mathbf{x}_i\|^2)$  o współczynniku  $\gamma = 0,8$ , współczynniku regularyzacji  $C = 1500$ , tolerancji błędu  $\varepsilon = 0,001$ .

Normalizacja (2) generuje dane uczące o zupełnie innym rozkładzie wartości, stąd zastosowane sieci różniły się zarówno strukturą jak i wartościami parametrów. Sieć MLP o globalnym charakterze funkcji aproksymującej, nawet przy precyzyjnie dobranej strukturze 27-24-24, okazała się mało wydajna w stosunku do sieci typu lokalnego (RBF i SVM). Parametry tych ostatnich różniły się nieco w stosunku do normalizacji (1). Najlepsze wyniki uzyskano: dla sieci RBF o  $\sigma = 0,55$ , dla sieci SVM z jądrem gaussowskim o współczynniku jądra  $\gamma=0,85$ , współczynniku regularyzacji  $C = 20000$  i tolerancji błędu  $\varepsilon = 0,01$ . Wartości parametrów zostały dobrane metodą prób przy kombinacji różnych wartości poszczególnych elementów podlegających adaptacji. Za optymalne uznano te, które

generowały najmniejsze błędy testowania na grupie danych walidacyjnych (około 1/5 danych uczących wyłączonych z procesu uczenia).

Wyniki uzyskane przez poszczególne predyktory podlegały integracji uśredniającej [11] dla sprawdzenia jak zespół jest w stanie poprawić wyniki indywidualnych predyktorów. W wyniku takiego uśrednienia błędy popełniane przez predyktory indywidualne w poszczególnych godzinach doby podlegały częściowej kompensacji wzajemnej, redukując w ten sposób błąd całkowity zespołu. Tego typu integracja wymaga od członków zespołu podobnej jakości działania. Stąd w dalszych eksperymentach poddano badaniu różne konfiguracje zespołu.

### Wyniki statystyczne prognozowania

Każda z sieci neuronowych była trenowana na zbiorze danych uczących z lat 2014-2016, a następnie testowana na zbiorze danych z roku 2017 nie uczestniczących w uczeniu (wspólnym dla wszystkich sieci). Wyniki testowania podlegały integracji w różnej konfiguracji członków zespołu. W operacji uśredniającej brano pod uwagę jedynie członków zespołu o porównywalnym poziomie dokładności.

W tabeli 5 przedstawiono wyniki badań w postaci błędów MAPE (*Mean Absolute Percentage Error*) [10] testowania sieci neuronowych na danych nie uczestniczących w uczeniu. Odnoszą się one do danych znormalizowanych wg wzoru (1) i (2) i stanowią wartości średnie wielu prób. Niezależnie od sposobu normalizacji najmniejszym błędem MAPE wśród pojedynczych predyktorów odznaczała się sieć RBF, dla której przy pierwszym sposobie normalizacji  $MAPE_{RBFnorm1}=1,68\%$  a przy drugim  $MAPE_{RBFnorm2}=1,26\%$ . Tylko nieco gorzej wypadła sieć SVM natomiast najmniej dokładne statystyczne wyniki prognozy zostały osiągnięte przy prognozowaniu z użyciem sieci MLP.

Wyniki integracji w dużej mierze zależą od liczby zintegrowanych sieci. Przy podobnym poziomie dokładności działania wszystkich trzech predyktorów dla normalizacji (1) błąd MAPE został zredukowany do wartości  $MAPE=1,44\%$  przy trzech członkach zespołu. Uwzględnienie jedynie dwu najlepszych członków dało wynik nieco gorszy  $MAPE=1,50\%$ .

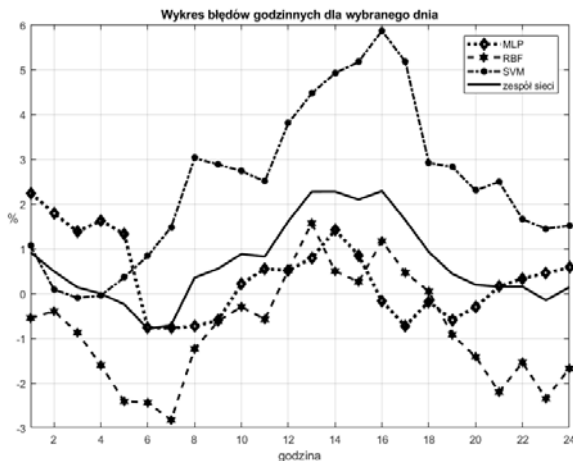
Odmienne wyniki uzyskano przy zastosowaniu normalizacji (2). Normalizacja taka zmienia radykalnie profile dzienne obciążenia. Zastosowanie aproksymatora globalnego (MLP) okazało się nieskuteczne generując dużą wartość błędu średniego w stosunku do pozostałych rozwiązań. Wyklucza to jego zastosowanie w integracji.

Tabela 5. Zestawienie błędów MAPE testowania sieci neuronowych dla danych znormalizowanych według wzorów (1, 2)

Dane wejściowe	Błąd MAPE testowania				
	MLP	RBF	SVM	zespół 2 sieci RBF i SVM	zespół 3 sieci
Normalizacja (1)	1,89%	1,68%	1,75%	1,50%	<b>1,44%</b>
Normalizacja (2)	5,66%	<b>1,26%</b>	1,63%	1,30%	-

Poza tym wystąpiła wyraźna różnica w poziomie błędów dla sieci RBF ( $MAPE=1,26\%$ ) i SVM ( $MAPE=1,63\%$ ). Stąd integracja tylko tych dwu sieci nie poprawiła wyniku końcowego sieci najlepszej. Tym nie mniej nawet w tym przypadku zespół ma sens, gdyż jednocześnie uległy zmniejszeniu błędy maksymalne predykcji (pojedyncze wartości odstające od reszty). Wyniki szczegółowe działania pojedynczych predyktorów i zespołu po odpowiedniej integracji przedstawiono w tabeli 5.

Działanie zespołu sieci najlepiej można wytłumaczyć poprzez zaprezentowanie błędów predykcji dla poszczególnych godzin wybranej doby, jak to przedstawiono dla normalizacji (1) na rysunku 3. Dla różnych godzin doby błędy poszczególnych predyktorów mają różny znak co daje możliwość wzajemnej kompensacji, redukując w efekcie finalny błąd prognozy. Widoczne jest również zmniejszenie maksymalnych odchyłek od wartości średniej w przypadku zastosowania zespołu.



Rys.3. Wykres błędów predykcji dla 24 godzin wybranego dnia (24.01.2017 roku), popełnionych przez poszczególne sieci neuronowe oraz zespół 3 sieci dla danych poddanych normalizacji (1)

Interesujące jest zestawienie średnich błędów MAPE dla różnych miesięcy roku. W tabeli 6 przedstawiono przykładowe wyniki statystyczne dla poszczególnych miesięcy roku 2017 dla danych poddanych normalizacji (1). Można zaobserwować pewną zależność przy dokładności prognozowania w poszczególnych miesiącach roku. Dla miesiąca kwietnia błąd MAPE testowania jest znacznie większy niż dla miesięcy sąsiednich (wynosi 2,95% dla sieci RBF, 2,00% dla sieci MLP oraz 1,94% dla sieci SVM).

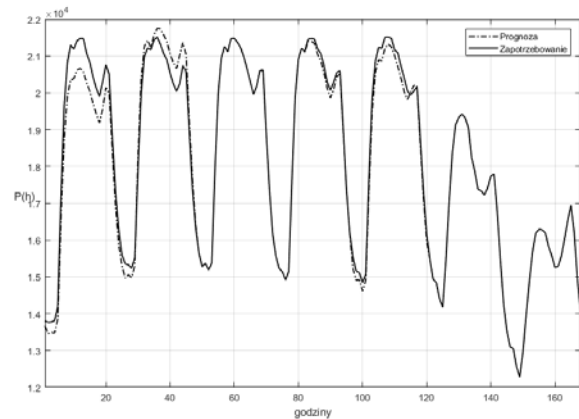
Tabela 6. Wyniki statystyczne błędu MAPE testowania dla roku 2017 dla trzech sieci oraz porównawczo dla zespołu – dane znormalizowane wg wzoru (1)

Miesiąc	MAPE				
	MLP	RBF	SVM	zespół 2 sieci RBF i SVM	zespół 3 sieci
styczeń	2,05%	1,12%	1,45%	1,15%	1,17%
luty	1,49%	1,57%	1,25%	1,09%	1,04%
marzec	1,88%	2,50%	1,83%	2,05%	1,81%
kwiecień	2,00%	2,95%	1,94%	2,30%	1,95%
maj	1,53%	0,59%	1,61%	0,96%	0,96%
czerwiec	2,63%	1,49%	2,55%	1,88%	2,04%
lipiec	1,80%	2,10%	1,91%	1,58%	1,43%
sierpień	2,01%	2,35%	2,41%	2,27%	1,82%
wrzesień	1,64%	1,05%	1,25%	1,13%	1,14%
październik	1,79%	1,30%	1,86%	1,20%	1,28%
listopad	1,57%	0,90%	1,10%	0,88%	1,02%
grudzień	2,25%	2,19%	1,90%	1,51%	1,62%
<b>Średnia</b>	<b>1,89%</b>	<b>1,68%</b>	<b>1,75%</b>	<b>1,50%</b>	<b>1,44%</b>

Dla różnych miesięcy poszczególne sieci neuronowe prognozują obciążenie dając lepsze lub gorsze wyniki według miary MAPE. Dla miesięcy: lipiec i sierpień najlepiej prognozują sieć MLP, dla miesięcy styczeń, maj, czerwiec,

wrzesień, październik, listopad – sieć RBF, natomiast dla miesięcy luty, marzec, kwiecień, grudzień – sieć SVM. Stosunkowo niewielkie różnice między sieciami występują w przypadku miesiąca lutego (0,32%) i grudnia (0,35%) natomiast największe dla miesiąca czerwca (1,14%)

Na rysunku 4 przedstawiono przykładowy wykres profilu obciążeń w wybranym tygodniu roku 2017 (od 15 maja – poniedziałek do 21 maja – niedziela). Krzywa ciągła reprezentuje rzeczywiste zapotrzebowanie mocy w systemie, natomiast kreskowana reprezentuje prognozę wskazaną przez zespół sieci neuronowych. Dane zostały poddane normalizacji zgodnie ze wzorem (2). Widoczna jest duża zgodność obu krzywych – z wyjątkiem środy i piątku.



Rys.4. Zestawienie zapotrzebowania w dniach 15-21 maja 2017 roku z wartościami prognozowanymi poprzez integrację sieci neuronowych

Najlepsze dopasowanie obydwu krzywych w analizowanym tygodniu można zaobserwować dla środy, soboty i niedzieli (17, 20, 21 maja). Natomiast dopasowanie krzywych w przypadku poniedziałku – 15 maja jest mniejsze i może wynikać z przejścia z modelu charakterystyki zapotrzebowania weekendowego na charakterystykę dnia powszedniego.

Interesujące jest porównanie dokładności działania systemu prognostycznego dla danych z KSE i małego rejonu energetycznego, analizowanego w [1]. Najlepszy uśredniony wynik prognozy zespołu dla małego rejonu był równy MAPE=1,80%. Jest to znacznie gorszy wynik niż ten uzyskany dla KSE. Wynika to głównie ze skali i specyfiki wielkości prognozowanych. Mały system energetyczny jest bardziej wrażliwy na przypadkowe zmiany obciążeń, podczas gdy w KSE obserwuje się znacznie węższe działanie kompensujące zmian obciążeń pochodzących od różnych użytkowników.

## Wnioski

Przeprowadzone badania sieci neuronowych w roli predyktorów wykazały, że połączenie wielu pojedynczych członków w zespół pozwala znacząco poprawić wynik prognozy. Dotyczy to zarówno wartości średnich błędów jak również wartości maksymalnych, które podlegają znaczącej redukcji przy zastosowaniu zespołu. Warunkiem poprawy jest porównywalny poziom dokładności działania poszczególnych członków zespołu.

Ważnym czynnikiem wpływającym na jakość działania poszczególnych predyktorów jest sposób normalizacji danych. W pracy poddano badaniom dwa rodzaje normalizacji: według wartości maksymalnych oraz standaryzację statystyczną uwzględniającą wartości średnie i odchylenie standardowe. Dane znormalizowane w różny sposób charakteryzują się różnym rozkładem statystycznym, wymagającym w końcowym projekcie sieci

prognostycznej indywidualnego podejścia do struktury sieci i doboru optymalnych parametrów.

Wyniki badań przeprowadzonych przy zastosowaniu obu rodzajów normalizacji wykazały wyższość aproksymatorów lokalnych (RBF i SVM). Oba rozwiązania generowały prognozę o najwyższym poziomie dokładności, niezależnie od rodzaju normalizacji. Sieć MLP okazała się bardzo wrażliwa na sposób normalizacji.

**Autorzy:** dr inż. Tomasz Ciechulski, Wojskowa Akademia Techniczna, Instytut Systemów Elektronicznych, Zakład Systemów Informacyjno-Pomiarowych, ul. gen. Witolda Urbanowicza 2, 00-908 Warszawa, E-mail: [tomasz.ciechulski@wat.edu.pl](mailto:tomasz.ciechulski@wat.edu.pl);  
prof. dr hab. inż. Stanisław Osowski, Politechnika Warszawska, Instytut Elektrotechniki Teoretycznej i Systemów Informacyjno-Pomiarowych, Zakład Elektrotechniki Teoretycznej i Informatyki Stosowanej, ul. Koszykowa 75, 00-662 Warszawa oraz Wojskowa Akademia Techniczna, Instytut Systemów Elektronicznych, Zakład Systemów Informacyjno-Pomiarowych, ul. gen. Witolda Urbanowicza 2, 00-908 Warszawa, E-mail: [sto@iem.pw.edu.pl](mailto:sto@iem.pw.edu.pl).

## LITERATURA

- [1] Ciechulski T., Osowski S., *Local dynamic integration of ensemble of predictors in load forecasting*, Przegląd Elektrotechniczny, nr 10/2017, ss. 57-59
- [2] Pappadimitis E., Sapatinas T., *Short-term load forecasting: the similar shape functional time-series predictor*, IEEE Transactions on Power Systems, vol. 28, no. 4, ss. 3818-3825, 2013
- [3] Marzecki J., *Metody prognozowania obciążeń terenowych stacji SN/nn*, Przegląd Elektrotechniczny, nr 4/2017, ss. 26-29
- [4] Lin C.J., Chang C.C., *Libsvm: a library for support vector machines*, <https://www.csie.ntu.edu.tw/~cjlin/libsvm>
- [5] MathWorks, *Matlab manual user's guide*, Natick, 2018
- [6] Osowski S., Siwek K., Szupiluk R., *Ensemble neural network approach for accurate load forecasting in the power system*, Applied Mathematics and Computer Science, 2009, vol. 19, No. 2, ss. 303-313
- [7] Polskie Sieci Elektroenergetyczne, <https://www.pse.pl>
- [8] Wan H., *Load forecasting via deep neural network*, Procedia Computer Science, 2017, vol. 122, ss. 308-314
- [9] Shah S., Nagaraja H. N., Chakravorty J., *Short term load forecasting model for UGVCL, MGVL, DGVCL and PGVCL using ANN*, International Journal of Recent Trends in Electrical & Electronics Eng., 2017, vol. 5, pp. 21-30.
- [10] Tan P.N., Steinbach M., Kumar V., *Introduction to Data Mining*, 2006, Pearson Education Inc., Boston.
- [11] Kuncheva L., *Combining Pattern Classifiers: Methods and Algorithms*, 2004, Wiley, New York.