

Kazimierz DAŚAL*, Tomasz POPLAWSKI**, Karolina RUSEK***

Prognozy długoterminowe energii elektrycznej w KSE – wybrane problemy

STRESZCZENIE. Istnieje wiele czynników mających znaczenie w procesie prognozowania. Do najważniejszych zaliczyć należy wybór modelu oraz dobór zmiennych objaśniających w przypadku wielowymiarowego modelu ekonometrycznego. W artykule skoncentrowano się na doborze zmiennych, który odgrywa kluczową rolę. Spośród wielu istniejących i znanych algorytmów doboru zmiennych, w artykule przedstawiono autorsko wybrane trzy metody: metodę pojemności integralnej Hellwiga, metodę współczynnika korelacji cząstkowej oraz metodę współczynnika korelacji wielorakiej. Dla każdej z zastosowanych metod podano syntetyczny opis wyjaśniający istotę jej działania. Wybór modelu do opracowania prognoz potraktowano jako mniej istotny, ponieważ dla każdej metody doboru model jest ten sam. Obliczenia wykonano modelem MRK (Model Rozkładu Kanonicznego), zamieszczając syntetyczny opis modelu. Wykazano na rzeczywistych danych jak zastosowana metoda doboru zmiennych objaśniających wpływa na uzyskane wyniki prognozy zużycia energii elektrycznej na poziomie kraju. W przykładzie obliczeniowym dysponowano potencjalnym zbiorem z siedmioma zmiennymi objaśniającymi. W opracowaniu zawarto 14 różnych prognoz otrzymanych w wyniku zastosowania 3 metod wyboru zmiennych. Wyniki dopasowania modeli (prognoz wygasłych) oraz prognoz do 2030 roku, przedstawiono w tabelach i na wykresach. We wnioskach zawarto uwagi dotyczące prognoz długoterminowych podstawowych wielkości dotyczących Krajowego Systemu Elektroenergetycznego, mogące mieć wpływ na poprawność ocen tych prognoz. Zwrócono uwagę na problem poprawności oceny prognoz długoterminowych dotyczącej zużycia energii czy mocy szczytowych. Wskazanie prognoz

* Dr inż. – Instytut Elektroenergetyki, Politechnika Częstochowska, Zakład Urządzeń i Gospodarki Elektroenergetycznej, Częstochowa; e-mail: daka@el.pcz.czyst.pl

** Prof. nadzw. dr hab. inż. – Instytut Elektroenergetyki, Politechnika Częstochowska, Zakład Urządzeń i Gospodarki Elektroenergetycznej, Częstochowa; e-mail: poptom@el.pcz.czyst.pl

*** Mgr – Raiffeisen Bank

o mniejszym ryzyku popełnienia dużych błędów umożliwia równoległa analiza prognozy zużycia energii elektrycznej, prognozy szczytów obciążeń z jednoczesnym wyznaczeniem wielkości umożliwiających realność prognoz. Może to przykładowo być wyznaczenie rocznych stopni obciążenia, dla których mamy ściśle określony zakres zmienności.

W przypadku dysponowania jedynie prognozą zużycia energii elektrycznej możliwe są inne analizy, przykładowo porównania energochłonności PKB, energochłonności produktów czy gałęzi gospodarki w wybranych krajach. Jeszcze innym, powszechnie stosowanym wskaźnikiem jest roczne zużycie energii elektrycznej na osobę. Prognozy ludności są jednymi z dokładniejszych, stąd ten wskaźnik może być dobrym odniesieniem przy porównywaniu prognoz.

SŁOWA KLUCZOWE: prognozowanie, model ekonometryczny, system elektroenergetyczny, metoda doboru zmiennych, prognoza długoterminowa, moc szczytowa, zużycie energii

1. Charakterystyka problemu

Znajomość długoterminowych prognoz, między innymi zapotrzebowania na energię elektryczną, czy prognoz obciążeń szczytowych jest istotnym czynnikiem w planowaniu przyszłych źródeł wytwarzania i związanym z tym wzrostem wartości przesyłanych mocy (Lipko i in. 2010; Malko 1995). W horyzoncie długoterminowym prognozy wykonywane są przy założeniach probabilistycznych, zarówno przy ocenie zużycia energii, jak i określenia poziomu wymaganych mocy maksymalnych (szczytowych) w systemie (Maciejewski 2005; Maciejewski 2008).

Prognozy długoterminowe najczęściej wykonywane są wielowymiarowymi modelami ekonometrycznymi (Hirschhausen, Andres 2000). Najważniejsze decyzje, które należy podjąć, to wybór rodzaju modelu i wybór dostępnych zmiennych, objaśniających badany (prognozowy) proces. Wiedza o właściwościach modelu, użytego do wyznaczenia prognozy, jest oczywiście bardzo istotna, zwiększa wiarygodność i zaufanie odbiorcy do wyników, zmniejszając ryzyko (Jurdiak, Wiktorowicz 2009). Jest to szczególnie istotne w przypadku prognoz długoterminowych. W artykule skoncentrowano się na drugim czynniku, bardzo ważnym aspekcie modelowania, a mianowicie – ustaleniu zbioru zmiennych objaśniających wybieranych ze zbioru potencjalnego i wpływu tego wyboru na wartości prognoz. Wybór zmiennych oraz ich wpływ na wartości wygenerowanych prognoz są tematem dalszej części artykułu.

2. Metody doboru składowych

Procedura doboru powinna być skuteczna, czyli wyznaczać zmienne do modelu najlepiej wyjaśniające wariację procesu. Rozwiązaniem optymalnym jest jak najmniejsza liczba

nieskorelowanych wzajemnie składowych objaśniających, a w możliwie maksymalnym stopniu skorelowanych ze zmienną objaśnianą, które w dostatecznym stopniu wyjaśniają wariancję zmiennej objaśnianej. W tak przedstawianym problemie, teoretycznie każda liczba składowych nie większa od liczby obserwacji, może stanowić wejście do modelu. Zatem możliwych jest $2^m - 1$ kombinacji. Sformułowanie powyższe znalazło w literaturze (Dobrzańska i in. 2002; Grabiński i in. 1999; Gruszczyński i in. 1989) omawiającej tę problematykę rozwinięcie. Spośród wielu metod doboru zmiennych do wielowymiarowych modeli ekonometrycznych, w artykule wykorzystano trzy, mianowicie:

1. Metodę pojemności integralnej Hellwiga.
2. Metodę współczynnika korelacji cząstkowej.
3. Metodę współczynnika korelacji wielowymiarowej.

Sformalizowany zapis zbioru danych

Metody doboru łatwiej opisać, gdy dane przedstawimy w zapisie macierzowym. Mając zbiór zmiennych objaśniających w postaci macierzy X oraz wektor Y będący wartościami badanego procesu, w pierwszym kroku wyznacza się dwie macierze współczynników korelacji. Macierzy D oraz macierzy R . Elementami macierzy D są wartości współczynników korelacji zmiennych X wraz z wektorem Y . Elementami macierzy R są wartości współczynników korelacji zmiennych X .

$$D = \begin{bmatrix} 1 & r_{01} & r_{02} & \cdots & r_{0k} \\ r_{10} & 1 & r_{12} & \cdots & r_{1k} \\ r_{20} & r_{21} & 1 & \cdots & r_{2k} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ r_{k0} & r_{k1} & r_{k2} & \cdots & 1 \end{bmatrix} \quad R = \begin{bmatrix} 1 & r_{12} & \cdots & r_{1k} \\ r_{21} & 1 & \cdots & r_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ r_{k1} & r_{k2} & \cdots & r_{kk} \end{bmatrix} \quad (1)$$

Zadaniem algorytmu wyznaczającego optymalny zbiór zmiennych egzogenicznych jest wybranie z macierzy X jednej lub wielu zmiennych, które jak najlepiej (wg przyjętego kryterium) odwzorują zmienność procesu Y .

Metoda pojemności integralnych informacji Hellwiga

Punktem wyjścia jest wyznaczenie wektora R_0 współczynników korelacji między zmienną Y a kolejnymi składowymi objaśniającymi X_i oraz macierzy R współczynników korelacji między zmiennymi objaśniającymi. Następnie wyznacza się indywidualne pojemności nośników informacji X_i , wchodzących w skład różnych kombinacji ze zbioru zmiennych objaśniających. Dla m zmiennych objaśniających, liczba k -elementowych zbiorów, $1 \leq k \leq m$ wynosi $2^m - 1$. Zmienna X_j jest tym lepszym nośnikiem informacji o zmiennej Y , im bliższy jedności jest moduł współczynnika korelacji liniowej r_j . Zanieczyszczeniem indywidualnego nośnika informacji X_j nazywa się wielkość

$$g_j = \frac{1}{m-1} \sum_{i \neq j} |r_{ij}| \quad \text{zawsze zachodzi nierówność } 0 \leq g_j \leq 1 \quad (2)$$

Pojemność indywidualna nośnika informacji X_j o zmiennej Y wynosi

$$h_j = \frac{r_j^2}{1 + \sum_{i \neq j} |r_{ij}|} = \frac{r_j^2}{1 + (m-1)g_j} \quad (3)$$

Pojemności integralne nośników informacji wyrażają się wzorem

$$H_k = \sum_{j=1}^m h_j \quad k = 1, 2, \dots, 2^m - 1 \quad (4)$$

Wielkość H_k jest unormowana, w przedziale $\langle 0, 1 \rangle$ i interpretujemy ją jako miarę względnego poziomu pojemności informacji danej kombinacji zmiennych objaśniających. Pojemność H_k bliska jedności oznacza, że zmienne wchodzące w skład danej kombinacji dostarczają niemal pełnego zasobu informacji o zmiennej endogenicznej Y . Kryterium wyboru optymalnej kombinacji zmiennych objaśniających sprowadza się do wybrania maksymalnej wartości H_m .

Metoda analizy współczynników określoności cząstkowej

Metody oparte na analizie jedynie zwykłych współczynników korelacji zmiennej zależnej z wyróżnioną pojedynczą zmienną niezależną poddawane są krytyce gdyż zróżnicowanie zmiennej zależnej Y praktycznie zawsze jest pochodną wielu czynników, wśród których tylko jednym z wielu jest wyróżniona zmienna. Z tego powodu ustalenie wielkości wpływu zmiennej niezależnej X na zmienną zależną Y powinno wskazywać, jaka część ogólnego, wynikającego ze zmian wszystkich zmiennych niezależnych – przyrostu zmiennej Y , wynika ze zmiany wyróżnionej zmiennej niezależnej X .

Korelacja cząstkowa bada związek dwóch cech, zakładając, że istnieją oraz mają wpływ na badane cechy, inne zmienne, które możemy świadomie wyeliminować z obliczeń. Szacowanie wpływu zmiennej X na zmienną Y oznacza więc wyznaczenie cząstkowego wpływu zmiennej X na zmienną Y przy założeniu, że pozostałe zmienne niezależne nie zmieniają się. Z matematycznego punktu widzenia narzędziem określającym cząstkowy wpływ zmiennej X_k na zmienną Y jest pierwsza pochodna cząstkowa funkcji względem zmiennej X_k . Praktycznie współczynnik korelacji cząstkowej między składową X_i i zmienną objaśnianą Y , można traktować, jak zwykły współczynnik korelacji, lecz na zmiennych przekształconych następująco:

$$X^*_k = X_k - \left(\sum_{q=1}^m \alpha_{qk} X_q + \alpha_{k0} \right) \quad (5)$$

$$Y^*_k = \sum_k \gamma_{ik} X_k \quad (6)$$

gdzie:

$$k = i, j \quad q \neq i, j \quad k = 1, 2, \dots, m \quad k \neq i$$

parametry α_{qk} , α_{k0} i γ_{ik} należy dobrać tak, aby zachodziło:

$$E(X^*_j)^2 = \min \quad \text{oraz} \quad E(X^*_i)^2 = \min$$

W praktyce najczęściej oblicza się kwadrat współczynnika korelacji cząstkowej, korzystając z zapisu macierzowego w następujący sposób:

$$R^2_{yj, 1, \dots, j-1, j+1, \dots, k} = 1 - \frac{\det R}{\det R_{jj}} \quad (7)$$

gdzie: $\det R$ – wartość wyznacznika macierzy R ,

R_{jj} – wartość dopełnienia algebraicznego elementu r_{jj} .

Po uszeregowaniu malejącym uwzględnia się w modelu te zmienne, które mają największą wartość R^2 .

Korelacja wieloraka

Współczynnik korelacji wielorakiej jest syntetyczną miarą współzależności między zmienną Y a wybranymi zmiennymi X , traktowanymi łącznie. Określa on zależność jednej zmiennej zależnej od zespołu k zmiennych niezależnych. Liczbowa wartość współczynnika determinacji R^2 określa, jaka część zmienności zmiennej zależnej jest spowodowana przez badane zmienne niezależne. Interpretacja wielkości współczynników korelacji wielorakiej jest podobna jak modułów wartości współczynników korelacji cząstkowej i całkowitej.

Korzystając z macierzy D oraz R (wzór 1) można – podobnie jak w przypadku współczynników korelacji cząstkowej – wyznaczyć wartości korelacji wielorakiej.

Wykorzystując wyznaczniki obu macierzy współczynnik korelacji wielorakiej oblicza się następująco:

$$R_{0,1,2,\dots,k} = \sqrt{1 - \frac{|D|}{|R|}} \quad (8)$$

gdzie: $\det R$ – wyznacznik macierzy współczynników korelacji,

$\det D$ – wyznacznik macierzy D powstałej przez skreślenie odpowiednich k wierszy i k kolumny macierzy współczynników korelacji R .

Współczynnik korelacji może przyjmować wartości w przedziale: 0–1.

3. Model zastosowany do obliczeń

W części praktycznej użyto modelu MRK opracowanego w Instytucie Elektroenergetyki Politechniki Częstochowskiej, wielokrotnie stosowanego w praktyce. Wybór modelu ma zresztą dla przedstawianego problemu znacznie mniejsze znaczenie. Zadanie bowiem polega na sprawdzeniu wpływu doboru zmiennych przy zastosowaniu tego samego modelu, na wartości wyznaczonych prognoz. Model MRK korzysta z rozkładu kanonicznego wektora losowego oraz empirycznych rozkładów warunkowych. Zapis modelu jest następujący. Przyjmując, że badany proces opisany jest wektorem losowym X , którego składowe X_i ($i = 1, 2, \dots, m$) są ze sobą skorelowane, dokonuje się ortogonalizacji wektora X na inny wektor V . Po przekształceniach model dla jednej chwili (indeks czasu pominięto) można zapisać następująco:

$$\begin{aligned}X_{01} &= V_1 \\X_{02} &= a_{21}V_1 + V_2 \\X_{03} &= a_{31}V_1 + a_{32}V_2 + V_3 \\&\vdots \\X_{0m} &= a_{m1}V_1 + a_{m2}V_2 + a_{m,m-1}V_{m-1} + V_m\end{aligned}\tag{9}$$

$$\hat{X}_i = \sum_{j=1}^{i-1} a_{ij}V_j + \hat{V}_i + \bar{x}_{Xi}\tag{10}$$

gdzie: m – liczba składowych wektorów X , $X_{0m} = X_m - \bar{x}_m$,
 a_{ij} – współczynniki rozkładu kanonicznego tak dobrane, aby zortogonalizować składowe V_i ,
 V_i – składowe wektora V po rozkładzie kanonicznym wektora X ,
 \bar{x}_{Xi} – wartość średnia składowej X_i .

Nieznaną w równaniu (10) wartość \hat{V}_i , wyznacza się z empirycznych rozkładów wyznaczonych z danych historycznych. Prognozowana jest zmienna $X_{0m} = Y_0$, rolę zmiennych objaśniających pełnią zmienne $X_{01} \dots X_{0m-1}$. Pełny opis procedury rozkładu kanonicznego wraz z jej zastosowaniem w procesie predykcji zamieszczono w (Dobrzańska i in. 2002). Model wykorzystano między innymi do prognozy szczytów zimowych w KSE, w perspektywie do 2030 roku dla potrzeb PSE Operator (Popławski i in. 2010).

4. Przykład obliczeniowy – prezentacja uzyskanych wyników

Procesem prognozowanym są wartości Y rocznego zużycia energii elektrycznej brutto w KSE. Potencjalnymi zmiennymi objaśniającymi, tworzącymi macierz X są:

- ✧ X_1 – szczyty roczne w KSE [MW],
- ✧ X_2 – całkowite zużycie energii pierwotnej w 27 krajach UE [PJ],
- ✧ X_3 – zużycie ogółem energii elektrycznej netto w 27 krajach UE [GWh],
- ✧ X_4 – energochłonność – całkowite zużycie energii pierwotnej na jednego dolara PKB w 27 krajach UE (ceny Exchange Market) [Btu/USD],
- ✧ X_5 – ludność Polski [tys.],
- ✧ X_6 – PKB Polski w cenach bieżących [mln zł],
- ✧ X_7 – energochłonność – całkowite zużycie energii pierwotnej na jednego dolara PKB w Polsce (ceny Exchange Market) [Btu/USD],
- ✧ Y – Zużycie energii elektrycznej brutto w kraju [GWh].

Dane historyczne obejmują okres 19 lat, od 1990 do 2008 roku. Prognozy wykonano do 2030 roku. Na podstawie źródłowych danych historycznych wyznaczono macierze współczynników korelacji między zmiennymi. Tabela 1 przedstawia pełną macierz współczynników korelacji – D zgodnie ze wzorem (1). Pola cieniowane w tabeli 1 są wartościami dla macierzy R także zgodnie ze wzorem (1).

Korelacje z tabeli 1 są bazą dla wykonania obliczeń trzema metodami doboru zmiennych. Metoda Hellwiga przy 7 zmiennych objaśniających daje 127 możliwych kombinacji zmiennych. Zakres zmienności współczynników pojemności integralnych Hellwiga zawarty jest w przedziale 0,92–0,49, z czego prawie 96% jest większych od 0,7, wskazując na istotną statystycznie wartość.

TABELA 1. Współczynniki korelacji – macierz D

TABLE 1. Correlation coefficients – D matrix

	Y	X_1	X_2	X_3	X_4	X_5	X_6	X_7
Y	1	0,840	0,839	0,874	-0,833	-0,696	0,886	-0,800
X_1	0,840	1	0,575	0,596	-0,537	-0,655	0,617	-0,510
X_2	0,839	0,575	1	0,978	-0,926	-0,677	0,962	-0,957
X_3	0,874	0,596	0,978	1	-0,975	-0,655	0,987	-0,969
X_4	-0,833	-0,537	-0,926	-0,975	1	0,545	-0,984	0,967
X_5	-0,696	-0,655	-0,677	-0,655	0,545	1	-0,636	0,635
X_6	0,886	0,617	0,962	0,987	-0,984	-0,636	1	-0,977
X_7	-0,800	-0,510	-0,957	-0,969	0,967	0,635	-0,977	1

Źródło: opracowanie własne

W tabeli 2 zamieszczono sześć wybranych kombinacji zmiennych, dla których współczynniki pojemności integralnej Hellwiga są największe oraz sześć najlepszych kombinacji z największymi wartościami współczynnika korelacji wielorakiej.

W tabeli 3 oraz na rysunku 1 zamieszczono współczynniki korelacji cząstkowej. Dla porównania w tabeli 3 oprócz współczynników korelacji cząstkowej zamieszczono także wartości współczynników korelacji zwykłej. Warty podkreślenia jest fakt znacznych zmian w wartościach współczynników korelacji cząstkowej, a nawet możliwa zmiana znaku.

Ze zbioru najlepszych kombinacji zmiennych otrzymanych z 3 metod, zbudowano 14 modeli, 6 według wskazań metody Hellwiga, 6 według rankingu wartości współczynnika korelacji wielorakiej oraz dwa według wartości współczynników korelacji cząstkowej.

TABELA 2. Wybrane kombinacje z najwyższymi wartościami współczynników pojemności integralnej Hellwiga oraz współczynników korelacji wielorakiej

TABLE 2. Chosen combinations with the highest values of the coefficients of integral capacity by Hellwig and coefficients of multiple correlation

Nr	Metoda Hellwiga		Metoda współczynnika korelacji wielorakiej	
	kombinacja zmiennych	wartość współczynnika pojemności integralnej Hellwiga	kombinacja zmiennych	wartość współczynnika korelacji wielorakiej
1.	1-6	0,9209	1-2-3-4-5-6	0,9998
2.	1-3	0,9199	1-2-3-4-5-6-7	0,9759
3.	1-3-6	0,9151	2-3-4-5-6-7	0,9680
4.	1-4-5-6	0,8904	1-2-6	0,9609
5.	1-2-4-5-6	0,8767	1-2-5	0,9458
6.	1-2-7	0,8755	1-7	0,9444

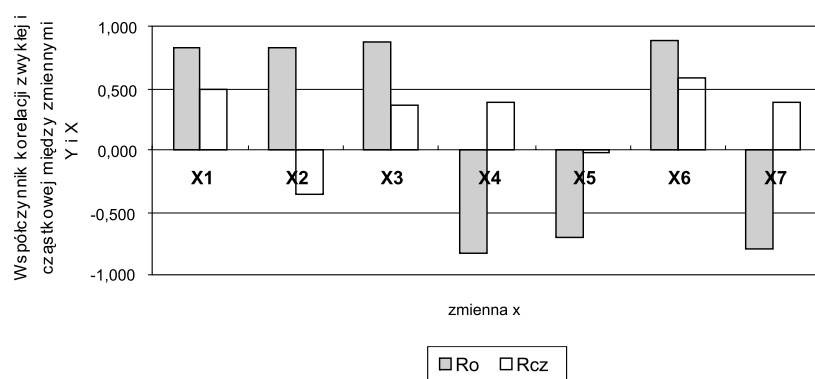
Źródło: opracowanie własne

TABELA 3. Współczynniki korelacji zwykłej oraz korelacji cząstkowej zmiennej Y ze zmiennymi objaśniającymi przy wyłączonych pozostałych

TABLE 3. Coefficients of ordinary and partial correlation of the variable Y with the explanatory variables, keeping other variables "switched off"

	X_1	X_2	X_3	X_4	X_5	X_6	X_7
Korelacja zwykła R_o	0,840	0,839	0,874	-0,833	-0,696	0,886	-0,800
Korelacja cząstkowa z wyłączeniem pozostałych sześciu zmiennych R_{cz}	0,495	-0,357	0,367	0,386	-0,007	0,593	0,389

Źródło: opracowanie własne



Rys. 1. Wykres współczynników korelacji zwykłej oraz korelacji cząstkowej zmiennej Y ze zmiennymi objaśniającymi przy wyłączonych pozostałych

Źródło: opracowanie własne

Fig. 1. The chart of coefficients of ordinary and partial correlation of the variable Y with the explanatory variables, keeping other variables “switched off”

TABELA 4. Wartości błędów dopasowania modelu dla badanych kombinacji zmiennych

TABLE 4. Values of fitting errors for chosen combinations of variables

Lp.	Metoda wyboru zmiennych	Wariant ze zmiennymi	Błąd prognozy wygasłej
1.	Metoda Hellwiga	1-6	1,05%
2.	Metoda Hellwiga	1-3	0,97%
3.	Metoda Hellwiga	1-3-6	1,08%
4.	Metoda Hellwiga	1-4-5-6	1,01%
5.	Metoda Hellwiga	1-2-4-5-6	0,98%
6.	Metoda Hellwiga	1-2-7	1,10%
7.	Korelacja cząstkowa	1-2-4-6-7	0,80%
8.	Korelacja cząstkowa	1-2-3-6-7	1,06%
9.	Korelacja wieloraka	1-2-3-4-5-6	0,85%
10.	Korelacja wieloraka	1-2-3-4-5-6-7	0,67%
11.	Korelacja wieloraka	2-3-4-5-6-7	0,86%
12.	Korelacja wieloraka	1-2-6	1,06%
13.	Korelacja wieloraka	1-2-5	1,14%
14.	Korelacja wieloraka	1-7	1,23%

Źródło: opracowanie własne

Kombinacje zmiennych według współczynnika korelacji cząstkowej (tab. 3) wybrano tak, aby nie były powtórzeniem kombinacji wcześniej wybranych dwiema pozostałymi metodami. Ustalono dwie następujące kombinacje:

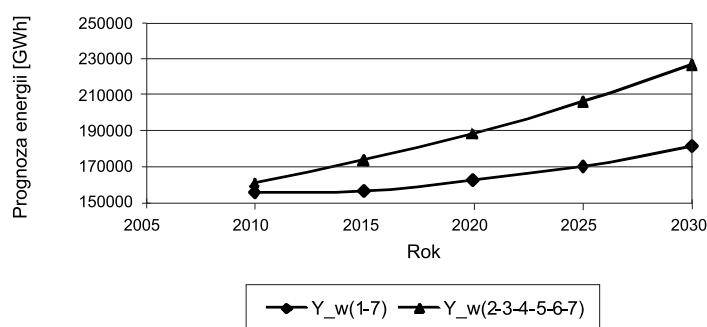
$$Y = f(X_1, X_2, X_4, X_6, X_7)$$

$$Y = f(X_1, X_2, X_3, X_6, X_7)$$

Wyniki dopasowania modeli dla danych historycznych wyrażonych średnim błędem względnym w % przedstawiono w tabeli 4. Najmniejsze błędy w każdej z metod wyróżniono w tabeli pogrubieniem. Wszystkie błędy prognoz wygasłych należy uznać za małe i każdy z 14 modeli może być użyty do wykonania prognoz właściwych.

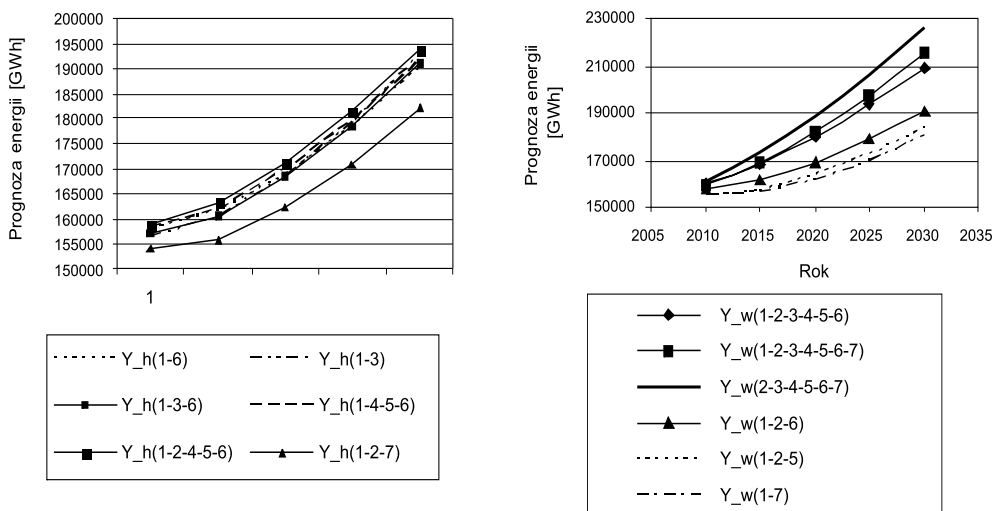
5. Prognozy w horyzoncie do 2030 roku

Wobec niewielkich błędów prognoz wygasłych prognozy właściwe w horyzoncie do 2030 roku wykonano dla wszystkich 14 kombinacji. Scenariusze dla danych wejściowych – z wyjątkiem prognozy ludności (opracowanie GUS) – są opracowaniem autorskim. Prognozy pokazano na rysunkach 2 i 3. Rozpiętość prognoz jest wyraźna w każdej z metod doboru. Mniejsza jest dla metody Hellwiga i metody współczynnika korelacji cząstkowej, około 10 GWh, większa dla metody współczynnika korelacji wielorakiej, dochodząca do 40 GWh. Pasma różnicy między prognozami skrajnymi dla metod Hellwiga i korelacji cząstkowej jest prawie stałe, zaś dla metody korelacji wielorakiej wykazuje lekką tendencję wzrostową. Na rysunku 2 pokazano dwie skrajne prognozy ze wszystkich 14. Obydwie powstały według metody współczynnika korelacji wielorakiej. Różnica w 2030 roku między nimi jest znacząca i wynosi około 46 000 GWh.



Rys. 2. Zakres prognoz minimalnych i maksymalnych
Źródło: opracowanie własne

Fig. 2. Range of minimum and maximum forecasts



Rys. 3. Prognoza według metody doboru Hellwiga – strona lewa oraz według metody wielorakiej – strona prawa
 Źródło: opracowanie własne.

Fig. 3. A ranked chart of model fitting errors for the considered combinations of variables

Podsumowanie i wnioski

Interesujące jest porównanie prognoz walidacyjnych dla 2009, 2010 i 2011 roku. Okazuje się, że zbyt dobre dopasowanie modelu dla danych historycznych, skutkowało w niektórych wariantach większymi błędami prognoz walidacyjnych. Zjawisko to dodatkowo utrudnia wybór wariantu. Naturalne zatem staje się pytanie, który wariant zmiennych rekomendować. Trudno bez dodatkowych informacji i wiedzy podjąć właściwą decyzję. Pomocą w ocenie realności prognoz z całą pewnością jest pogłębiona analiza wykorzystująca wiedzę ekspercką o badanym procesie. Na pewno pełniejszą ocenę prognozy zużycia energii można wykonać, znając w tym samym czasie dla rozpatrywanego horyzontu także prognozę obciążeń szczytowych. Znając obie te prognozy można wyznaczyć wskaźniki dobrze zdefiniowane, wiążące te dwie wielkości w każdym systemie (Dobrzańska i in. 2002). Przykładami jest wskaźnik określający strukturę odbiorców, nazywany średnim rocznym statycznym – względnie dynamicznym – stopniem obciążenia m_r , którego wartość zawiera się w przedziale (0–1). W przypadku dysponowania jedynie prognozą zużycia energii elektrycznej możliwe są inne analizy, przykładowo: porównania energochłonności PKB, wybranych produktów czy gałęzi gospodarki w wybranych krajach. Jeszcze innym, powszechnie stosowanym wskaźnikiem jest roczne zużycie energii elektrycznej na osobę. Prognozy ludności są jednymi z lepiej dopracowanych, stąd ten wskaźnik może być dobrym odniesieniem przy porównywaniu prognoz. Wykonane badania dotyczące tylko jednego z wielu aspektów modelowania oraz przedstawione wyniki

uświadamiają jak trudny i odpowiedzialny jest proces prognozowania, szczególnie w długich horyzontach.

Literatura

- DOBRZAŃSKA i in. 2002 – DOBRZAŃSKA I., DĄSAL K., ŁYP J., POPLAWSKI T., SOWIŃSKI J., 2002 – Prognozowanie w elektroenergetyce. Zagadnienia wybrane. Wyd. PCz., Częstochowa.
- GRABIŃSKI i in. 1999 – GRABIŃSKI T., WYDYMUS S., ZELIAŚ A., 1999 – Metody doboru zmiennych w modelach ekonometrycznych. PWN Warszawa.
- GRUSZCZYŃSKI M., i in., 1989 – Miary zgodności, metody doboru zmiennych, problemy współliniowości. PWN.
- von HIRSCHHAUSEN Ch., ANDRES M., 2000 – Long-Term Electricity Demand in China-From Quantitative to Qualitative Growth. Energy Policy, Vol. 28, 231–241.
- JURDZIAK L., WIKTOROWICZ J., 2009 – Prognozowanie poziomu ryzyka finansowego dla układu kopalni węgla brunatnego i elektrowni. Polityka Energetyczna t.12, z. 2/2. Wyd. IGSMiE PAN, Kraków, s. 205–216.
- LIPKO i in. 2010 – LIPKO K., PARCZEWSKI Z., TATAREWICZ I., KLIMPEL A., 2010 – EPC S.A., Długoterminowe prognozy popytu na energię i moc elektryczną w kraju dla potrzeb rozwojowych PSE Operator SA, ELEKTROENERGETYKA nr 1(3).
- MACIEJEWSKI Z., 2005 – Prognozy a możliwości krajowego systemu elektroenergetycznego. Polityka Energetyczna t. 8, z. spec. Wyd. IGSMiE PAN, Kraków, s. 205–215.
- Maciejewski Z., 2008 – Sieci przesyłowe jako element bezpieczeństwa elektroenergetycznego Polski. Polityka Energetyczna t. 11, z. 1. Wyd. IGSMiE PAN, Kraków, s. 285–298.
- MALKO J., 1995 – Wybrane zagadnienia prognozowania w elektroenergetyce. OWPW, Wrocław.
- POPLAWSKI i in. 2010 – POPLAWSKI T., DĄSAL K., DUDEK G., ŁYP J., SOWIŃSKI J., STARCZYŃSKA E., 2010 – Opracowanie i weryfikacja modeli do prognozowania zapotrzebowania na moc elektryczną w krajowym systemie elektroenergetycznym (KSE) dla potrzeb opracowania w PSE Operator S.A. planów koordynacyjnych. Centrum Zastosowań Zaawansowanych Technologii Sp. z o.o., Warszawa, (niepublikowany).

Kazimierz DAŚAL, Tomasz POPLAWSKI, Karolina RUSEK

Long-term forecasts of electric energy consumption in the domestic power engineering system – selected issues

Abstract

There exist a number of important factors in forecasting processes. The most significant in the case of a multi-dimensional econometric model are the choice of the model and the explanatory variables. This paper focuses on the choice of variables, which plays a crucial role. Among many existing and recognized algorithms for the selection of variables, the following three chosen arbitrarily by the authors are presented: the method of integral capacity by Hellwig, the partial correlation coefficient, and multiple correlation coefficient. For each considered method, a synthetic description explaining its action is given. The choice of the model for making forecasts was treated as less significant because, for each method, the same model is used. Calculations were made using the MRK (Model of Canonical Distribution) model. The synthetic description of the model is also provided. Using real-life data, the analysis demonstrates how the method applied in choosing explanatory variables influences the obtained forecast results concerning the consumption of electric energy on a national scale. In the example calculation, a potential dataset of seven explanatory variables was used. The paper summarizes fourteen different forecasts obtained from three methods of variables selection. The results of model fittings (extinct forecasts) and forecasts until 2030 are presented in the form of tables and charts. Concluding remarks concern long-term forecasts of fundamental quantities related to the Domestic Power Engineering System, which may influence the correctness of these forecasts. Particular attention was paid to the issue of the correctness of long-term forecasts related to energy consumption and peak power. An analysis of the forecast of energy consumption, processed parallel to the determination of quantities assessing the reality of these forecasts, makes it possible to indicate the forecasts endowed with the lowest risk of making excessive errors. For example, it may be possible to consider the determination of annual load levels for which a pre-determined, exact level of variability is given.

In situations where only the forecast of electric energy consumption is available, it is possible to develop further analyses such as a comparison of energy consumption for a Domestic Brutto Product, energy consumption for products, or economic branches in selected countries. Yet another indicator in common use is annual energy consumption *per capita*. Forecasts *per capita* are among the most exact; therefore, this indicator may be a useful tool for comparison of different forecasts.

KEY WORDS: Forecasting, econometric model, power engineering system, method of the choice of variables, long-term forecast, peak power, energy consumption