

IDENTYFIKACJA KOMPONENTÓW DESTRUKCYJNYCH W MODELACH PREDYKCYJNYCH W PODEJŚCIU WIELOMODELOWYM

Ryszard Szupiluk, Paweł Rubach

Kolegium Analiz Ekonomicznych

Szkoła Główna Handlowa w Warszawie

e-mail: ryszard.szupiluk@sgh.waw.pl; pawel.rubach@sgh.waw.pl

Streszczenie: W niniejszym artykule przedstawimy metodę identyfikacji komponentów destrukcyjnych występujących w podejściu wielomodelowym wykorzystującym algorytmy ślepej separacji sygnałów. Ocena charakterystyki poszczególnych komponentów dokonana zostanie na podstawie autorskich mierników zmienności/gładkości sygnałów. W celu potwierdzenia skuteczności prezentowanej metody przedstawimy praktyczny eksperyment poprawy wyników prognozy zużycia energii elektrycznej.

Słowa kluczowe: predykcja, agregacja modeli, ślepa separacja, identyfikacja szumów

WPROWADZENIE

Rosnąca popularność podejścia wielomodelowego jest naturalną konsekwencją rozwoju metod tworzenia modeli w takich dyscyplinach jak Data Mining, uczenie maszynowe, sztuczna inteligencja czy modelownie nieparametryczne. Dojrzałość podstawowych metod predykcji powoduje, że dalszej poprawy jakości predykcji upatrywać można w działaniach prowadzących do połączenia wiedzy pochodzącej z różnych modeli ujmujących badane zjawisko w odmienny sposób. Jest wiele popularnych metod agregacji, z których większość oparta jest na różnych formach uśredniania wyników lub parametrów modeli [Clements 1989, Gatnar 2008]. Jednym z zasadniczych ograniczeń takich metod jak boosting lub bagging jest zawężenie agregacji do modelu w zasadzie tego samego typu, gdzie zróżnicowanie wiąże się przede wszystkim z uczeniem na różnych podzbiorach [Breiman 1996]. W istotny sposób ogranicza to ideę agregacji różnych modeli jako koncepcję łączenia różnych źródeł informacji.

Tych ograniczeń nie posiada podejście agregujące oparte na metodach ślepej separacji [Szupiluk i in. 2004, Szupiluk 2013]. W podejściu tym, traktujemy zbiór rezultatów predykcji generowanych przez różne modele, jako jedną wielowymiarową zmienną. Zakładamy, że zmienna ta zawiera ukryte komponenty konstruktywne i destrukcyjne dla wyniku prognozy. Oczekujemy przy tym, że owe ukryte komponenty bazowe mają pewien „fizyczny charakter”, zaś eliminacja komponentów destrukcyjnych powinna skutkować poprawą prognozy mierzoną różnymi kryteriami. Ten aspekt eliminacji „niejako” fizycznych zakłóceń zidentyfikowanych metodami ślepej separacji sygnałów (ang. Blind Signal Separation BSS) odróżnia prezentowane podejście od agregacji opartej na optymalizacji wybranego pojedynczego kryterium błędu.

W niniejszym artykule skoncentrujemy się na aspekcie identyfikacji komponentów destrukcyjnych. Jest to zagadnienie o tyle złożone, że komponenty owe mają zwykle charakterystykę statystyczną znacząco odmienną od klasycznych białych szumów – często przyjmowanych jako elementy zakłócające. Powoduje to konieczność poszukiwania nowych miar i charakterystyk – w stosunku do standardowych metod opartych na metodach korelacyjnych, R/S lub spektralnych [Hurst 1951, Vasegi 1997, Therrien 1992]. W niniejszym artykule rozwiniemy klasę miar zmienności mogących stanowić alternatywę lub uzupełnienie dla owych klasycznych metod. Całokształt rozważań zostanie zobrazowany praktycznym przykładem prognozowania energii.

AGREGACJA BSS

Załóżmy istnienie zbioru m rezultatów predykcji x_i , $i = 1, \dots, m$. Wyniki te zbierzemy w jednej wielowymiarowej zmiennej $\mathbf{x} = [x_1, \dots, x_m]^T$. Przyjmijmy dalej, że wyniki są akceptowalnej jakości co oznacza, iż prognozowane wartości w pewnym stopniu odpowiadają wartościom obserwowanym a w pewnym się różnią. Możemy więc powiedzieć, że dany rezultat jest kombinacją komponentów konstruktywnych \hat{s}_j , $j = 1, \dots, p$ związanych z podobieństwem oraz destrukcyjnych \tilde{s}_l , $l = 1, \dots, q$ związanych z różnicami wartości przewidywanych i obserwowanych. Wszystkie te komponenty potraktujemy jako ukryte komponenty bazowe zawarte w wielowymiarowej zmiennej \mathbf{x} . W przypadku liniowego sposobu (systemu) mieszania można więc zapisać

$$\mathbf{x}(k) = \mathbf{A}\mathbf{s}(k), \quad (1)$$

gdzie k oznacza numer obserwacji lub indeks czasu, macierz $\mathbf{A} = [a_{ij}] \in \mathbb{R}^{m \times n}$ reprezentuje system mieszający zaś wektor $\mathbf{s}(k) = [\hat{s}_1(k), \dots, \hat{s}_p(k), \tilde{s}_{p+1}(k), \dots, \tilde{s}_{p+q}(k)]^T$ reprezentuje zbiór komponentów bazowych, gdzie $n = p + q$. Dla uproszczenia przyjmijmy, że $m = n$, macierz \mathbf{A} jest nieosobliwa, zaś $E\{\mathbf{s}\} = \mathbf{0}$. Zauważmy dalej,

że identyfikując system mieszający \mathbf{A} oraz komponenty bazowe \mathbf{s} oraz eliminując komponenty destrukcyjne (stawiając odpowiednio $\tilde{s}_i = 0$) otrzymamy

$$\hat{\mathbf{x}}(k) = \mathbf{A}[\hat{s}_1(k), \dots, \hat{s}_p(k), 0_{p+1}(k), \dots, 0_n(k)]^T, \quad (2)$$

gdzie $\hat{\mathbf{x}} = [\hat{x}_1, \dots, \hat{x}_m]^T$ jest „poprawianą” wersją rezultatów predykcji \mathbf{x} . Kluczowym zagadnieniem w rozważanej procedurze, jest znalezienie takiej transformacji, która rzeczywiście prowadzi do dekompozycji danych \mathbf{x} na komponenty konstruktywne i destrukcyjne. A następnie poprawne określenie które komponenty są destrukcyjne.

Punktem wyjścia w poszukiwaniu komponentów bazowych może być oczekiwanie, że mają one pewien „fizyczny” charakter związany z czynnikami determinującymi proces modelowania. Na cały proces możemy także spojrzeć jak na próbę separacji nieznanymi komponentów bazowych zmieszanych w nieznanym systemie. Prowadzi to wprost do tzw. problemu ślepej separacji i rozwiązań w nim stosowanych [Comon i in. 2010, Hyvarinien i in. 2001], co można przedstawić jako poszukiwanie transformacji odwrotnej do (1) takiej, że

$$\mathbf{y}(k) = \mathbf{W}\mathbf{x}(k) \approx \mathbf{s}(k), \quad (3)$$

gdzie macierz $\mathbf{W} = \mathbf{A}^{-1}$ pełni rolę systemu separującego odwrotnego do \mathbf{A} .

Istnieje wiele metod BSS znajdowania \mathbf{W} takich jak algorytmy analizy składowych niezależnych (ang. Independent Component Analysis-ICA) czy też algorytmy AMUSE lub SOBI [Common i in. 2010, Hyvarinien i in. 2001, Szupiluk 2013]. Znaczną część tych rozwiązań można sprowadzić, rozszerzyć lub uogólnić do przypadku eksploracji uogólnionej macierzy kowariancji postaci

$$\mathbf{R}_F = E \left\{ \left(\lambda \mathbf{y}(k) + \beta \mathbf{f}(\mathbf{y}(k)) \right) \left(\sum_{j \in J} a_j \mathbf{y}(k-j) \right)^T \right\}, \quad (4)$$

gdzie λ, β są parametrami określającymi proporcje między macierzą kowariancji oraz nieliniową macierzą kowariancji, J jest zbiorem wartości opóźnień pozwalających uwzględnić strukturę czasową sygnałów. Eksploracja macierzy (4) łączy w sobie charakterystyki wykorzystywane zarówno przez algorytmy ICA bazujące na statystykach wyższych rzędów jak również charakterystyki wykorzystywane przez algorytmy typu BSS – SOS (ang. Second Order Statistics) eksplorujące strukturę czasową sygnałów. Zauważmy, że dla $\lambda = 1$ $\beta = 0$ oraz

$J = \{0\}$ otrzymujemy zwykłą macierz kowariancji $\mathbf{R}_F = E \{ \mathbf{y}\mathbf{y}^T \}$, dla $\lambda = 1$, $\beta = 0$, $J \neq \{0\}$ otrzymujemy macierz kowariancji z opóźnieniami

$\mathbf{R}_F = E \left\{ \mathbf{y}(k) \left(\sum_{j \in J} a_j \mathbf{y}(k-j) \right)^T \right\}$ zaś dla $\lambda = 0$ $\beta = 1$, $J = \{0\}$, otrzymujemy

nieliniową macierz kowariancji $\mathbf{R}_F = E\{\mathbf{f}(\mathbf{y})\mathbf{y}^T\}$. Zasadniczą kwestią jest dobór nieliniowości występujących w (4). Teoretycznie jej optymalna postać dla algorytmów ICA ma postać $f(y) = -\partial \log(p(y)) / \partial y$ gdzie $p(y)$ jest funkcją gęstości prawdopodobieństwa zmiennej y .

Ze względu na ograniczoną informację a priori co do postaci separowanych sygnałów jako podstawa wyprowadzenia nieliniowości proponowane są najczęściej ogólne modele rozkładów pozwalające na elastyczne dopasowanie funkcji rozkładu do przetwarzanych danych.

Uwzględniając powyższe zależności można podać algorytm separacji postaci

1. Niech $\mathbf{z}(t) = \mathbf{x}(t)$, $\mathbf{W} = \mathbf{I}$
2. Oblicz generalną postać macierzy korelacji opóźnień \mathbf{R}_F zgodnie z wzorem (4)
3. Znajdź macierz ortogonalną \mathbf{Q}_p (np. za pomocą SVD), która diagonalizuje $\mathbf{R}_F + \mathbf{R}_F^T$ i wykonaj dekorelację $\mathbf{y}(t) = \mathbf{Q}_p \mathbf{z}(t)$
4. Niech $\mathbf{W} \leftarrow \mathbf{Q}_p \mathbf{W}$, $\mathbf{z}(t) \leftarrow \mathbf{y}(t)$ i powróć do kroku nr 2 dopóki nie zostanie osiągnięte zamierzone kryterium sukcesu.

Jak wcześniej powiedziano w zależności od parametrów macierzy \mathbf{R}_F otrzymujemy rodzaj eksplorowanej charakterystyki. Ma to istotny wpływ na operacjonalizację powyższego algorytmu. W przypadku dominacji statystyk drugiego rzędu liczbę iteracji można ograniczyć nawet do dwóch, co w wielu przypadkach wystarcza do efektywnej separacji. Ograniczeniem jest jednak charakterystyka separowanych sygnałów, muszą one mieć strukturę czasową o różnych funkcjach autokorelacji. Ograniczenia tego nie posiada eksploracja zależności nieliniowych, jednak w tym wypadku liczba iteracji musi być znacznie większa. Efektywne działanie algorytmu łączącego jednocześnie aspekty liniowości, nieliniowości oraz opóźnień jest otwartym i złożonym zagadnieniem.

IDENTYFIKACJA KOMPONENTÓW DESTRUKCYJNYCH

Przedstawiony powyżej schemat konceptualny definiujący proces łączenia modeli przy jednoczesnej redukcji zakłóceń, jest schematem ideowym. Zakładając jego pełną realizację moglibyśmy oczekiwać, że wydzielone komponenty destrukcyjne będą związane z konkretnymi czynnikami zniekształcającymi wyniki predykcji takie jak zakłócenia i szумы w danych uczących, nie w pełni adekwatny dobór modeli predykcyjnych czy niedokładności procesu optymalizacji. W efekcie eliminacja takiego „fizycznego” komponentu powinna skutkować poprawą predykcji na wszelkich kryteriach oceny. Owa motywacja przyświeca oparciu etapu dekompozycji na metodach ślepej separacji. Oczywiście w rzeczywistości ów czysty schemat ideowy rzadko może być w pełni zrealizowany.

W praktyce założenia co do statycznej i liniowej postaci modelu mieszającego oraz wybór określonej metody separacji np. ICA mogą być nie w pełni adekwatne. Rzeczywiste komponenty mogą być zmieszane w sposób nieliniowy lub dynamiczny zaś dla jego identyfikacji mogą być właściwe metody bazujące na nieujemności, gładkości lub rzadkości (ang. sparse) sygnałów. Należy jednak zaznaczyć, że te różnice choć mogące być istotnymi dla konkretnego przypadku nie zmieniają ogólnych ram naszego schematu badawczego. Ponieważ problematyka rozwiązywania nieliniowych oraz dynamicznych problemów BSS nadal jest w dużym stopniu otwarta, to istniejące efektywne algorytmy separacji dla przypadku liniowego, są stosowane w każdym typie modelu mieszającego. W kontekście metod ślepej separacji istotnym elementem jest także liczba sygnałów użytych w procesie mieszania/separacji, tu także nie dysponujemy wiedzą ile i jakiego typu modeli predykcyjnych należałoby użyć dla uzyskania idealnych komponentów konstrukcyjnych i destrukcyjnych. Wymiar przyjmowany jest zgodny z liczbą posiadanych modeli.

Wobec powyższych uwarunkowań można przyjąć, że choć celem dekompozycji wyników predykcji za pomocą metod ślepej separacji jest identyfikacja czystych komponentów destrukcyjnych, to w praktyce otrzymujemy najczęściej tylko pewien analityczny zbiór komponentów bazowych, których zarówno charakterystyki matematyczne jak i ich wpływ na wyniki predykcji nie są w pełni jednoznaczne. Owa niejednoznaczność oraz trudność oceny w dużej mierze wiąże się z właściwościami i charakterystyką standardowych metod i miar używanych do oceny zaszumienia i losowości sygnałów. Analiza wariancji, korelacji, autokorelacji, spektralna lub analiza R/S są efektywnymi narzędziami w przypadkach standardowych szumów modelowanych jako matematyczne białe szumy [Hamilton 1994, Shirayev 1999, Therrien 1992]. W szczególności odnosi się to do sygnałów gausowskich które są w pełni statystycznie określone statystykami drugiego rzędu. Niestety w przypadkach z jakimi mamy do czynienia przy dekompozycjach rzeczywistych sygnałów takich jak np. zużycie energii elektrycznej owe standardowe metody nie są w pełni skuteczne. Stąd potrzeba opracowań nowych koncepcji oceny zmienności i losowości sygnałów. Istnieją różne koncepcje takiej oceny [Szupiluk i inni 2010], obecnie rozwiniemy metodę opartą na zestawie miar zmienności.

Punktem wyjścia naszej koncepcji będzie powiązanie zakłóceń z szeroko rozumianą zmiennością (możliwe są tu także nawiązania do gładkości, rozproszenia, rozrzutu) dla której można zaproponować następującą miarę:

$$\Theta_1(y) = \frac{1}{N} \frac{\sum_{k=2}^N |y(k) - y(k-1)|}{I(\max(y) - \min(y))}, \quad (5)$$

gdzie: $I(u) = u$ dla $u \neq 0$ oraz $I(u) = 1$ dla $u = 0$.

Interpretacja miary (5) jest stosunkowo prosta i intuicyjna. Otrzymujemy średnią bezwzględną (całkowitą) zmienność sygnału w stosunku do zakresu wahań. Funkcja $I(u)$ pełni rolę indykatora zera i zapobiega dzieleniu przez zero dla sygnału stałego. Formuła (5) wartość maksymalną równą 1 gdy wszystkie zmiany sygnału równe są zakresowi wahań, przyjmuje zaś wartość minimalną równą 0 dla sygnału stałego. Miara (5) i jej interpretacja stanowią punkt wyjścia dla jej zmodyfikowanych wersji. Przede wszystkim w zastosowaniach optymalizacyjnych w miejsce czasem kłopotliwej wartości bezwzględnej można zastosować formułę postaci

$$\Theta_2(y) = \frac{1}{N} \frac{\sum_{k=2}^N (y(k) - y(k-1))^2}{I((\max(y) - \min(y))^2)}, \quad (6)$$

lub

$$\Theta_3(y) = \frac{1}{N} \frac{\sum_{k=2}^N \log(\cosh(y(k) - y(k-1)))}{I(\log(\cosh(\max(y) - \min(y))))}. \quad (7)$$

Powyższe miary (5)-(7) oprócz bezpośredniej miary zmienności/gładkości w różnych postaciach analitycznych można wiązać także z innymi charakterystykami. Wyrażenie (5) można także interpretować w kategoriach rzadkości (ang. sparse) sygnałów. Z kolei formułę (6) można powiązać z analizą R/S na bazie eksploracji ilorazu statystyki drugiego rzędu i rozstępu. Formuła (7) ze względu na popularną w obszarze machine learning funkcję $\log(\cosh(y))$ może być interesująco interpretowana w kategoriach statystyk wyższych rzędów [Oja 1995].

W efekcie, choć każda z powyższych miar może być wykorzystana do oceny zmienności/gładkości, to głębsze praktyczne badanie charakterystyki sygnałów dokonane powinno być raczej z łącznym użyciem powyższych miar. W tym celu możemy rozważyć charakterystyki sygnałów w pewnej przestrzeni określonej wymiarami $\Theta_1(y)$, $\Theta_2(y)$, $\Theta_3(y)$. Ze względu na łatwość graficznej inspekcji wygodnym może być rozważanie par charakterystyk w układzie kartezjańskim, co np. dla $\Theta_1(y)$, $\Theta_2(y)$ prowadzi do możliwości intuicyjnej miary łącznej zmienności postaci

$$\Theta_E(y) = \sqrt{(\Theta_1(y))^2 + (\Theta_2(y))^2}. \quad (8)$$

Miara (8) określa odległość euklidesową sygnału od środka układu określonego przez charakterystyki $\Theta_1(y)$, $\Theta_2(y)$ i może być rozszerzona na większą liczbę zmiennych.

EKSPERYMENT PRAKTYCZNY

W niniejszym eksperymencie przedstawimy skuteczność prezentowanej metody dla poprawy predykcji krótkoterminowego zużycia energii elektrycznej. W badaniu wykorzystano godzinowe dane o łącznym zużyciu energii pochodzące z polskiego systemu energetycznego z lat 1987-1996. Dla agregacji wykorzystano sześć modeli neuronowych typu MLP optymalizowanych algorytmem Levenberga-Marquarda. Do agregacji wykorzystano próbę 2000 obserwacji. W efekcie uzyskano prognozy o następujących wielkościach błędu MSE oraz MAPE (tabela 1).

Tabela 1. Błędy prognozy dla poszczególnych modeli neuronowych

	X ₁	X ₂	X ₃	X ₄	X ₅	X ₆
MSE (10 ⁻⁴)	8,5616	8,3066	6,9451	8,2068	6,2407	8,8686
MAPE (10 ⁻²)	2,1454	2,1408	1,9121	2,1108	1,8243	2,1855

Źródło: opracowanie własne

W kolejnym kroku dokonano dekompozycji prognoz na komponenty bazowe $y_1 \dots y_6$ za pomocą algorytmu AMUSE zgodnie z równaniem (3). Następnie podjęto próbę identyfikacji komponentów konstruktywnych i destruktywnych przy użyciu badania odległości euklidesowej pomiędzy wektorami wartości prognoz poszczególnych komponentów. Wyniki zebrane w tabeli 2 nie pozwalają jednak na jednoznaczną kategoryzację komponentów. Można jedynie zauważyć, że od pozostałych komponentów wyraźnie odróżnia się komponent y_1 .

Tabela 2. Odległości euklidesowe między wektorami komponentów bazowych

	y ₁	y ₂	y ₃	y ₄	y ₅	y ₆
y ₁	0,00	25 693,78	25 700,56	25 579,62	25 656,51	25 679,86
y ₂	25 693,78	0,00	3 639,29	2 989,31	3 335,86	3 730,19
y ₃	25 700,56	3 639,29	0,00	3 609,59	3 577,76	3 919,25
y ₄	25 579,62	2 989,31	3 609,59	0,00	3 276,68	3 415,62
y ₅	25 656,51	3 335,86	3 577,76	3 276,68	0,00	3 675,32
y ₆	25 679,86	3 730,19	3 919,25	3 415,62	3 675,32	0,00

Źródło: opracowanie własne

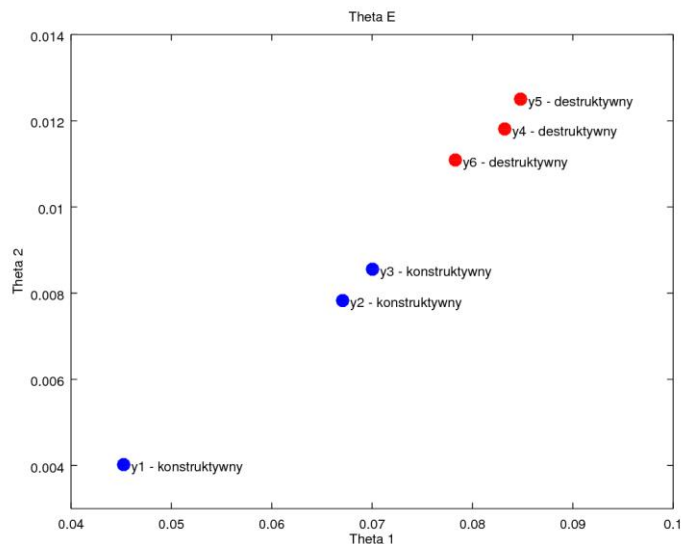
W następnym etapie eksperymentu dokonano obliczeń wariancji, współczynników Hursta oraz miar zmienności $\Theta_1(y)$, $\Theta_2(y)$, $\Theta_3(y)$ oraz $\Theta_E(y)$ zaproponowanych w niniejszej pracy. Wyniki przedstawiono w tabeli 3.

Tabela 3. Wariancja, współczynnik Hursta oraz miary zmienności $\Theta(y)$

	y_1	y_2	y_3	y_4	y_5	y_6
$D^2(y)$	0,0134	1,3331	1,4180	1,4616	1,4896	1,7895
Hurst	0,8700	0,6400	0,6300	0,6300	0,7000	0,5300
$\Theta_1(y)$	0,0452	0,0670	0,0700	0,0832	0,0848	0,0783
$\Theta_2(y)$	0,0040	0,0078	0,0086	0,0118	0,0125	0,0111
$\Theta_3(y)$	0,0042	0,0293	0,0332	0,0410	0,0430	0,0442
$\Theta_E(y)$	0,0454	0,0675	0,0706	0,0840	0,0857	0,0791

Źródło: opracowanie własne

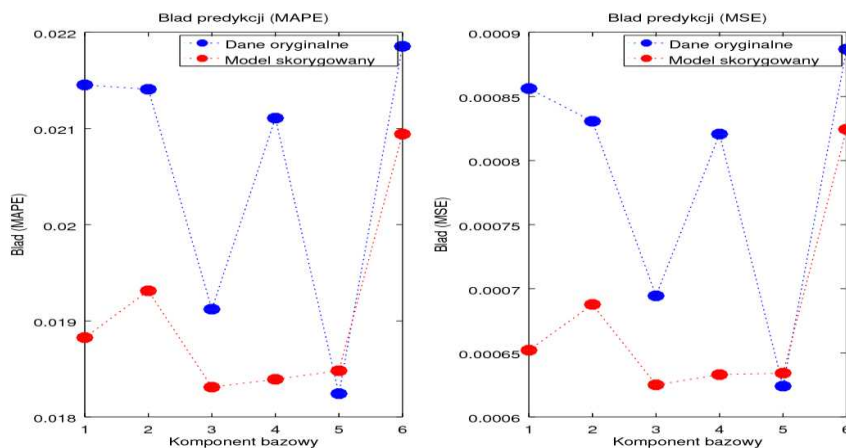
Jak wynika z tabeli 3 pierwszy z komponentów y_1 z uwagi na niską wariancję oraz wysoki współczynnik Hursta należy najprawdopodobniej do komponentów konstruktywnych. Z drugiej jednak strony, wartości wariancji oraz współczynnika Hursta nie pozwalają na odróżnienie komponentów destrukcyjnych i konstrukcyjnych spośród pozostałych pięciu komponentów $y_2 \dots y_6$. Można natomiast analizując wartości miar zmienności dojść do wniosku, że destrukcyjne w tym przypadku są ostatnie trzy komponenty y_4, y_5, y_6 , gdyż w ich przypadku wartości miary $\Theta_E(y)$ są istotnie wyższe niż w przypadku pierwszych trzech komponentów (powyżej 0,79) co pokazuje rysunek 1, w którym miara jest widoczna jako odległość poszczególnych punktów od punktu (0,0) na wykresie.

Rysunek 1. Miary zmienności $\Theta_1(y)$, $\Theta_2(y)$ oraz $\Theta_E(y)$ 

Źródło: opracowanie własne

Na podstawie powyższej obserwacji dokonano eliminacji 3 komponentów bazowych y_4 , y_5 , y_6 w wyniku, którego średni błąd predykcji zużycia energii elektrycznej zmniejszył się średnio o 7,8% jeśli zmierzmy go za pomocą metody MSE oraz 3,8% w przypadku MAPE. Rysunek 2 ilustruje szczegółowe zmiany błędów predykcji dla poszczególnych komponentów bazowych.

Rysunek 2. Błędy predykcji przed i po eliminacji komponentów destrukcyjnych



Źródło: opracowanie własne

PODSUMOWANIE

Przedstawiona koncepcja identyfikacji komponentów destrukcyjnych opiera się na nowych charakterystykach zmienności pozwalających dokonywać stosunkowo precyzyjnych pomiarów wpływu poszczególnych komponentów na wyniki prognoz. Zagadnienie identyfikacji zakłóceń/szumów przedstawione zostało w ramach predykcyjnego podejścia wielomodelowego, jednak prezentowana metoda może zostać zastosowana w dowolnym problemie, w którym problemy detekcji losowości sygnałów lub oceny zakłóceń wiążemy z charakterystykami zmienności lub gładkości. W ramach badanego przykładu opracowana metoda pozwala na efektywną identyfikację komponentów destrukcyjnych dla wyników prognozy. Prezentowane metody otwierają duże możliwości dalszych badań teoretycznych oraz aplikacyjnych w tym obszarze.

BIBLIOGRAFIA

- Breiman L. (1996) Bagging Predictors. *Machine Learning*, 24, 123-140.
Clements R. T. (1989) Combining Forecasts: A Review and Annotated Bibliography. *International Journal of Forecasting*, 5, 559-581.

- Comon P., Jutten C. (2010) Handbook of Blind Source Separation: Independent Component Analysis and Applications. Academic Press.
- Gatnar E. (2008) Podejście wielomodelowe w zagadnieniach dyskryminacji i regresji. PWN.
- Hamilton J. D. (1994) Time Series Analysis. Princeton University Press, Princeton.
- Hurst H. E. (1951) Long Term Storage Capacity of Reservoirs. Transactions of the American Society of Civil Engineers, 116, 770-799.
- Hyvarinen A., Karhunen J., Oja E. (2001) Independent Component Analysis. John Wiley, New York.
- Oja E. (1995) PCA, ICA and Nonlinear Hebbian Learning. Proc. Int. Conf. on Artificial Neural Networks ICANN-95, Oct. 9-13, France, 89-94.
- Shiryaev A. N. (1999) Essentials of Stochastic Finance: Facts, Models, Theory. Singapore, World Scientific.
- Szupiluk R., Wojewnik P., Ząbkowski T. (2004) Model Improvement by the Statistical Decomposition. Lecture Notes in Computer Science, 3070, 1199-1204.
- Szupiluk R. (2013) Dekompozycje wielowymiarowe w agregacji predykcyjnych modeli data mining. Oficyna Wydawnicza Szkoły Głównej Handlowej w Warszawie.
- Szupiluk R., Wojewnik P., Ząbkowski T. (2010) Noise Detection for Ensemble Methods. Lecture Notes in Artificial Intelligence, 6113, 471-478.
- Therrien C. W. (1992) Discrete Random Signals and Statistical Signal Processing. Prentice Hall, New Jersey.
- Vasegi S. V. (1997) Advanced Signal Processing and Digital Noise Reduction. John Wiley and Sons, Chichester.

IDENTIFICATION OF DESTRUCTIVE COMPONENTS IN PREDICTIVE MODELS WITH A MULTI-MODEL APPROACH

Abstract: In this paper we present a method of identification of destructive components in predictive models. This method may be applied in case of a multi-model approach and uses algorithms of blind signal separation. The evaluation of the characteristics of individual components will be based on the proposed metrics for evaluating the variation or smoothness of signals. In order to confirm the effectiveness of the presented method, we will present a practical experiment in which the results of the forecast of short-term electricity consumption are improved.

Keywords: prediction, blind separation, ensemble methods, Theta noise measure