

METODYKA ANALIZY I PROGNOZY OBCIĄŻEŃ ELEKTROENERGETYCZNYCH SYSTEMÓW LOKALNYCH

Jacek Łyp

Instytut Elektroenergetyki, Politechnika Częstochowska, Częstochowa

1 WPROWADZENIE

Różnorodne aspekty zjawisk gospodarczych związanych z kształtowaniem zapotrzebowania na energię elektryczną wymagają różnych informacji o istocie tego procesu. Informacji takich dotyczących przyszłości, czyli prognoz elektroenergetycznych, wymaga się zwłaszcza w procesach planowania rozwoju i eksploatacji systemu elektroenergetycznego. Znaczenie prognozowania elektroenergetycznego znalazło swoje formalne odzwierciedlenie w uregulowaniach prawnych. W ustawie Prawo energetyczne [10] i w rozporządzeniach Ministra Gospodarki [11] czytamy o obowiązkach prognostycznych operatora systemu przesyłowego, operatorów systemów rozdzielczych i gmin. Decentralizacja organizacyjno-techniczna energetyki znajduje również wyraz w przyjętych założeniach polityki energetycznej państwa [8].

Wzrost znaczenia prognozowania regionalnego przy jednoczesnej uzasadnionej obawie o niedostatki aktualnego stanu archiwizacji danych i rzeczywistych możliwościach ich pozyskiwania stanowił genezę niniejszej pracy, w której podjęto problematykę analizy i prognozy zmienności obciążeń w lokalnych systemach elektroenergetycznych.

2 CEL PRACY

Celem pracy jest wskazanie i zweryfikowanie efektywnej metody średnio-/długoterminowej prognozy przebiegów dobowych obciążenia systemów lokalnych dla dni charakterystycznych, która powinna działać w oparciu o zbiór

realnie dostępnych w obecnych warunkach danych statystycznych i scenariusz zapotrzebowania na energię roczną, i moc szczytową, przygotowany odrębną metodą.

3 PROBLEMATYKA ZADANIA I KONCEPCJA ROZWIĄZANIA

Szkieletem proponowanego rozwiązania jest algorytm prognozy wykorzystujący wyniki analizy zmienności obciążeń zbioru systemów lokalnych na wzór modelu opracowanego przez K. Kopeckiego [6], a rozwiniętego przez I. Dobrzańską [4, 8]. Tabele 1, 2 prezentują w skrócie najważniejsze wielkości modelu analizy.

Tabela 1. Kluczowe wskaźniki zmienności obciążeń ([8]).

$m_r = \frac{A_r}{T_r \cdot P_{Sr}}$	Roczny stopień obciążenia A_r - praca roczna, P_{Sr} - moc szczytowa roku, T_r - czas trwania roku (8760 h)
$\sigma'_{rA} = \frac{A_{d\acute{s}r7}}{A_{d\acute{s}r12}}$	Stopień zmienności średniej miesięcznej pracy dobowej $A_{d\acute{s}r i}$ - średnia praca dobowa i-tego miesiąca
$\sigma'_{rA} = \frac{P_{S7}}{P_{Sr}}$	Stopień zmienności szczytów miesięcznych P_{Si} - moc szczytowa i-tego miesiąca)
$m_d = \frac{A_d}{T_d \cdot P_{Sd}}$	Dobowy stopień obciążenia A_d - praca dobowa P_{Sd} - moc szczytowa doby T_d - 24 h
$\frac{1}{m_{dh}} = \frac{T_d \cdot P_h}{A_d}$	Chwilowy stopień wyrównania doby w godzinie h P_h - moc w godzinie h

Procedury analizy dostarczają zestaw charakterystyk wiążących ze sobą różne wskaźniki (tabela 3). W funkcji kilku kluczowych wielkości możliwych do uzyskania bezpośrednio lub

pośrednio z wartości rocznego zapotrzebowania na energię i moc szczytową, można wyznaczyć z pomocą poszczególnych charakterystyk wartości wszystkich szczegółowych wielkości umożliwiających skonstruowanie przebiegu obciążenia dobowego dla danego systemu. Wymaganiem elementem algorytmu prognozy dobowej zmienności obciążeń, oprócz zestawu charakterystyk, są prognozy rocznej energii i mocy szczytowej uzyskane odrębnymi metodami.

Tabela 2. Obliczanie wielkości statycznych ([8]).

$\alpha_A = \frac{A_r - A_{r-1}}{A_{r-1}}$	Względny roczny przyrost energii
$\alpha_S = \frac{P_{Sr} - P_{S,r-1}}{P_{S,r-1}}$	Względny roczny przyrost mocy
$A_r^{stat} = \frac{A_r}{1 + 0,54 \cdot \alpha_A}$	Roczna energia statyczna
$P_{Sr}^{stat} = \frac{P_{Sr}}{1 + \frac{\alpha_A + \alpha_S}{2}}$	Roczny szczyt statyczny
$P_{Si}^{stat} = \frac{P_{Si}}{1 + \frac{i}{12} \cdot \frac{\alpha_A + \alpha_S}{2}}$	Statyczny szczyt i-tego miesiąca
$A_{d\ sr\ i}^{stat} = \frac{A_{d\ sr\ i}}{1 + \frac{i}{12} \cdot \alpha_A}$	Statyczna praca średniodobowa i-tego miesiąca

Z próby zastosowania historycznych formuł do analizy zmienności obciążeń współczesnych systemów lokalnych odpowiadających obszarom działalności przedsiębiorstw dystrybucyjnych elektroenergetyki wynika m.in. nieprzydatność charakterystyk do ich użytecznej aproksymacji, z powodu dużej dyspersji punktów empirycznych (stąd nazwa „funkcje punktowe” [3]) co ilustruje rys. 1.

Tabela 3. Kluczowe relacje modelu analizy ([8]).

$m_{dz}^{stat} = f(m_r^{stat})$	$\sigma_r^{stat} = f(m_r^{stat})$	$\sigma_{rA}^{stat} = f(m_r^{stat})$
$m_{dl}^{stat} = f(m_r^{stat})$	$\frac{1}{m_{dz\ h}} = f\left(\frac{1}{m_{dz}}\right)$	$\frac{1}{m_{dl\ h}} = f\left(\frac{1}{m_{dl}}\right)$

Fakt ten stał się przyczyną podjęcia wysiłku opracowania skutecznej metody identyfikacji wartości funkcji punktowych, która wraz ze wspomnianym algorytmem prognozy (rys. 2) stanowi osiągnięcie założonego celu pracy.

Formalna teza pracy brzmi następująco: „Posiadając zestaw charakterystyk wskaźników opisujących zmienność obciążeń systemów elektroenergetycznych w postaci funkcji punktowych o ogólnej postaci $y = f_p(x)$:

$$m_{dz}^{stat} = f(m_r^{stat}), \quad (1)$$

$$m_{dl}^{stat} = f(m_r^{stat}), \quad (2)$$

$$\sigma_r^{stat} = f(m_r^{stat}), \quad (3)$$

$$h=1 \div 24 \quad \frac{1}{m_{dz\ h}} = f\left(\frac{1}{m_{dz}}\right), \quad (4)$$

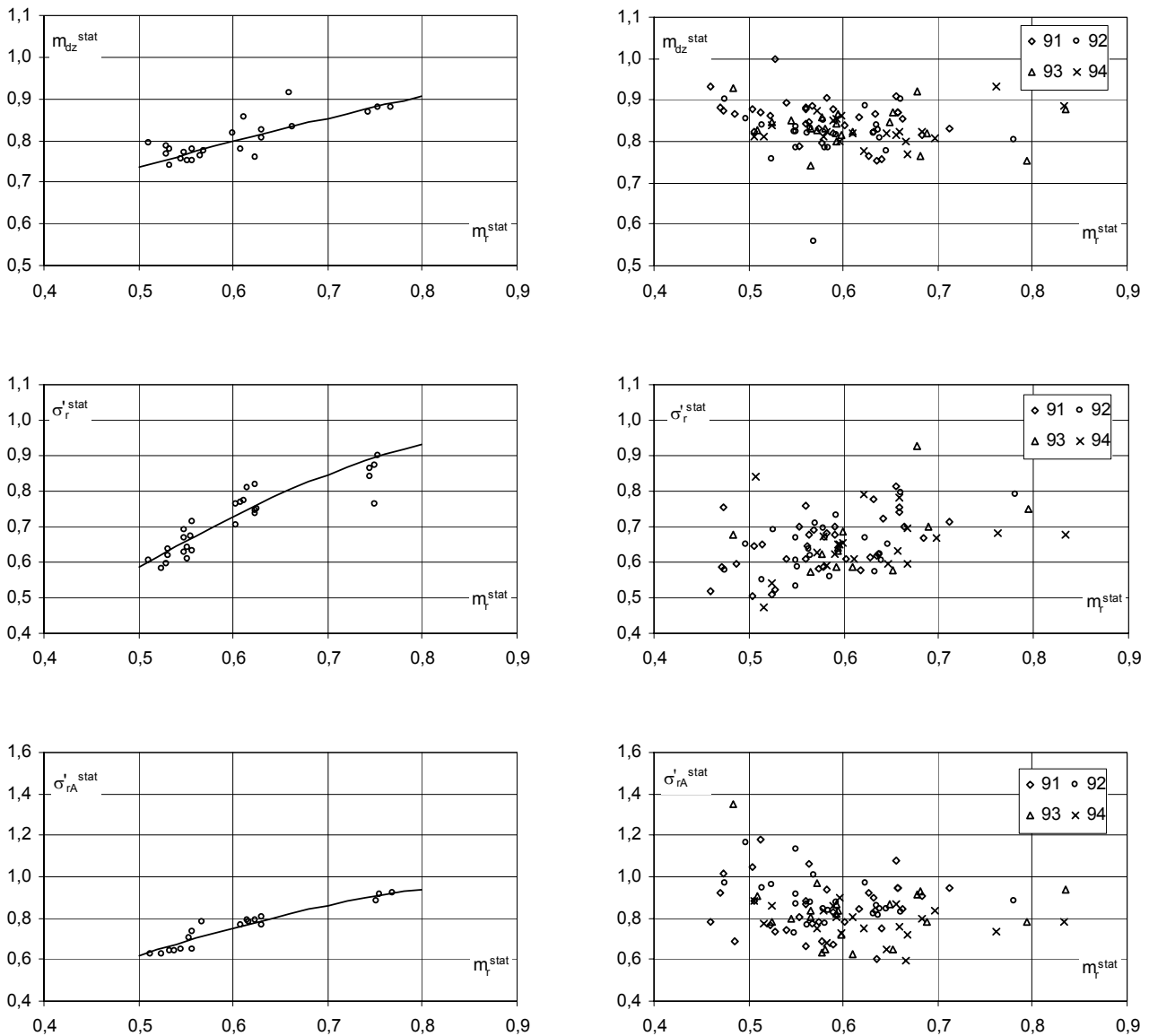
$$h=1 \div 24 \quad \frac{1}{m_{dl\ h}} = f\left(\frac{1}{m_{dl}}\right), \quad (5)$$

wykonany wg *n*-letniej statystyki systemów odpowiadających obszarom spółek dystrybucyjnych, można wyznaczyć z niego wartości \hat{y} dla żądanych \hat{x} z efektywnością pozwalającą na prognozę krzywych obciążenia dobowego systemów, wg zadanego zapotrzebowania na moc szczytową i energię roczną, z dokładnością nie mniejszą od wymaganej²”.

Do realizacji zadania wykorzystano: technikę predykcji wektora losowego z użyciem metody rozkładu kanonicznego, optymalizację algorytmem genetycznym i inne techniki obliczeniowe.

² Oczekiwania jakościowe stawiane przed prognozą zależą od celu prognozowania. Formułowane są w postaci ograniczeń bądź wymagań typu optymalizacyjnego polegającego na minimalizacji jednego lub więcej wskaźników jakości prognozy; jak np.: pułapu wartości średniego błędu absolutnego, błędu średniokwadratowego, wymogu nieobciążoności, pułapu błędu prawdopodobnego lub wysoce prawdopodobnego. W literaturze spotyka się również bardziej ogólnikowe określenia wymagań w postaci wartości błędów procentowych. Z uwagi na zróżnicowanie w przeznaczeniu prognoz nie sposób nie wziąć pod uwagę wynikającego stąd subiektywizmu w ocenie przydatności prognoz. W pracy [7] podano przykładowe oczekiwania dokładności prognoz średnioterminowych w różnych obszarach ich funkcjonowania:

- harmonogramowanie obsługi i remontów – względnie dokładne (10%);
 - koszty produkcji, harmonogramowanie elektrowni wodnych, marketing hurtowy – w zakresie kilku procent;
 - planowanie sieciowe, planowanie emisji – lepiej niż 10%.
- Zapewnienie błędu wysoce prawdopodobnego prognozy na poziomie kilku procent wydaje się zaspokajać wymagania dokładności w wymienionych obszarach.



Rys. 1. Porównanie wybranych relacji funkcyjnych dla systemów historycznych (lewa kolumna) i współczesnych (prawa kolumna).

4 POSZUKIWANIE EFEKTYWNEJ PROCEDURY IDENTYFIKACJI WARTOŚCI FUNKCJI PUNKTOWEJ

Funkcję punktową można rozumieć jako zbiór punktów na płaszczyźnie, gdzie współrzędne (x_t, y_t) każdego punktu są parą wartości konkretnych wskaźników opisujących stan w chwili t jednego spośród podlegających analizie systemów lokalnych.

4.1 Metoda koncentracji wymiarem Hausdorffa (mkwh)

Tok postępowania wg [3] polega na wyznaczeniu wartości funkcji punktowej $\hat{y}_{t+1} = f_p(x_{t+1})$

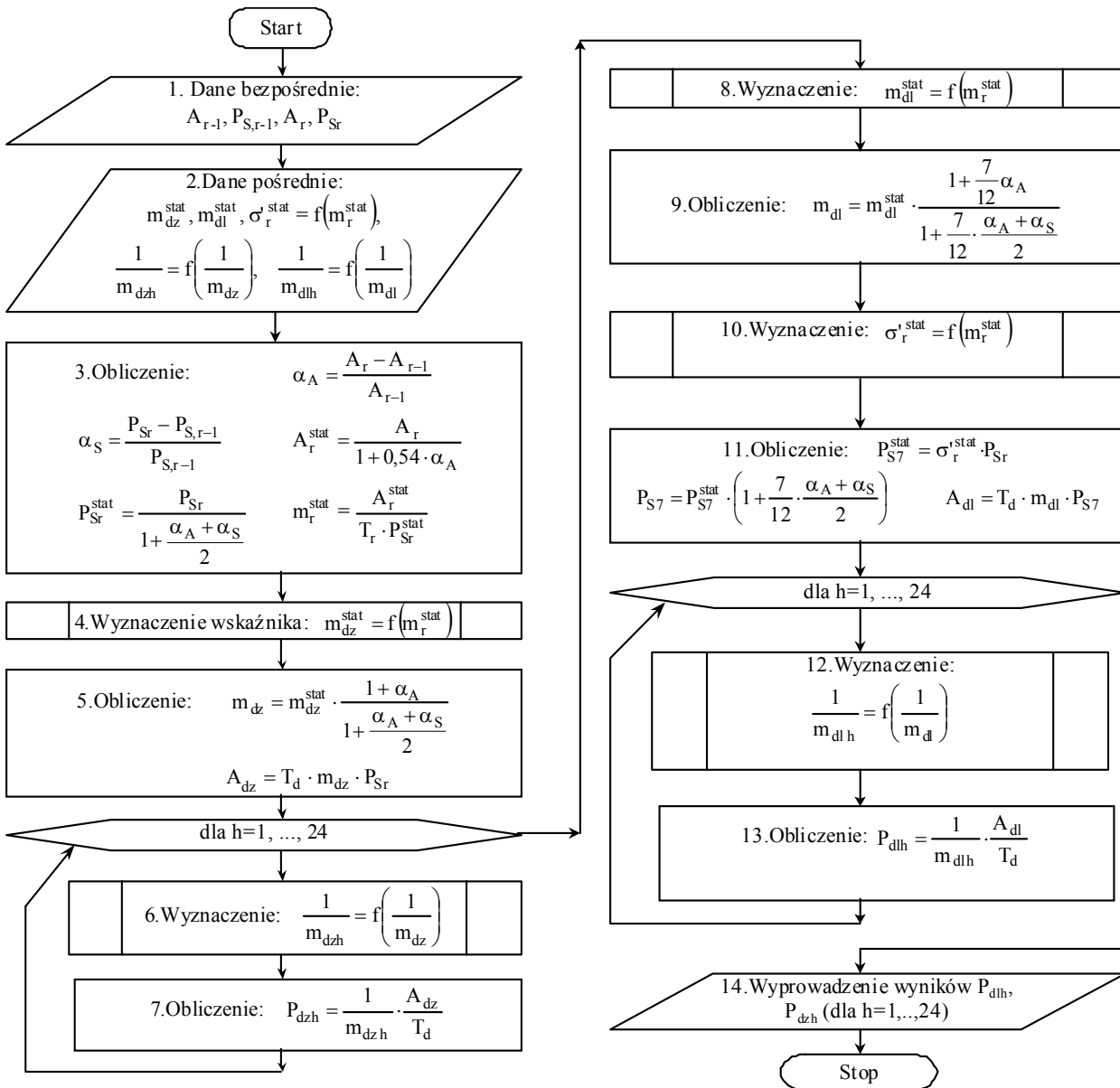
dla zadanego x_{t+1} , jako średniej ze spreparowanych wartości rzędnych punktów z pewnego przedziału Δx określonego w otoczeniu x_{t+1} . Spreparowanie rzędnych polega na skoncentrowaniu ich wokół wartości y_t . Poniżej przytoczono kolejne czynności procedury.

1 Określenie zbioru P punktów do wyznaczenia średniej:

$$x \in \left(\frac{(x_t + x_{t+1}) \mp c_1 |x_t - x_{t+1}|}{2} \right) \quad (6)$$

gdzie:

c_1 - stała zapewniająca obecność w przedziale ok. 15÷40% wszystkich punktów.



Rys. 2. Podstawowy algorytm prognozy godzinowych przebiegów obciążenia dla doby letniej i zimowej.

2 Eliminacja ze zbioru P niepożądanych (zakłócających) punktów wg kryterium:

$$|y_i - \bar{y}| > c_2 \cdot \bar{\Delta}_y \quad (7)$$

gdzie:

c_2 - stała $c_2 \in (1,25 - 2,5)$,

$\bar{y} = \frac{1}{n-1} \sum_{i=1}^{n-1} y_i$ - średnia z rzędnych $n-1$ punktów zbioru P, (z pominięciem y_t),

$\bar{\Delta}_y = \frac{1}{n-1} \sum_{i=1}^{n-1} |y_i - \bar{y}|$ - przeciętne absolutne odchylenie od średniej.

3 Wyznaczenie kresu górnego $y_{g \max}$ i dolnego $y_{g \min}$ dla $y_i > y_t$ $i=1, \dots, n$.

4 Przedstawienie kresów w postaci wskaźników:

$$W_{g \max} = \frac{y_{g \max}}{y_t}, \quad (8)$$

$$W_{g \min} = \frac{y_{g \min}}{y_t}. \quad (9)$$

5 Obliczenie wymiaru Hausdorffa [2]:

$$D_g = \frac{\ln(w_{g \min})}{\ln(w_{g \max})}, \quad (10)$$

6 Wyznaczenie skoncentrowanych rzędnych dla $y_i > y_t$:

$$y_{k_{gi}} = y_t + D_g(y_i - y_t) \quad (11)$$

7 Koncentracja $y_i < y_t$ (analogicznie).

8 Wyznaczenie górnego i dolnego \hat{y}_{t+1} :

$$\hat{y}_{t+1,g} = \frac{1}{n_g} \left(y_t + \sum_{i=1}^{n_g-1} y_{k_{gi}} \right), \quad (12)$$

$$\hat{y}_{t+1,d} = \frac{1}{n_d} \left(y_t + \sum_{i=1}^{n_d-1} y_{k_{di}} \right), \quad (13)$$

gdzie:

n_g, n_d - licznosci zbiorów.

9 Obliczenie wartości średniej \hat{y}_{t+1} metodą momentów:

$$\hat{y}_{t+1} = \frac{\hat{y}_{t+1,g}^2 + \hat{y}_{t+1,d}^2}{\hat{y}_{t+1,g} + \hat{y}_{t+1,d}}. \quad (14)$$

Średnia \hat{y}_{t+1} zastępuje w następnym kroku prognozy wieloletniej y_t i staje się podstawą do obliczenia wielkości $\hat{y}_{t+2,d}$, $\hat{y}_{t+2,g}$ oraz średniej \hat{y}_{t+2} .

4.2 Metoda wag podobieństwa (mwp)

Wagę odzwierciedlającą stopień podobieństwa struktury odbiorców dwóch systemów może być wartość współczynnika korelacji liniowej, wyznaczona dla zmiennych, którymi są obciążenia w poszczególnych godzinach doby typowej (zimowej i/lub letniej) porównywanych systemów w roku t:

$$k_{si} = \frac{\text{Cov}(P_s \cdot P_i)}{\sigma_{P_s} \cdot \sigma_{P_i}}, \quad (15)$$

gdzie:

P_s, P_i – obciążenie odpowiednio s-tego i i-tego systemu;

$\sigma_{P_s}, \sigma_{P_i}$ – odchylenia standardowe zmiennych

P_s, P_i ,

Wynik wyznaczenia wartości funkcji punktowej dla danego systemu s, (którego stan począt-

kowy określony jest punktem $(x_{s,t}, y_{s,t})$ w odciętej $x_{s,t+1}$, jest średnią ważoną z rzędnych punktów z przedziału $(x_{s,t+1}-\Delta x; x_{s,t+1}+\Delta x)$. Przedział powinien pokrywać ok. 20% całej populacji funkcji punktowej.

W celu wzmocnienia różnicującego działania wag dokonuje się przeskalowania:

$$(dla k_{si} \geq 0) \quad w_{si} = 1 - \sqrt{1 - k_{si}^2}, \quad (16)$$

Średnia ważona $\hat{y}_{s,t+1}$ wg metody momentów:

$$\hat{y}_{s,t+1} = \frac{\sum y_i w_{si}}{\sum w_{si}} \quad (17)$$

Dodatkowym zabiegiem poprawiającym dokładność identyfikacji jest zwielokrotnienie wagi w_{ss} ($w_{ss}=1$). Powoduje to określenie $\hat{y}_{s,t+1}$ bliżej wartości $y_{s,t}$. Rozmiar zwielokrotnienia (wspólny dla wszystkich funkcji punktowych) optymalizuje się ze względu na minimum błędów prognoz wygasłych przebiegów obciążenia dobowego.

4.3 Metoda wag podobieństwa z poprawką (mwp+)

Elementy procedury są identyczne z mwp, a różnica polega na wprowadzeniu do obliczeń dynamicznie wyznaczanej korekty.

Po pierwszym etapie obliczeniowym wykonanym zgodnie z mwp, otrzymany rezultat $y_{s,t+1}$ traktuje się w taki sposób, jakby był wraz z $x_{s,t+1}$ stanem początkowym systemu s. Następnie dokonuje się prognozy „w tył” otrzymując $y_{s,t}$. Stąd poprawka względna λ :

$$\lambda = \frac{y_{s,t} - y_{s,t}}{y_{i \max} - y_{i \min}}, \quad (18)$$

gdzie:

$y_{i \max} - y_{i \min}$ – rozstęp zbioru biorącego udział w cyklu prognozy „w tył”.

Ostateczny wynik to:

$$\hat{y}_{s,t+1} = \frac{\sum y_i w_{si}}{\sum w_{si}} + \lambda(y_{i \max} - y_{i \min}) \quad (19)$$

gdzie:

$y_{i \max} - y_{i \min}$ – rozstęp zbioru punktów biorących udział w cyklu „w przód”.

4.4 Model grawitacyjny (mg)

Przyjmuje się, że zależności między punktami funkcji punktowej mają analogię w interakcjach ciał podlegających fizycznym prawom grawitacji tzn. punkty funkcji punktowej posiadają umowne masy i działają na siebie siłami przyciągania proporcjonalnymi do iloczynu mas i odwrotnie proporcjonalnymi do kwadratu odległości między nimi. Rozpatrywane są siły działające na punkt opisujący stan s -tego systemu. Masy pozostałych i -tych punktów pozostają względem siebie w stosunku odpowiadającym podobieństwu i -tych systemów do systemu s . W ten sposób punkty systemów „podobniejszych” przyciągają punkt s -ty z większą siłą niż inne. Należy określić położenie punktu o odciętej $x_{s,t+1}$ takie, aby składowa rzędna wypadkowej sił działających na ten punkt była zerowa.

Jako wartości mas można wykorzystać zdefiniowane wg wzorów (15) i (16) wagi.

Model wymaga wprowadzenia uzupełnienia w postaci ograniczenia od góry sił przyciągania między bliskimi punktami. Realizuje się to poprzez przyporządkowanie każdemu i -temu punktowi pewnej indywidualnej wielkości r_{i0} sztucznie powiększającej odległość od punktu s -tego systemu, przy wyznaczaniu wartości siły przyciągania.

Zadaniem jest odnalezienie $\hat{y}_{s,t+1}$ - wartości rzędnej punktu o odciętej $\hat{x}_{s,t+1}$ i o masie w_{ss} , dla której składowa rzędna wypadkowej sił przyciągania jest równa zero:

$$\sum_i F_i^y = 0, \quad (20)$$

gdzie:

$$F_i^y = \frac{w_{si} \cdot w_{ss}}{(r_{si} + r_{i0})^2} \sin\left(\frac{y_i - \hat{y}_{s,t+1}}{r_{si}}\right), \quad (21)$$

$$r_{si} = \sqrt{(y_i - \hat{y}_{s,t+1})^2 + (x_i - \hat{x}_{s,t+1})^2} \quad (22)$$

Dobór zbioru wielkości r_{i0} realizuje się algorytmem Rosenbrocka, w którym funkcją celu jest $\sum F_i^y = 0$ dla punktu $(x_{s,t}, y_{s,t})$. Miejsce zerowe F^y wyznacza się numerycznie (np. metodą „reguła fałsi”, zakładając, że będzie to wartość najbliższa $y_{s,t}$).

4.5 Zastosowanie MRK

Dotychczas opisane metody wyznaczały wartość funkcji punktowej systemu s w roku $t+1$ w funkcji: zadanej wartości $x_{s,t+1}$, stanu systemu w roku t i stanu wszystkich innych systemów do roku t włącznie. Zasadniczym ograniczeniem tych metod jest założenie o niewielkich zmianach w strukturze jakościowej systemów lokalnych. Zawęża to pole zastosowań tych metod do takich horyzontów czasowych, w których może być zapewnione spełnienie założenia.

W następnej kolejności zostaną przedstawione metody pozwalające na wyznaczenie wartości funkcji punktowych na rok $t+1$ w postaci najbardziej prawdopodobnej reakcji systemu na zadany rodzaj przemiany; reakcji określonej na podstawie obserwowanego zbioru przemian, jakim podlegały wszystkie systemy lokalne do roku t włącznie. Metody działające wg powyższej zasady można stosować dopóki dla rodzaju przemiany, jakiej podlega rozpatrywany system da się znaleźć odpowiedniki w skali posiadanego zbioru realizacji dla wszystkich systemów lokalnych. Dzięki temu pole zastosowań staje się szersze w porównaniu z dotychczas przedstawionymi metodami.

Wyznacznikami przemian, jakim podlegają systemy lokalne są oczywiście wskaźniki zmienności obciążeń w postaci bezpośredniej, w postaci indeksów wzrostu, lub ich funkcji.

Narzędziem użytym do realizacji tak ujętego zadania identyfikacji wartości funkcji punktowych jest metoda predykcji wektora losowego oparta na jego postaci kanonicznej.

Podstawowa procedura predykcji tzw. MRK, została opisana w pracy [1]. Metoda ta wykorzystuje sposób przekształcenia wektora losowego \mathbf{X} w wektor \mathbf{V} o składowych nieskorelowanych.

4.5.1 Postać kanoniczna wektora losowego

Dany wektor losowy \mathbf{X} posiada m składowych X_i ($i=1,2,\dots,m$) o tej właściwości, że każda i -ta składowa jest skorelowana z poprzednimi ($i-1$) składowymi oraz jest od nich zależna. Postać kanoniczną tego wektora uzyskuje się wg [1] w następujący sposób:

$$\text{dla } i=1 \quad X_{01} = V_1, \quad (23)$$

$$i>1 \quad X_{0i} = \sum_{j=1}^{i-1} a_{ij} V_j + V_i, \quad (24)$$

gdzie:

- $X_{0i}=X_i-m_{xi}$ - wartość scentrowana składowej X_i ,
- $m_{xi}=E(X_i)$ - wartość średnia,
- V_i - nieskorelowane ze sobą składowe wektora V o zerowych wartościach średnich.

Współczynniki rozkładu a_{ij} wyznacza się na podstawie pewnej liczby obserwacji składowych wektora X :

$$a_{ii} = \frac{k_{ii}}{\text{Var}(V_i)}, \quad (25)$$

$$\text{dla } j>1 \ a_{ij} = \frac{1}{\text{Var}(V_j)} \left(k_{ij} - \sum_{z=1}^{j-1} a_{iz} a_{jz} \text{Var}(V_z) \right), \quad (26)$$

gdzie:

- dla $i \neq j$ $k_{ij}=\text{Cov}(X_{0i},X_{0j})$ – kowariancja zmiennych X_{0i} i X_{0j} ,
- $k_{ii}=\text{Var}(X_{0i})$ – wariancja zmiennej X_{0i} .

4.5.2 Tok predykcji

Jeśli wejściem do prognozy jest $n-1$ z n składowych wektora X to prognoza n -tej składowej:

$$\hat{X}_n = \sum_{j=1}^{n-1} a_{nj} V_j + \hat{V}_n + m_{xn} \quad (27)$$

Po wykonaniu (na podstawie danych z historii procesu) przekształcenia wektora X na wektor V , niewiadomą w (27) stanowi jedynie reszta \hat{V}_n . Predykcja polega na wygenerowaniu najbardziej prawdopodobnej wartości \hat{V}_n z empirycznego rozkładu gęstości prawdopodobieństwa warunkowego $g(V_n | X_{0\ n-1})$. Wykorzystuje się do tego macierz dystrybuant empirycznych $F(V_n | X_{0\ n-1})$, z których każda odwzorowana jest przy pomocy linii łamanej skonstruowanej z odcinków prostych, których krańce odpowiadają granicom klas, na jakie dzieli się zakres zmienności V_n . Dokładność predykcji \hat{V}_n zależy od liczby klas na jakie dzieli się zakres zmienności $V_n : l_v$, oraz liczby klas na jakie dzieli się zakres zmienności $X_{0\ n-1} : l_x$. Wg rezultatów osiągniętych w opisanym dalej postępowaniu celową okazuje się indywidualna optymalizacja wielkości l_x, l_v przeprowadzona dla osiągnięcia maksymalnej dokładności predykcji.

4.5.3 Adaptacja MRK do zadania identyfikacji wartości funkcji punktowej

Przedstawiono sposób zastosowania MRK do zadania identyfikacji wartości funkcji punktowej $y = f(x)$ dla żądanego x . W dalszej części wektor o składowych skorelowanych oznaczany będzie jako Z (do tego miejsca stosowano oznaczenie X - zgodnie z materiałami źródłowymi).

1 Przygotowanie zbioru danych wejściowych

Każdą funkcję punktową przyjmujemy za ciąg kolejnych realizacji odpowiedniego procesu. Każda realizacja to para odpowiadających sobie parametrów (x,y) opisujących pewien stan danego systemu elektroenergetycznego. Poniżej przedstawiono sposoby komponowania składowych wektora losowego Z , realizujących rolę wyznaczników przemian, jakim podlegają ewoluujące systemy lokalne. Predykcji podlega ostatnia składowa wektora.

– Wektor o czterech składowych (4S):

$$Z_1 = x_{\tau-1}, \quad (28a)$$

$$Z_2 = y_{\tau-1}, \quad (b)$$

$$Z_3 = x_{\tau}, \quad (c)$$

$$Z_4 = y_{\tau}, \quad (d)$$

gdzie:

dla $\tau=t+1$: Z_1, Z_2, Z_3 jest wejściem do prognozy, a wynikiem prognozy jest:

$$\hat{y}_{t+1} = \hat{Z}_4. \quad (29)$$

– Wektor o dwóch składowych (2S):

$$Z_1 = G_1 = \ln(y_{\tau-1}) / \ln(x_{\tau-1}), \quad (30a)$$

$$Z_2 = G_2 = \ln(y_{\tau}) / \ln(x_{\tau}), \quad (b)$$

gdzie:

$G = 1 - D$, D - wymiar Hausdorffa [2],

dla $\tau=t+1$: Z_1 - wejście do prognozy,

\hat{Z}_2 - wyjście, stąd szukany wynik:

$$\hat{y}_{t+1} = e^{\hat{Z}_2 \ln(x_{t+1})}. \quad (31)$$

– Wektor o trzech składowych (3S):

$$Z_1 = G_1 = \ln(y_{\tau-2}) / \ln(x_{\tau-2}), \quad (32a)$$

$$Z_2 = \ln(y_{\tau-2} \cdot y_{\tau-1}) / \ln(x_{\tau-2} \cdot x_{\tau-1}), \quad (b)$$

$$Z_3 = \ln(y_{\tau-2} \cdot y_{\tau-1} \cdot y_{\tau}) / \ln(x_{\tau-2} \cdot x_{\tau-1} \cdot x_{\tau}), \quad (c)$$

gdzie:

dla $\tau=t+1$: Z_1, Z_2 – wejście prognozy,

\hat{Z}_3 - wyjście, stąd wynik:

$$\hat{y}_{t+1} = \frac{e^{\hat{Z}_3 \ln(x_{t-1} \cdot x_t \cdot x_{t+1})}}{y_{t-1} \cdot y_t} \quad (33)$$

Kolejne trzy sposoby polegają na przekształceniu każdego z powyższych wektorów \mathbf{Z} w wektor \mathbf{Z}^* , którego składowe są wskaźnikami wzrostu w interwale T odpowiednich składowych wektora \mathbf{Z} .

– Wektor o czterech składowych (W4S):

$$Z_1^* = x_{\tau-1} / x_{\tau-1-T}, \quad (34a)$$

$$Z_2^* = y_{\tau-1} / y_{\tau-1-T}, \quad (b)$$

$$Z_3^* = x_{\tau} / x_{\tau-T}, \quad (c)$$

$$Z_4^* = y_{\tau} / y_{\tau-T}, \quad (d)$$

gdzie:

T-interwał wskaźnika wzrostu (w latach),

dla $\tau=t+1$: poszukiwane rozwiązanie:

$$\hat{y}_{t+1} = \hat{Z}_4^* \cdot y_{t+1-T}. \quad (35)$$

– Wektor o dwóch składowych (W2S):

$$Z_1^* = \frac{\ln(C \cdot y_{\tau-1}) / \ln(C \cdot x_{\tau-1})}{\ln(C \cdot y_{\tau-1-T}) / \ln(C \cdot x_{\tau-1-T})}, \quad (36a)$$

$$Z_2^* = \frac{\ln(C \cdot y_{\tau}) / \ln(C \cdot x_{\tau})}{\ln(C \cdot y_{\tau-T}) / \ln(C \cdot x_{\tau-T})}, \quad (b)$$

gdzie:

C – stała (C>1) zabezpieczająca przed zerowaniem się licznika wyrażeń (36); dla $\tau=t+1$ szukana wartość:

$$\hat{y}_{t+1} = \frac{1}{C} \cdot e^{\frac{\ln(C \cdot y_{t+1-T}) \cdot \ln(C \cdot x_{t+1}) \cdot \hat{Z}_2^*}{\ln(C \cdot x_{t+1-T})}}. \quad (37)$$

– Wektor o trzech składowych (W3S):

$$Z_1^* = \frac{\ln(C \cdot y_{\tau-2}) / \ln(C \cdot x_{\tau-2})}{\ln(C \cdot y_{\tau-2-T}) / \ln(C \cdot x_{\tau-2-T})}, \quad (38a)$$

$$Z_2^* = \frac{\frac{\ln(C \cdot y_{\tau-2} \cdot y_{\tau-1})}{\ln(C \cdot x_{\tau-2} \cdot x_{\tau-1})}}{\frac{\ln(C \cdot y_{\tau-2-T} \cdot y_{\tau-1-T})}{\ln(C \cdot x_{\tau-2-T} \cdot x_{\tau-1-T})}}, \quad (b)$$

$$Z_3^* = \frac{\frac{\ln(C \cdot y_{\tau-2} \cdot y_{\tau-1} \cdot y_{\tau})}{\ln(C \cdot x_{\tau-2} \cdot x_{\tau-1} \cdot x_{\tau})}}{\frac{\ln(C \cdot y_{\tau-2-T} \cdot y_{\tau-1-T} \cdot y_{\tau-T})}{\ln(C \cdot x_{\tau-2-T} \cdot x_{\tau-1-T} \cdot x_{\tau-T})}}, \quad (c)$$

gdzie dla $\tau=t+1$ wynik:

$$\hat{y}_{t+1} = \frac{e^{\frac{\hat{Z}_3^* \ln(C \cdot y_{t-1-T} \cdot y_{t-T} \cdot y_{t+1-T}) \ln(C \cdot x_{t-1} \cdot x_t \cdot x_{t+1})}{\ln(C \cdot x_{t-1-T} \cdot x_{t-T} \cdot x_{t+1-T})}}}{C \cdot y_{t-1} \cdot y_t}. \quad (39)$$

K-letnia statystyka dostarcza k realizacji w czasie dla każdego systemu. Zakładając, że systemy podlegają tym samym prawom rządzącym ich zmiennością obciążeń, można poszerzyć zbiór realizacji przez odwrócenie procesu w czasie, otrzymując w rezultacie po 2·k-1 punktów na każdy system.

W celu zapewnienia nietrywialnego charakteru przebiegu dystrybuant, dokonuje się sztucznego rozmnożenia zbioru danych. Dla każdego systemu osobno, dzieli się każdy przedział pomiędzy kolejnymi realizacjami wektora na taką samą liczbę L równych odcinków. W ten sposób otrzymuje się dodatkowe (L-1) punktów na każdą parę następujących po sobie realizacji, dla każdego systemu. Liczbę L dobiera się tak aby łącznie otrzymać nie mniej niż 260 realizacji w każdej z podanych postaci wektorów. Zgodność rozkładu prawdopodobieństwa tak zmodyfikowanych składowych z rozkładem oryginalnych składowych, zweryfikowano pozytywnie statystycznym testem serii Walda-Wolfowitza.

2 Problem optymalizacji wymiaru macierzy dystrybuant warunkowych

Poszukiwane są l_z i l_v , dla których otrzymuje się najlepsze dokładności predykcji. Przestrzeń poszukiwań określono arbitralnie: $l_z, l_v \in \langle 1, 256 \rangle$ co daje 65536 możliwych par (l_z, l_v).

Kryterium doboru (l_z, l_v) jest najmniejszy błąd średniokwadratowy liczony z populacji uzyskanych w wyniku predykcji estymatorów każdej oryginalnej realizacji ostatniej składowej wektora losowego \mathbf{Z} lub \mathbf{Z}^* :

$$\min \{ \Delta \}, \quad (40a)$$

$$\Delta = \sqrt{\frac{\sum_{j=1}^r (\hat{Z}_{n,j} - Z_{n,j})^2}{r}} = f(l_z, l_v), \quad (b)$$

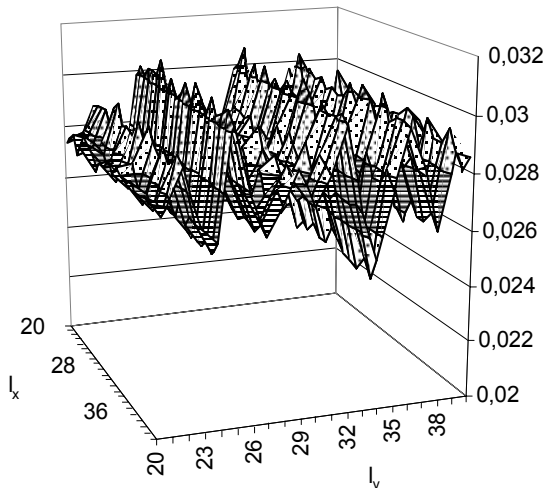
$$\text{lub } \Delta = \sqrt{\frac{\sum_{j=1}^r (\hat{Z}_{n,j}^* - Z_{n,j}^*)^2}{r}} = f(l_z, l_v), \quad (c)$$

gdzie:

r - liczba oryginalnych realizacji wektora losowego (takich, których nie pozyskano

w wyniku odwracania procesu lub rozmnażania zbioru wejściowego).

Optymalizację $\Delta(l_z, l_v)$ należy przeprowadzić dla każdej z 51 funkcji punktowych potrzebnych do prognozy przebiegów obciążenia dobowego w dniu zimowym i letnim. Daje to 306 wykonań procedury optymalizacyjnej.



Rys. 3. Fragment funkcji $\Delta(l_z, l_v)$.

Ze względu na czas trwania obliczeń sukcesywne przeszukiwanie enumeracyjne w tak rozległym zakresie jest nieefektywne. Zaś ze względu na wielomodalny i dyskretny charakter funkcji $\Delta(l_z, l_v)$ proste gradientowe metody wyznaczania ekstremum globalnego nie mogą dać rezultatów.

Wstępnie użyto do minimalizacji funkcji $\Delta(l_z, l_v)$ połączenia metody błędzenia losowego z metodą gradientową, która dała umiarkowanie dobre wyniki. Jednakże próby zastosowania algorytmu genetycznego (AG) okazały się efektywniejsze.

3 Opis algorytmu genetycznego (AG)

W AG użytym do minimalizacji funkcji $\Delta(l_z, l_v)$ zastosowane zostało kodowanie binarne 16-bitowych ciągów kodowych. Bity 1-8 oznaczają wartość l_z , bity 9-16: l_v . Dekodowanie ciągu kodowego polega na przeliczeniu dwóch przedstawionych liczb w postaci dwójkowej na liczby w systemie dziesiętnym i dodanie do obu wartości 1. W ten sposób uzyskuje się możliwy zakres zmienności l_z, l_v od 1 do 256.

Parametry AG:

- stała licznosc populacji $l_{pop} = 40$;
- liczba pokoleń $l_{gen} = 40$;

- losowy wybór punktu krzyżowania z dopuszczeniem do reprodukcji osobników bez krzyżowania;
- prawdopodobieństwo krzyżowania $p_{cross} = 0.9$;
- prawdopodobieństwo mutacji $p_{mut} = 0.01$;
- funkcja przystosowania:

$$\text{dla } C > \Delta \quad F_{fit}(l_z, l_v) = C - \Delta, \quad (41a)$$

$$\text{dla } C \leq \Delta \quad F_{fit}(l_z, l_v) = 0, \quad (b)$$

gdzie:

- C - stała, której wartość powinna zabezpieczać otrzymanie dodatnich wartości F_{fit} w dużym obszarze przestrzeni poszukiwań;
- reprodukcja wg „reguły ruletki” wykalibrowanej proporcjonalnie do przystosowania ciągów kodowych:

$$w_{fit}(k) = \frac{F_{fit}(l_z, l_v)_k}{\sum_{j=1}^{l_{pop}} F_{fit}(l_z, l_v)_j}, \quad (42)$$

gdzie:

- w_{fit} - wskaźnik przystosowania k-tego ciągu kodowego w l_{pop} - licznej populacji;
- liniowe skalowanie przystosowania ciągów kodowych zapewniające regulację oczekiwanej liczby kopii w populacji potomnej; wartość współczynnika zwielokrotnienia ustalono na 2.0;

$$\begin{cases} F'_{fit\ sr} = F_{fit\ sr} \\ F'_{fit\ max} = C_{mult} \cdot F_{fit\ sr} \\ F'_{fit} = a \cdot F_{fit} + b \end{cases} \quad (43)$$

gdzie:

C_{mult} - wskaźnik zwielokrotnienia; a, b wyznacza się wg pierwszych dwóch równań;

jeśli wg (43) $F'_{fit\ min} < 0$ wtedy:

$$\begin{cases} F'_{fit\ sr} = F_{fit\ sr} \\ F'_{fit\ min} = 0 \\ F'_{fit} = a \cdot F_{fit} + b \end{cases} \quad (44)$$

Przed wykonaniem l_{gen} cykli AG, dokonuje się zainicjowania populacji ciągów kodowych. Generuje się losowo ciąg $16 \times l_{pop}$ zer i jedynek, którymi zapełnia się ciągi kodowe. Wyznacza się ich przystosowanie wg (42), i przystosowanie wyskalowane wg (43) lub (44). Następnie wykonuje się l_{gen} cykli złożonych ze standardo-

wych operatorów AG tzn. reprodukcji, krzyżowania i mutacji.

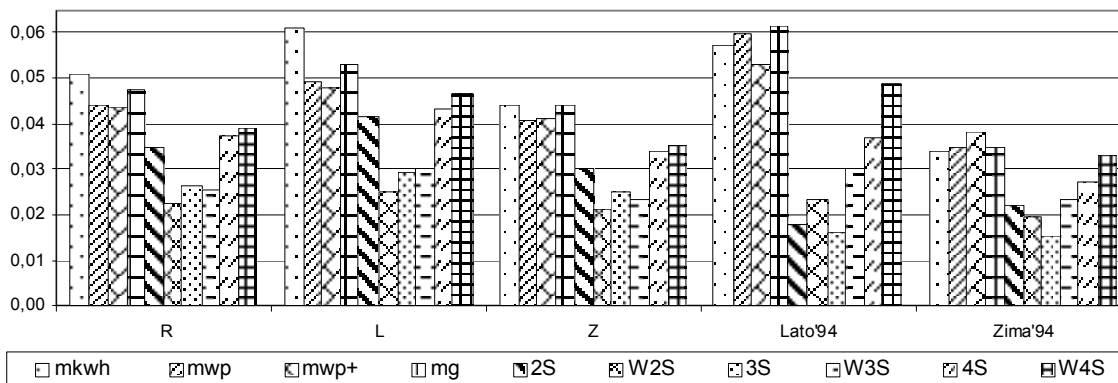
5 STATYSTYCZNA ANALIZA DOKŁADNOŚCI BADANYCH METOD PROGNOZOWANIA

Bazą do badań były statystyki obejmujące: przebiegi obciążenia dobowego systemów przedsiębiorstw dystrybucyjnych energetyki, z dób roboczych grudnia i lipca lat 1990÷1994, przebiegi obciążenia dobowego Krajowego Systemu Elektroenergetycznego z dób jw., zestaw miesięcznych wielkości sprzedaży energii w 33 spółkach z lat 1990÷1997, oraz dane dotyczące rocznych strat energii w przedsiębiorstwach dystrybucyjnych.

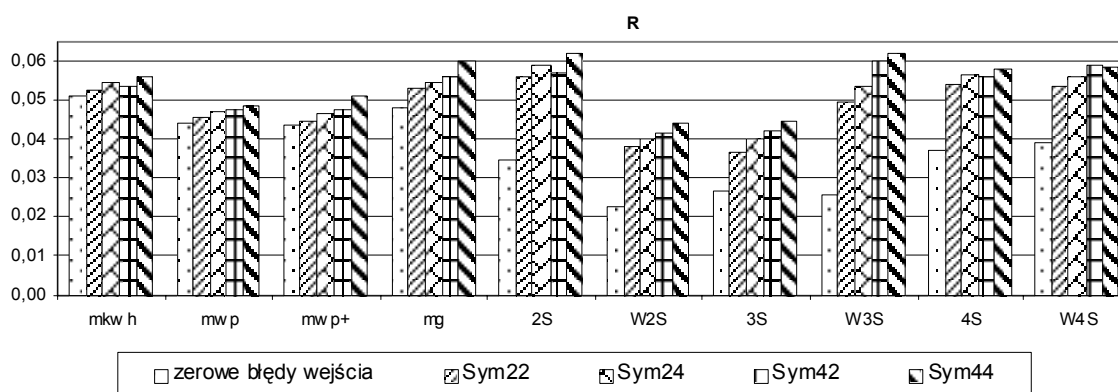
Ze zróżnicowanego w trzech przekrojach (dla dób lata (L), zimy (Z) i łącznego (R)) zbioru błędów prognoz wygasłych z wyprzedzeniem rocznym, wykonanych dla ośmiu systemów

w maksymalnym możliwym do wykonania zakresie, dla każdej z metod predykcji wykonano następujące czynności: sprawdzono warunki nieobciążoności prognoz, zweryfikowano zgodność z dopasowanymi teoretycznymi rozkładami prawdopodobieństwa; wyznaczono: przedziały błędów prawdopodobnych (tj. takich, że prawdopodobieństwo wystąpienia błędów nie większych od nich $P\{\delta_p \geq \delta\} = 0,68$) i wysocę prawdopodobnych ($P\{\delta_{wp} \geq \delta\} = 0,95$), średnie błędy absolutne, dystrybuanty empiryczne błędów absolutnych i inne przydatne w ocenie właściwości populacji błędów statystyki.

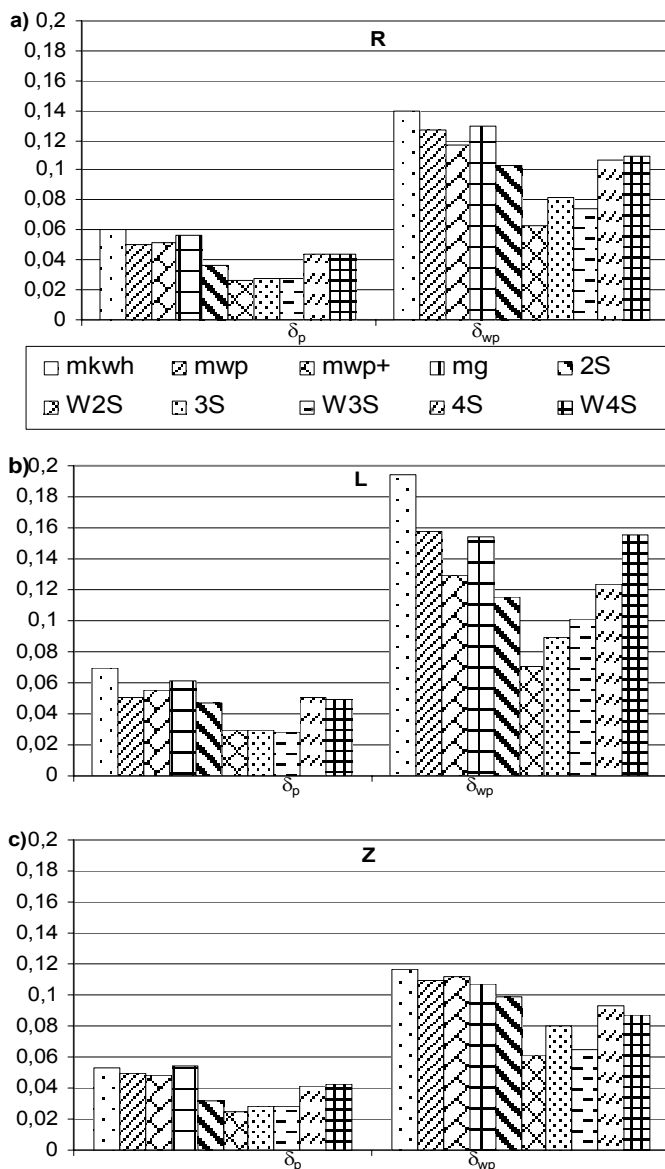
Biorąc pod rozwagę specyfikę algorytmu prognozy, który bazuje na otrzymanych odrębnymi metodami prognozach rocznego zapotrzebowania na energię i moc szczytową, przeprowadzono dodatkowe badanie mające na celu określenie stopnia wrażliwości rozpatrywanych metod na niedokładności prognoz wejściowych.



Rys. 4. Absolutne błędy średnie prognoz wygasłych (L-doby letnie, Z – doby zimowe, R – doby letnie i zimowe łącznie).



Rys. 5. Absolutne błędy średnie z symulacji Sym22, Sym24, Sym42, Sym44 dla zbioru łącznego.



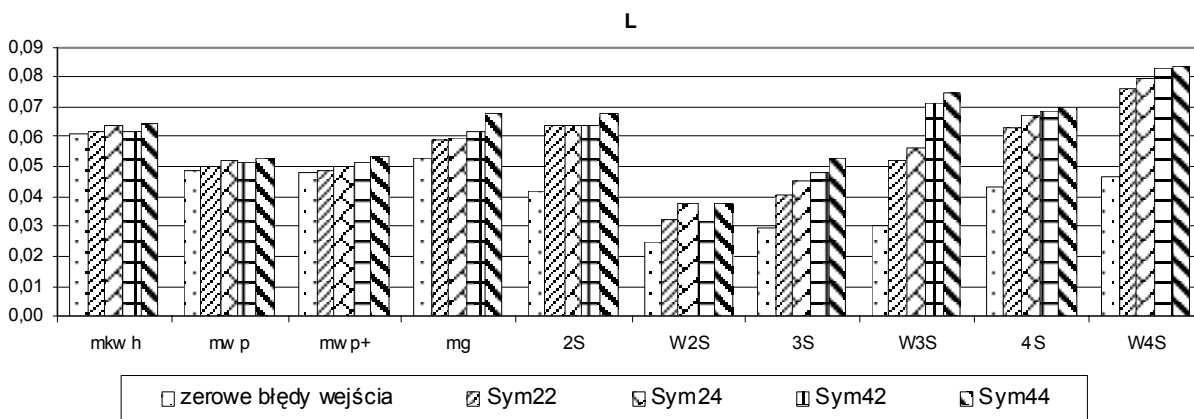
Rys. 6. Błędy prawdopodobne i wysoce prawdopodobne prognoz wygasłych (L-doby letnie, Z – doby zimowe, R – doby letnie i zimowe łącznie).

Badanie to polegało na ocenie właściwości zbioru błędów otrzymanego w wyniku wielokrotnego powtarzania cyklu prognozy z danymi wejściowymi każdorazowo obciążanymi odchyleniem pseudolosowym z rozkładu prawdopodobieństwa o ściśle określonych parametrach.

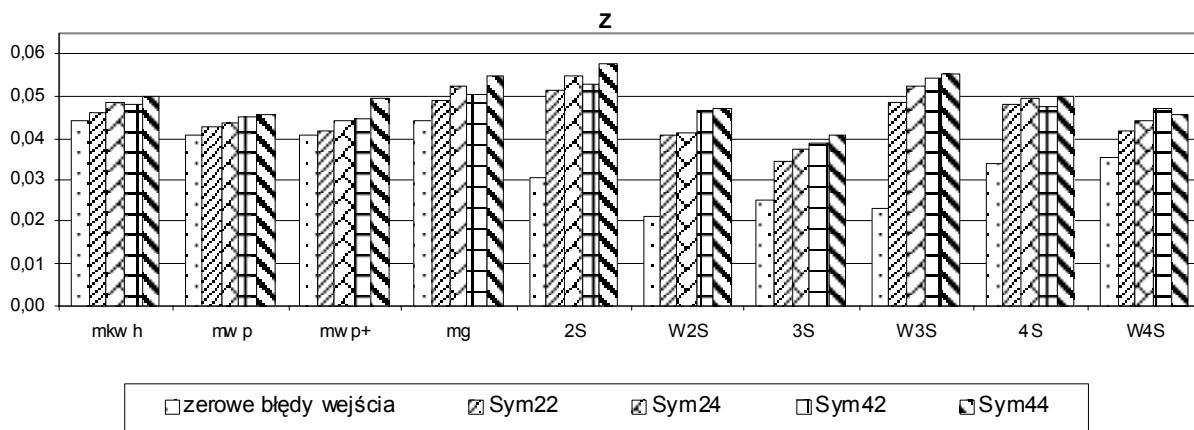
Wykonano 4 symulacje, w których obciążano energią roczną i moc szczytową parą błędów względnych losowanych z rozkładu normalnego o zerowej średniej i o odchyleniach standardowych odpowiednio: 0,02 i 0,02 (Sym22), 0,02 i 0,04 (Sym24), 0,04 i 0,02 (Sym42), oraz 0,04 i 0,04 (Sym44).

5.1 Wnioski

- ◆ Metody wykorzystujące procedury identyfikacji wartości funkcji punktowych działające w oparciu o MRK charakteryzują się dużą wrażliwością na jakość danych wejściowych. Szczególnie duży spadek dokładności stwierdzono dla metod 2S i W3S. Jednak najlepsze oceny pozostają przy wysoko ocenionych na podstawie „czystych prognoz wygasłych” metodach W2S i 3S.
- ◆ Cechą stałą wydaje się być wyraźnie zaznaczona przewaga metody W2S w zastosowaniu do prognozy przebiegów letnich, a metody 3S w odniesieniu do przebiegów zimowych.
- ◆ W warunkach wysokiej niepewności prognozy zapotrzebowania na energię i moc szczytową celowe może być stosowanie metod mwp i mwp+ jako stosunkowo mało wrażliwych i prostszych w użyciu.

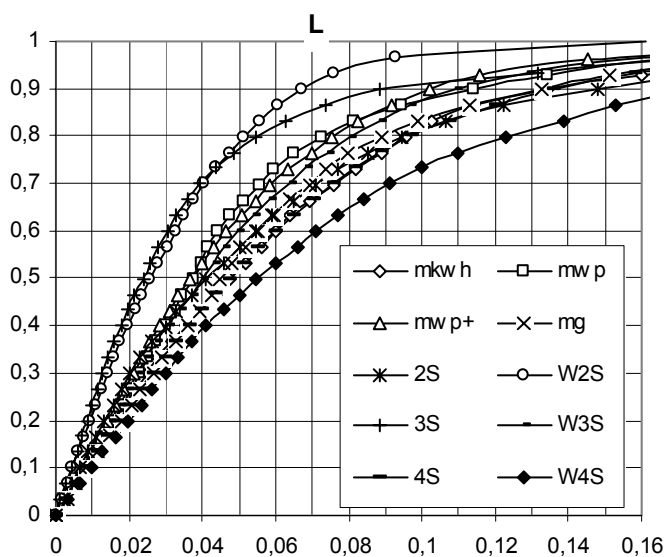


Rys. 7. Absolutne błędy średnie z symulacji Sym22, Sym24, Sym42, Sym44 dla zbioru błędów dób letnich.



Rys. 8. Absolutne błędy średnie z symulacji Sym22, Sym24, Sym42, Sym44 dla zbioru błędów dób zimowych.

◆ Typ rozkładu prawdopodobieństwa zbiorów łącznych błędów prognozy jest efektem ich składania się z wielu prób prostych (zbiorów błędów w poszczególnych godzinach danej doby) o rozkładach normalnych (lub zbliżonych), o bliskich zeru wartościach średnich i różnych wariancjach. Skutkuje to kształtem rozkładu zbioru łącznego zbliżonym do rozkładu Weibulla.



Rys. 9. Dystrybuanta empiryczna błędów dla dób letnich z symulacji Sym22.

6 UWAGI KOŃCOWE

Skuteczność działania opracowanych metod identyfikacji zbadano dla jednorocznego horyzontu prognozy, co stanowi wg przyjętych definicji górną granicę wyprzedzeń dla prognoz średnioterminowych. Nie ma technicznych

przeszkód do stosowania przedstawionej metodyki do prognoz w dłuższych horyzontach, choć należy wziąć pod uwagę, że w takim wypadku każda prognoza dla kolejnego roku byłaby wykonywana krokowo tzn. w trybie prognoza-z-prognozy, co może prowadzić do oczywistego pogorszenia dokładności. Równie gruntowna jak przeprowadzona na potrzeby tej pracy weryfikacja rozpatrywanych metod dla dłuższych horyzontów, wymagałaby odpowiednio obszerniejszych, w stosunku do posiadanych, danych statystycznych. W związku z tym przydatność przedstawionych metod w prognozowaniu długoterminowym w przeciwieństwie do średnioterminowego należy uznać za nierozstrzygniętą.

Specyfika przedstawionych metod sprawiła, że postać zbioru informacji potrzebnych do prognozy przebiegów obciążenia dla pojedynczego systemu nie jest tak zwarta i prosta w wykorzystaniu jak miało to miejsce w przypadku historycznego modelu prognozy. Zamiast zestawu kilkudziesięciu matematycznych wzorów krzywych regresji dla poszczególnych charakterystyk wskaźników opisujących stan systemów elektroenergetycznych, przetwarzane są zbiory par liczb stanowiące funkcje punktowe, będące efektem obróbki materiału statystycznego, obejmującego minimalnie okres od jednego do trzech lat, w zależności od metody. Pomimo to, dane takie nadal stanowią podstawowe wielkości bezpośrednio dotyczące prognozowanego procesu obciążenia dobowego. Utrudnienia wyniknąć mogą jedynie z faktu zapotrzebowania na dane dotyczące nie tylko systemu, dla którego wykonywana jest prognoza, ale również pewnej ilości innych systemów, w miarę możliwości stanowiących przekrój pod

względem rozpiętości możliwych struktur odbiorców, kierunków rozwoju gospodarczego i dynamiki tego rozwoju. Ilość ta jest jednocześnie liczbą systemów będących podmiotem przeprowadzonych tu prognoz i stanowi ok. 1/4 wszystkich systemów odpowiadających obszarowo przedsiębiorstwom dystrybucyjnym elektroenergetyki w kraju. Subiektywnie oceniając ilość ta wydaje się wystarczająca, choć z pewnością powiększenie jej, jak i rozciągnięcie okresu obserwacji (zwłaszcza dla metod bazujących na MRK) mogłoby mieć pozytywny wpływ na wyniki.

W pracy, której niniejsze streszczenie dotyczy, zgodnie z brzmieniem tezy wykazano, że stosując odpowiednie metody obliczeniowe, na bazie statystyk zwanych funkcjami punktowymi, można wyznaczyć prognozowane wartości wskaźników opisujących zmienność obciążenia lokalnych systemów elektroenergetycznych, z precyzją wystarczającą do skonstruowania użytecznej prognozy dobowych przebiegów obciążenia. W szczególności zaś:

- omówiono obszernie historyczny model analizy zmienności obciążeń,
- przedstawiono ogólny algorytm prognozy,
- przeprowadzono analizę zmienności obciążeń współczesnych systemów lokalnych wg historycznych formuł,
- przedstawiono 10 rozwiązań metod identyfikacji wartości funkcji punktowych,
- przedstawiono wyniki statystycznej i teoretycznej analizy dokładności ww. metod,
- wskazano efektywne rozwiązania.

Elementy nowości:

- przeprowadzenie teoretycznej (symulacyjnej) analizy dokładności (historycznego) algorytmu prognozy;
- spośród rozwiązań metodologicznych identyfikacji wartości funkcji punktowych całkowicie autorskimi rozwiązaniami są: mwp, mwp+ i mg;
- adaptacja znanej wcześniej metody predykcji wektora losowego – MRK do wyznaczania wartości funkcji punktowych (metody: 2S, W2S, 3S, W3S, 4S, W4S);
- efektywne zastosowanie algorytmu genetycznego do optymalizacji wymiarów macierzy dystrybuant warunkowych w MRK.

BIBLIOGRAFIA

- 1) Dobrzańska I. 1991. MRK - *Parametryczna metoda predykcji wektora losowego*. Częstochowa: monografia.
- 2) Dobrzańska I. 1991. Wymiar Hausdorffa jako narzędzie predykcji krokowej. *Badania operacyjne i decyzje*. Nr 2.
- 3) Dobrzańska I., Dąsał K., Łyp J., Popławski T. 1995. Metodyka prognozowania zapotrzebowania na moc i energię wraz z wykresami obciążenia dla typowych dób reprezentatywnych dla kraju i regionów systemu elektroenergetycznego. Dokumentacja PBZ-10-03, Częstochowa.
- 4) Dobrzańska I., Gallus E. 1962. Wykresy obciążeń doby grudniowej i lipcowej polskiego systemu energetycznego. Instytut Energetyki. Nr inw. 2871. Katowice.
- 5) Goldberg D. E. 1995. *Algorytmy genetyczne i ich zastosowania*. Warszawa: WNT.
- 6) Kopecki K. Analiza zmienności obciążeń systemów energetycznych w Polsce. *Zeszyty naukowe Politechniki Gdańskiej*. Elektryka Nr 2/1958.
- 7) Malko J. 1998. Restrukturyzacja elektroenergetyki i zadania prognozowania – doświadczenia amerykańskie. *Materiały konferencyjne: IV Konferencja Naukowa Prognozowanie w elektroenergetyce PE'98*. Częstochowa.
- 8) Ministerstwo Gospodarki. Założenia polityki energetycznej Polski do roku 2020. <http://www.cup.gov.pl/prawo/energia>.
- 9) Praca zbiorowa. 1971. *Analiza i prognoza obciążeń elektroenergetycznych*. Warszawa: WNT.
- 10) Prawo energetyczne. Dz. U. Nr 54, poz. 348.
- 11) Rozporządzenie Ministra Gospodarki Dz.U. Nr 135, poz. 881.

