

Jacek ŁYP*

Prognozy krótkoterminowe obciążeń małych odbiorców energii elektrycznej

STRESZCZENIE. Artykuł opisuje propozycję rozwiązania modelu prognozy miesięcznego zapotrzebowania na energię elektryczną małych podmiotów, opartego o technikę sztucznych sieci neuronowych. Zamieszczono opis metodologii i wyniki przeprowadzonych testów.

SŁOWA KLUCZOWE: prognozowanie, elektroenergetyka, sieci neuronowe

Wprowadzenie

Zgodnie z obowiązującymi zapisami ustawy Prawo energetyczne [1], od 1 lipca 2007 r., wszyscy odbiorcy energii elektrycznej w Polsce nabyli prawo do zakupu energii elektrycznej od wybranego przez siebie sprzedawcy. Wraz z wejściem w życie odpowiednich uregulowań (w szczególności kolejnych wariantów instrukcji ruchu i eksploatacji sieci przesyłowej i sieci rozdzielczych), okoliczność ta wzmogła zainteresowanie zagadnieniami związanymi z możliwościami uzyskania oszczędności w procesie kontraktowania i rozliczania kosztów energii elektrycznej. Zainteresowanie dotyczy zarówno odbiorców potencjalnie korzystających z nabytego prawa wyboru sprzedawcy, jak i samych sprzedaw-

* Dr inż. — Instytut Elektroenergetyki, Politechnika Częstochowska, Częstochowa.

Recenzent: prof. dr hab. inż. Eugeniusz MOKRZYCKI

ców, a dotyczy m.in. problematyki sporządzania krótkoterminowych prognoz obciążenia jako bazowych narzędzi w procesie skutecznego kontraktowania.

W przypadku odbiorców o mocy umownej powyżej 40 kW, dla których obowiązkowe są układy pomiarowe umożliwiające uzyskanie profili obciążenia z godzinową rozdzielczością, adekwatnie do dobowo-godzinowego sposobu rozliczania, przedmiotem prognoz będą podobnie jak dla dotychczasowych aktywnych uczestników rynku bilansującego, krzywe obciążenia dobowego. Nowością jest sposób traktowania odbiorców o mocy poniżej 40 kW, dla których Operator Systemu Dystrybucyjnego przydziela zestaw tzw. standardowych profili zużycia energii, które wraz z odczytami zużycia energii w okresie rozliczeniowym, służą za podstawę procedury rozliczeń.

Według przepisów, wielkość energii zamówionej można korygować nie częściej niż dwa razy do roku, co mocno ogranicza możliwości manewru. Z drugiej strony, procedury te są nowością i można domniemywać ich ewolucji w przyszłości. Faktem jest, że kluczową rolę odgrywa w tym przypadku prognoza zapotrzebowania za okres rozliczeniowy, np. miesiąc. Zainteresowane sporządzaniem takich prognoz do celów tak technicznych jak i marketingowych, mogą być podmioty pełniące rolę sprzedawcy z urzędu jak i sprzedawcy wybranego, świadczące usługi bilansowania handlowego, konsultingowe itp.

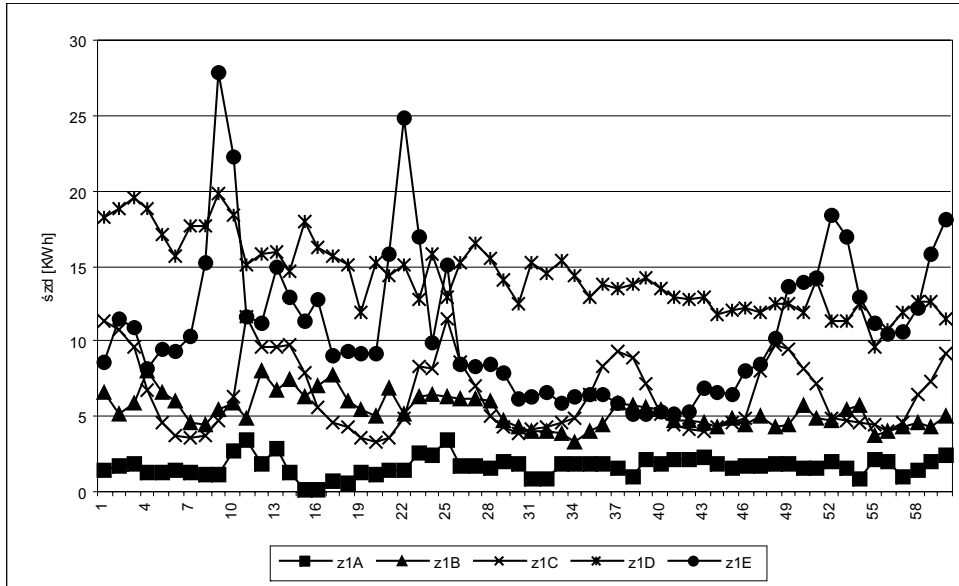
Potrzeby prognozowania miesięcznego zużycia energii dla najmniejszych odbiorców dawały znać o sobie w przeszłości. Spośród bardziej niecodziennych pobudek można przytoczyć prognozy w roli narzędzi monitorowania i wykrywania ewentualnych nieprawidłowości w przebiegu procedur akwizycji danych (np. celowe zawyżanie odczytów przed podwyżką cen itp.). Prognoza odgrywa tu odmienną rolę: nie jest przewidywaniem przyszłości, ale narzędziem weryfikacji przeszłości. Inne nietypowe przyczyny zainteresowania takimi prognozami wywodzą się z zamierzeń związanych z redukcją kosztów obsługi końcowego odbiorcy poprzez wprowadzenie systemu rozliczeń ryczałtowych.

Poniżej przedstawiono efekty prób rozwiązania metodyki takich właśnie prognoz.

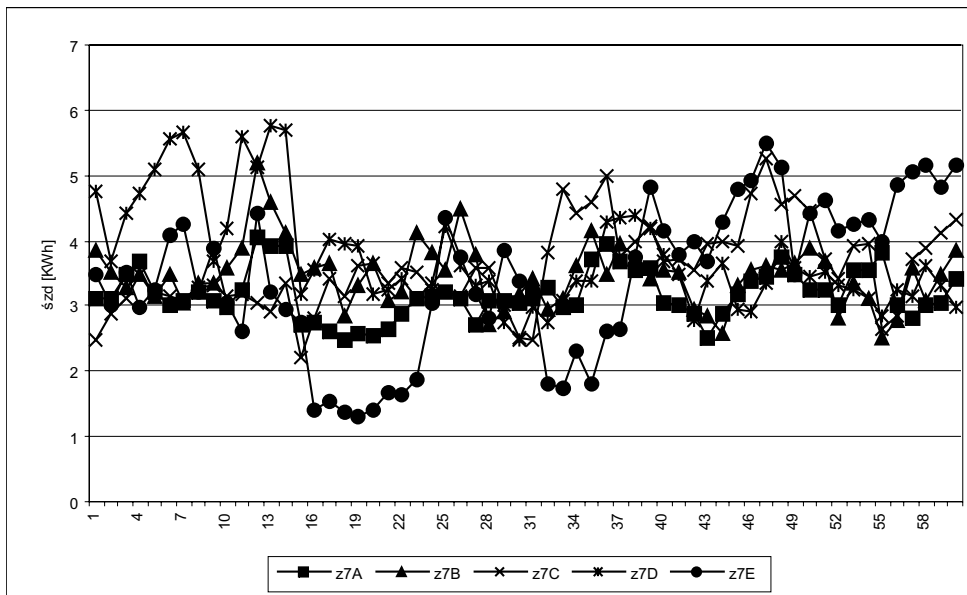
1. Dane wejściowe

Materiałem statystycznym są odczyty liczników energii 75 odbiorców wylosowanych w liczbie po 5 w 15 grupach z populacji o licznosci ponad 35 000, z pięcioletniego ciągłego okresu pomiarów. Wydzielono następujące grupy odbiorców: odbiorcy indywidualni (gospodarstwa domowe) z taryf innych niż G: jednostrefowi i dwustrefowi (grupa 1 i 2); odbiorcy z taryfy G o zużyciu rocznym do: 2, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14 MW·h (grupy od 3 do 14); odbiorcy biznesowi (grupa 15).

Na rysunku 1 zamieszczono przykłady przebiegów średniodobowego zużycia energii w kolejnych miesiącach dla odbiorców grupy 1, natomiast na rysunku 2 – dla odbiorców grupy 7.



Rys. 1. Przykłady przebiegów średniodobowego zużycia energii w kolejnych miesiącach – grupa 1
 Fig. 1. Examples of the average daily energy consumption curves in following months – 1st group



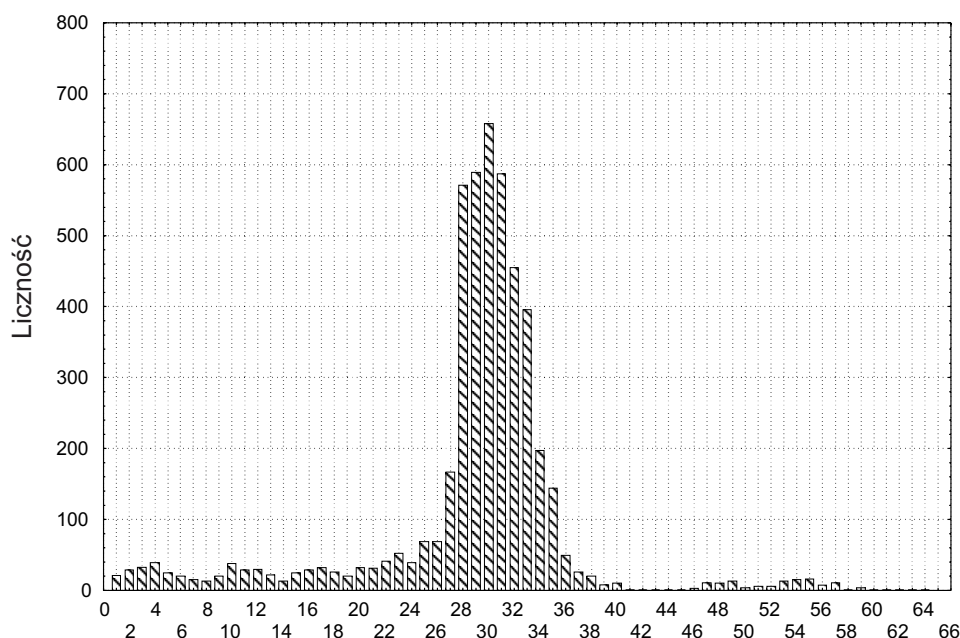
Rys. 2. Przykłady przebiegów średniodobowego zużycia energii w kolejnych miesiącach – grupa 7
 Fig. 2. Examples of the average daily energy consumption curves in following months – 7th group

1.1. Sprowadzenie danych do jednolitej podstawy czasowej

Wartości zużycia w pierwotnym zbiorze danych wejściowych dotyczą niejednakowych okresów czasu co wprowadza dodatkowe zakłócenia. Dla odbiorcy, który (przykładowo) przez luty i marzec obciążałby system stałą mocą P , przyrost zapotrzebowania na energię w marcu wynosiłby: $100\% = \frac{24 \cdot (31 - 28) \cdot P}{24 \cdot 28 \cdot P} = 10,7\%$.

Około 68% wszystkich odczytów dotyczy zakresu od 29 do 33 dni (co daje rozstęp 5 dniowy, czyli ok. 1/6 miesiąca). Jednak całkowity rozstęp okresów odczytowych wynosi 39 dni (1–40) dla grup innych niż 1 i 74 dni (1–75) dla grupy 1.

Na rysunku 3 zilustrowano histogram długości okresów odczytowych.



Rys. 3. Histogram długości okresów odczytowych

Fig. 3. Histogram of the reading cycles lengths

Konieczne jest sprowadzenie danych odczytowych do jednolitych podstaw czasowych, co wykonano według poniższego algorytmu:

- ✧ dla każdego okresu odczytowego obliczono zużycie średniodobowe energii,
- ✧ wyznaczono ciąg zużyć dobowych dla wszystkich dni danych danego odbiorcy, za które przyjęto zużycie średniodobowe wzięte z odpowiedniego dla danego dnia okresu odczytowego,
- ✧ wyznaczono wartości zużyć średniodobowych w kolejnych miesiącach kalendarzowych; wartości dla skrajnych miesięcy odrzucono jeśli zostały wyznaczone na podstawie mniej niż 14 dni.

2. Predyktory neuronowe

2.1. Opis modeli

Przeprowadzono próby zastosowania technik neuronowych do szeregów w postaci oryginalnej i przetworzonych na kilka różnych sposobów. Do przetworzenia użyto: dekompozycję, długoterminowy trend liniowy i składową sezonową, oraz różnicowania.

Dołączono zestaw prognoz uzyskanych z modeli uproszczonych, dalej nazywanych modelami kontrolnymi.

1. Model kontrolny oparty na średniej ruchomej (*śred*) – wykonano prognozę wygasłą poprzez ekstrapolację w przyszłość średniej z T ostatnich obserwacji prognozowanego procesu. Przyjęto okres średniej ruchomej T o długości 24 obserwacji.

$$\hat{A}_t = \frac{1}{T} \sum_{i=1}^T A_{t-i} \quad (1)$$

2. Model kontrolny oparty na trendzie liniowym i funkcji okresowej (*lincos*) – wykonano prognozę wygasłą, poprzez ekstrapolację dopasowanego trendu liniowego i funkcji cosinusoidalnej.

$$\hat{A}_t = b_1 \cdot t + b_2 + c_1 + c_2 \cos\left(t \cdot \frac{2\pi}{12} - c_3\right) \quad (2)$$

gdzie: b_1, b_2 – parametry trendu liniowego,
 c_1, c_2, c_3 – parametry funkcji okresowej.

3. Model kontrolny oparty na trendzie liniowym i wielomianie trygonometrycznym k -tego rzędu (*lintrik*) – wykonano prognozę wygasłą, poprzez ekstrapolację dopasowanego trendu liniowego i wielomianu trygonometrycznego k -tego rzędu.

$$\hat{A}_t = b_1 \cdot t + b_2 + W(t) \quad (3)$$

gdzie: b_1, b_2 – parametry trendu liniowego,
 $W(\tau)$ – wielomian trygonometryczny:

$$W(\tau) = a_0 + \sum_{i=1}^k c_i \cos\left(\frac{2\pi i}{t-1}(\tau-1)\right) + \sum_{i=1}^k d_i \sin\left(\frac{2\pi i}{t-1}(\tau-1)\right) \quad \tau = 1, \dots, t \quad (4)$$

$(t-1)$ – numer ostatniej znanej realizacji procesu,
 k – rząd wielomianu,
 c_0, c_i, d_i – współczynniki Eulera-Fouriera ($i = 1, \dots, k$).

Eksperymentalnie dobrano rząd wielomianu: $k = 5$.

4. Model oparty na nieprzetworzonym szeregu czasowym (*pros*):

$$\hat{A}_t = f(A_{t-1}, A_{t-2}, \dots, A_{t-k}) \quad (5)$$

k – rząd szeregu czasowego.

5. Model oparty na szeregu reszt $\{R_{\tau}\}$ uzyskanych z usunięcia trendu liniowego i sezonowości o postaci funkcji okresowej (*lincos_**):

$$\hat{A}_t = b_1 \cdot t + b_2 + c_1 + c_2 \cos\left(t \cdot \frac{2\pi}{12} - c_3\right) + \hat{R}_t \quad (6)$$

$$\hat{R}_t = f(R_{t-1}, R_{t-2}, \dots, R_{t-k}) \quad (7)$$

6. Model oparty na szeregu reszt $\{R_{\tau}\}$ uzyskanych z usunięcia trendu liniowego i sezonowości o postaci funkcji okresowej (2) (*lintri_**):

$$\hat{A}_t = b_1 \cdot t + b_2 + W(t) + \hat{R}_t \quad (8)$$

$$\hat{R}_t = f(R_{t-1}, R_{t-2}, \dots, R_{t-k}) \quad (9)$$

7. Model oparty na szeregu $\{D_{\tau}\}$ uzyskanym z różnicowania szeregu pierwotnego (*diff_**):

$$\hat{A}_t = A_{t-1} + \hat{D}_t \quad (10)$$

$$\hat{D}_t = f(D_{t-1}, D_{t-2}, \dots, D_{t-k}) \quad (11)$$

Parametry dekompozycyjne modeli (tj. parametry trendów liniowych, funkcji okresowych (2) i wielomianów (4)) estymowano krokowo, wykorzystując każdorazowo właściwą dla danego kroku liczbę realizacji.

Parametry właściwych modeli predykcyjnych (oznaczanych we wzorach (5), ..., (11) jako $f(\dots)$) wyznaczono w toku wykonania procedury uczenia sieci neuronowych na jednorazowo wyznaczonym ciągu realizacji zwanym zbiorem uczącym.

Dla szeregu reszt (7), (9) lub różnic (11) skonstruowano odpowiednio zbiory par wektorów: wektor wejściowy o k składowych i wektor wyjściowy o jednej składowej. W ujęciu matematycznym, funkcje (7), (9), (11) realizują autoregresję zmiennej będącej składową wektora wyjściowego.

2.2. Podział danych

Z każdego zbioru par wektorów, wydzielono dane o wektorach wyjściowych dotyczących okresu od stycznia 2002 włącznie. Tę część danych traktowano jako dane testowe. Nie brały one udziału w uczeniu sieci. Pozostałe dane (zbiór uczący), w sposób losowy dzielono na dwa podzbiory: zbiór treningowy i szacujący (walidacyjny). W przypadku dużej rozbieżności między wartościami błędu sieci dla części treningowej i szacującej przeprowadzano od nowa procedurę losowego podziału przypadków na część treningową i szacującą.

2.3. Typy sieci neuronowych

Użyto następujących typów sieci neuronowych:

- ✧ sieć liniowa (*lin*) [3, 5],
- ✧ sieć o radialnych funkcjach bazowych (*rbf*) [2, 4],
- ✧ wielowarstwowy perceptron z jedną warstwą ukrytą (*mlp3*) [2, 3, 5],
- ✧ wielowarstwowy perceptron z dwoma warstwami ukrytymi (*mlp4*) [2, 3, 5].

2.4. Uczenie

Procedurę uczenia wykonywano dla każdego odbiorcy osobno bez używania danych testowych. Niezależnie od rodzaju sieci, procedura była wielokrotnie powtarzana w celu wyznaczenia optymalnych parametrów sieci.

W przypadku sieci liniowych (*lin*), które pod względem matematycznym są modelem autoregresji liniowej, optymalizacji podlega jedynie rząd autoregresji, zaś samo „uczenie” polega na wyznaczeniu współczynników autoregresji, które w sieci pełnią rolę wag neuronów w warstwie wyjściowej. Wagi te są wyznaczone analitycznie metodą pseudoinwersji [2, 4]. Kontrola wartości błędów sieci liniowej uzyskanych dla zbiorów treningowego, szacującego i testowego, pozwala na dobranie odpowiedniego rzędu autoregresji.

Sieci *rbf* [2, 4] uczone są dwuetapowo. Pierwszy etap polega na wyznaczeniu tzw. centrów funkcji bazowych (realizowany metodą *k*-średnich [2]), drugi zaś polega na wyznaczeniu wag neuronów liniowej warstwy wyjściowej (realizowany analitycznie metodą pseudoinwersji [2, 4]).

Optymalizowane są zarówno rozmiar wejścia sieci (czyli rząd autoregresji) i wewnętrzna architektura sieci (czyli ilość neuronów w warstwach ukrytych).

Uczenie sieci *mlp3* i *mlp4* realizowano algorytmem „wstecznej propagacji błędów” [2, 3, 5], z kryterium tzw. „wcześniejszego stopu” [2, 3]. Optymalizowane tu są zarówno rozmiar wejścia jak i struktura warstw ukrytych.

2.5. Uwagi o uczeniu i efektywności sieci

Z uwagi na przeglądowy charakter eksperymentu intensywność przeprowadzonego cyklu procedur uczenia nie była forsowana. Przy indywidualnym potraktowaniu poszczególnych przypadków i bardziej czasochłonnym procesie poszukiwania optymalnych parametrów sieci, można uzyskać lepsze wyniki, zwłaszcza dla sieci *mlp*.

Fakt niestosowania w uczeniu kilkunastu ostatnich realizacji prognozowanego procesu, może pozbawić sieć istotnych informacji o najaktualniejszych wzorcach. W normalnym trybie zbiór testowy powinien być losowany, podobnie jak część treningowa i szacująca.

3. Analiza wyników

Jako bazowej miary błędu użyto względnego błędu procentowego.

3.1. Oznaczenia

N – liczność zbioru błędów,

PE – względny błąd procentowy (*Percentage Error*) $PE = 100$ (prognoza – realizacja) / realizacja,

$APE = |PE|$ – moduł błędu (*Absolute PE*),

$MPE = \frac{1}{N} \sum PE$ – średni błąd (*Mean PE*),

$MAPE = \frac{1}{T} \sum APE$ – średni moduł błędu (*Mean APE*),

$RMSPE = \sqrt{\frac{1}{N} \sum PE^2}$ – błąd średniokwadratowy (*Root Mean Square of PE*),

$SDPE = \sqrt{\frac{\sum (PE - MPE)^2}{N - 1}}$ – odchylenie standardowe błędu procentowego (*Standard Deviation of PE*),

$PAPE$ – błąd prawdopodobny (*Probable APE*) – prawdopodobieństwo $P[APE \leq PAPE] = 0,68$,

$HPAPE$ – błąd wysoce prawdopodobny (*High Probable APE*) – prawdopodobieństwo $P[APE \leq HPAPE] = 0,95$.

+/- – znakiem „plus” oznaczono pozytywnie zweryfikowaną hipotezę o nieobciążoności prognozy; znakiem „minus” oznaczono odrzuconą hipotezę o nieobciążoności.

3.2. Test nieobciążoności

Nieobciążoność (tzn. zerową wartość oczekiwaną błędu), w postaci hipotezy zerowej $H_0: E\{PE\} = \mu_{PE} = 0$, weryfikuje się poprzez wyznaczenie przedziału ufności, który z prawdopodobieństwem $1 - \alpha$ zawiera „prawdziwą” wartość średnią błędu μ_{PE} [6]:

$$MPE - \frac{SDPE}{\sqrt{N}} t_{\frac{\alpha}{2}} < \mu_{PE} < MPE + \frac{SDPE}{\sqrt{N}} t_{\frac{\alpha}{2}} \quad (12)$$

gdzie: $t_{\frac{\alpha}{2}}$ – kwantyl rozkładu Studenta o $N-1$ stopniach swobody, spełniający warunek:

$$P\{-t_{\frac{\alpha}{2}} < t < t_{\frac{\alpha}{2}}\} = 1 - \alpha \quad (13)$$

$$SDPE = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (PE_i - MPE)^2} \quad (14)$$

$$MPE = \frac{1}{N} \sum_{i=1}^N PE_i \quad (15)$$

Jeśli przedział ufności pokrywa wartość zero, nie ma podstaw do odrzucenia hipotezy H_0 na poziomie istotności $p = 1 - \alpha$ ($\alpha = 0,05$).

Pozytywną weryfikację hipotezy H_0 oznaczono znakiem „+” – tabela 1.

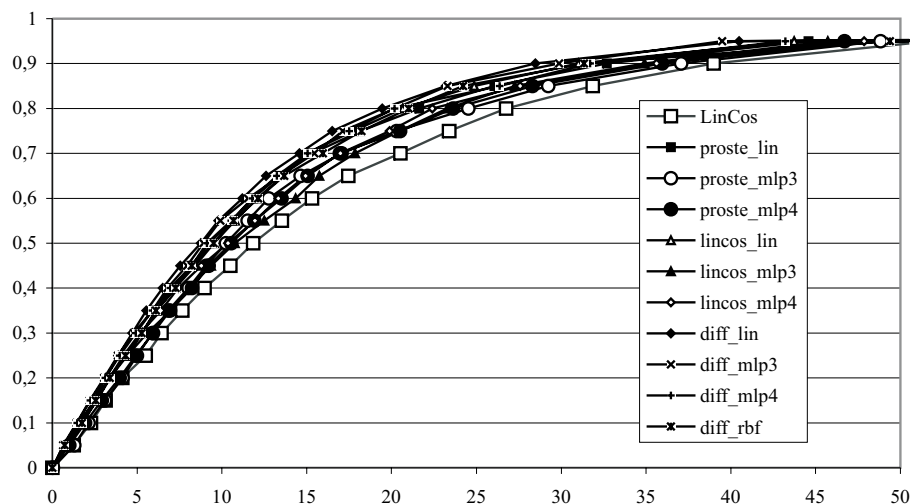
Rysunek 4 przedstawia wybrane dystrybuanty empiryczne MAPE.

Wnioski

Najlepszą jakość precyzji wydają się oferować wielowarstwowe perceptrony z jedną warstwą ukrytą (*mlp3*), jednakże pod pewnymi warunkami, których spełnienie w przypadku masowego stosowania może sprawiać trudności.

Wniosek ten postawiono w opozycji do zbiorczego zestawienia wyników, w którym osiągi *mlp3*, nie są klasyfikowane najwyżej, jednak przemawia za tym doświadczenie w pracy z sieciami neuronowymi i świadomość warunków przeprowadzonego eksperymentu.

Pełne powodzenie stosowania sieci *mlp3* wymaga zwykle indywidualnego podejścia do każdego przypadku prognozowanego procesu w dwóch etapach analizy: (1) zapewnienie reprezentatywności zbiorów treningowego, szacującego i testowego, (2) uczenie sieci z poszukiwaniem optymalnej architektury. Algorytmy automatyzujące te czynności (zwłaszcza



Rys. 4. Wybrane łączne dystrybuanty empiryczne MAPE

Fig. 4. Selected all-in empirical distributions of MAPE

TABELA 1. Zbiorcze zestawienie analizy błędów poszczególnych metod

TABLE 1. Specification of error analysis of particular methods

	MPE	MAPE	RMSPE	SDPE	PAPE	HPAPE	MinPE	MaxPE	testT
śred	3,5	20,9	33,5	33,3	21,6	58,3	-77,8	308,8	-
LinCos	2	18	28,1	28,1	19,3	51,8	-75,9	211,4	-
LinTri5	9	20,2	34,2	33	20,1	65,8	-88,2	330,6	-
pros_lin	2,4	16,5	39,6	39,5	14,3	44,6	-640,1	570,2	+
pros_rbf	6,8	21,4	38,4	37,8	20,1	69	-141,8	382,6	-
pros_ MLP3	3,6	16,9	29,2	29	16,4	48,8	-399,5	196,2	-
pros_ MLP4	3,4	16,4	26,8	26,6	16,1	46,7	-107,2	248,2	-
lincos_lin	1,1	15	28,3	28,3	14,5	43,7	-484,8	213,7	+
lincos_rbf	4	20,5	35,5	35,3	18,4	66,9	-303,9	285,2	-
lincos_ MLP3	0,6	16,8	30,7	30,7	17,1	45,7	-447,2	203,2	+
lincos_ MLP4	2,1	16,3	28,4	28,3	16,1	47,9	-341,8	215,7	-
lintri5_lin	68	187,2	445,6	440,6	129,6	807,9	-2453,1	3932	-
lintri5_rbf	8,7	20,8	34	32,9	21,3	67	-89,4	333,2	-
lintri5_ MLP3	32,1	53,3	91,6	85,9	52,3	172,9	-159,5	937,7	-
lintri5_ MLP4	30,3	48,5	89,4	84,2	49,9	138,3	-156	1039,3	-
diff_lin	1,9	13,5	22,3	22,3	13,8	40,5	-120,8	222,4	-
diff_ MLP3	3,5	13,9	23,5	23,2	14,8	39,5	-98,1	282,1	-
diff_ MLP4	2,2	14,9	26,5	26,4	14,2	43,2	-278,9	241,8	-
diff_rbf	4,6	15,2	25,5	25,1	15	49,4	-93	223,3	-

algorytmy uczenia połączonego z optymalizacją architektury) są czasochłonne i cechują się pewnym, trudnym do przewidzenia stopniem niepewności działania.

Z powodu niewielkiej liczebności rozpatrywanego ciągu realizacji procesu, reprezentatywność podzbiorów zbioru uczącego jest problematyczna. Jednak stosowanie dłuższego ciągu jest niecelowe, z uwagi na brak użytecznych wzorców we wcześniejszych latach (co wynika z trwałych zmian w charakterystykach obciążeniowych odbiorców).

Optymalizacja architektury sieci wymaga wielokrotnego powtarzania procedur uczenia i oceny dla różnych architektur. Jest to proces czasochłonny i łatwo popełnić podczas niego błąd tzw. przetrenowania sieci (zbytniego dopasowania sieci do danych).

Wobec tych spostrzeżeń, do masowego stosowania predestynowane wydają się być techniki o mniejszym koszcie obliczeniowym (chodzi głównie o stabilność procedur optymalizacji parametrów), takie jak sieci liniowe (model *diff_lin*), czy sieci o radialnych funkcjach bazowych (model *diff_rbf*).

Wyniki analizy błędów zwracają uwagę na znaczenie odpowiedniej procedury dekompozycji zadania. Najlepiej sprawdziło się różnicowanie o rzędzie jeden.

Literatura

- [1] Prawo energetyczne. Dziennik Ustaw 2006 r. Nr 89, poz. 625 z późniejszymi zmianami.
- [2] Praca zbiorowa. Prognozowanie w elektroenergetyce. Zagadnienia wybrane. Wydawnictwo Politechniki Częstochowskiej, Częstochowa 2002.
- [3] MASTERS T., 1996 — Sieci neuronowe w praktyce. WNT, Warszawa.
- [4] ORR M., 1996 — Introduction to Radial Basis Function Networks. <http://www.anc.ed.ac.uk/rbf/papers/intro.ps.gz>. Edinburgh.
- [5] TADEUSIEWICZ R. Sieci neuronowe. <http://winntbg.bg.agh.edu.pl/skrypty>.
- [6] HELLWIG Z., 1961 — Elementy rachunku prawdopodobieństwa i statystyki matematycznej. Państwowe Wydawnictwo Naukowe, Warszawa.

Jacek ŁYP

Short term load forecasts of small electric energy consumers

Abstract

The paper describes the proposal of the solutions of the monthly energy demand forecasting model for small objects based on the artificial neural networks. The paper contains description of the proposed methodology and presents the results obtained from experimental tests.

KEY WORDS: forecasting, electric power engineering, artificial neural networks