

Centralna Komisja do Spraw
Stopni i Tytułów

Wniosek
z dnia 15 grudnia 2011
o przeprowadzenie postępowania habilitacyjnego

w dziedzinie nauki techniczne w dyscyplinie elektrotechnika

1. Imię i Nazwisko:

Krzysztof Siwek

2. Stopień doktora:

doktor inżynier nauk technicznych w dyscyplinie elektrotechnika, Wydział
Elektryczny, Politechnika Warszawska, 2001

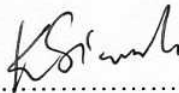
3. Wskazanie jednostki organizacyjnej do przeprowadzenia postępowania
habilitacyjnego

Wydział Elektryczny, Politechnika Warszawska

4. ~~Wnoszę o głosowanie komisji postępowania habilitacyjnego w trybie tajnym¹⁾.~~

5. Przyjmuję do wiadomości, iż wniosek wraz z autoreferatem zostanie opublikowany na
stronie internetowej Centralnej Komisji do Spraw Stopni i Tytułów, zgodnie z
obowiązującymi przepisami.

1) jeżeli niepotrzebne - skreślić



.....
podpis Wnioskodawcy

Załączniki (w 2 egz.):

1. kopia dyplomu nadającego stopień doktora,
2. autoreferat w języku polskim,
3. autoreferat w języku angielskim,
4. wykaz opublikowanych prac naukowych w języku polskim,
5. wykaz opublikowanych prac naukowych w języku angielskim,
6. oświadczenia współautorów prac zbiorowych wchodzących w skład
jednotematycznego cyklu publikacji, określające indywidualny wkład w jej powstanie,
7. informacja o osiągnięciach naukowych, współpracy naukowej i działalności
popularyzującej naukę,
8. informacja o osiągnięciach dydaktycznych,
9. dane kontaktowe do autora wniosku
10. kopie jednotematycznego cyklu 16 publikacji,
11. płyta CD z elektronicznymi wersjami wszystkich załączonych dokumentów.

16. GRU. 20

Wpłynęło dnia.....
Nr DOK VI-L-6020/14

Warszawa, 2.01.2012

dr inż. Krzysztof Siwek
Wydział Elektryczny
Politechnika Warszawska

Centralna Komisja do Spraw
Stopni i Tytułów

W sprawie wniosku z dnia 15 grudnia 2011
o przeprowadzenie postępowania habilitacyjnego

Proszę o dołączenie do dokumentacji wnioskowej następujących dokumentów:

1. Pisma od prof. Andrzeja Cichockiego, głównego autora i pomysłodawcy pakietów ICALAB, określającego mój wkład w powstanie tych pakietów.
2. Opublikowanego artykułu (pozycja 1 z jednotematycznego cyklu publikacji z załącznika 2 do wniosku):

1. KRZYSZTOF SIWEK, Stanisław Osowski, "Improving the accuracy of prediction of PM10 pollution by the wavelet transformation and an ensemble of neural predictors", **International Scientific Journal Engineering Applications of Artificial Intelligence**, Vol. 46, Nos. 3/4, 2011, pp. 199-215

.....
podpis Wnioskodawcy

Wpłynęło dnia 03.01.12
Nr DOK VI-L-6121/11

Autoreferat

1. Imię i Nazwisko: **Krzysztof Siwek**

2. Posiadane dyplomy, stopnie naukowe – z podaniem nazwy, miejsca i roku ich uzyskania.

- magister inżynier, Wydział Elektryczny, Politechnika Warszawska, 1995,
- doktor inżynier nauk technicznych w dyscyplinie elektrotechnika, Wydział Elektryczny, Politechnika Warszawska, 2001, rozprawa doktorska pod tytułem: „**Prognostowanie obciążeń w systemie elektroenergetycznym przy wykorzystaniu sztucznych sieci neuronowych**”.

3. Informacje o dotychczasowym zatrudnieniu w jednostkach naukowych/ artystycznych.

- 1995-obecnie – pracownik naukowo dydaktyczny, aktualnie na stanowisku adiunkta, Instytut Elektrotechniki Teoretycznej i Systemów Informacyjno-Pomiarowych, Wydział Elektryczny, Politechnika Warszawska

4. Wskazanie osiągnięcia* wynikającego z art. 16 ust. 2 ustawy z dnia 14 marca 2003 r. o stopniach naukowych i tytule naukowym oraz o stopniach i tytule w zakresie sztuki (Dz. U. nr 65, poz. 595 ze zm.):

jednotematyczny cykl publikacji pod tytułem:

Zastosowania metod sztucznej inteligencji w zadaniach predykcji i klasyfikacji

a) (autor/autorzy, tytuł/tytuły publikacji, nazwa wydawnictwa, rok wydania),

1. KRZYSZTOF SIWEK, Stanisław Osowski, “Improving the accuracy of prediction of PM10 pollution by the wavelet transformation and an ensemble of neural predictors”, **International Scientific Journal Engineering Applications of Artificial Intelligence**, zaakceptowany do publikacji w październiku 2011, druk w 2012, dostępne w doi:10.1016/j.engappai.2011.10.013, (2011).
2. KRZYSZTOF SIWEK, Stanisław Osowski, Mieczysław Sowiński, “Evolving the ensemble of predictors model for forecasting the daily average”, **International Journal of Environment and Pollution (IJEP)**, zaakceptowany do publikacji w marcu 2011, druk w 2012.
3. Stanisław Osowski, KRZYSZTOF SIWEK, Robert Siroic, “Neural system for heartbeats recognition using genetically integrated ensemble of classifiers”, **Computers in Biology and Medicine**, 2011, vol. 41, pp. 173-180.

Dodatkowo w kolumnie 4 zamieściłem wyniki działania zespołu 4 sieci neuronowych (MLP, RBF, Elmana, maszyny SVM), w którym każda sieć działała na rozkładzie falkowym danych.

Tabela 1. Wyniki prognozy wartości PM10 dla danych testujących, nie biorących udziału w uczeniu (definicje miar jakości w [1])

	Prognoza przy użyciu sieci SVM	Prognoza przy użyciu SVM i dekompozycji falkowej	Integracja dodatkową siecią SVM
1	2	3	4
MAE [$\mu\text{g}/\text{m}^3$]	9.45	4.62	4.46
RMSE [$\mu\text{g}/\text{m}^3$]	18.03	6.73	6.41
MAPE [%]	28.21	14.62	14.43
R	0.64	0.90	0.92
IA	0.71	0.91	0.92

• Integracja zespołu wielu predyktorów indywidualnych w jeden system finalny

Kolejną zrealizowaną przeze mnie koncepcją jest połączenie wielu systemów predykcyjnych w jeden zintegrowany zespół, podobnie jak w przypadku prognozowania obciążeń elektroenergetycznych.

Każdy z systemów predykcyjnych oblicza niezależną prognozę, obarczoną zwykle pewnym błędem. Rozkład statystyczny tego błędu jest różny dla każdego predyktora. System integrujący ma za zadanie takie połączenie prognoz, które zapewni najmniejszy błąd finalnej predykcji. Opracowałem i przetestowałem różne systemy integracji wyników działania poszczególnych elementów zespołu, w tym: uśrednianie, uśrednianie ważone, zastosowanie dekompozycji PCA, zastosowanie dekompozycji wielowymiarowej przy użyciu algorytmów ślepej separacji sygnałów oraz zastosowanie integratora neuronowego. Wyniki takiej integracji pokazały przewagę zespołu predyktorów nad najlepszym predyktorem indywidualnym (przykład wyników integracji zawarty jest w kolumnie 4 tabeli 1). Wyniki badań zostały zaprezentowane w pracach [1,2,4].

W wyniku współpracy z Zakładem Interdyscyplinarnych Zastosowań Fizyki Narodowego Centrum Badań Jądrowych (NCBJ Świerk, osoba kontaktowa: doc. dr Mieczysław Sowiński) samodzielnie zaprojektowałem i wykonałem aplikację komputerową do predykcji dobowego zanieczyszczenia powietrza działającą w trybie on-line. Aplikacja ta używana jest przez Instytut do monitoringu jakości powietrza atmosferycznego.

Rezultaty badań prowadzonych przeze mnie od 2007 w tematyce prognozowania zanieczyszczenia powietrza atmosferycznego pyłem zawieszonym PM10 zostały opublikowane w czasopismach [1,2,4] oraz zaprezentowane na krajowych i międzynarodowych konferencjach naukowych.

3. Zastosowanie metod sztucznej inteligencji w zadaniach klasyfikacji

Drugim nurtem moich badań było zastosowanie praktyczne metod sztucznej inteligencji w problemach klasyfikacji. Zaliczyć można do nich zastosowanie tych metod w inżynierii (diagnostyka obwodów elektrycznych i maszyn) i bioinżynierii (wsparcie diagnostyki zaburzeń rytmu serca, rozpoznawania komórek krwiotwórczych w diagnostyce białaczki oraz selekcji genów w chorobach nowotworowych).

3.1. Zastosowanie metod sztucznej inteligencji w zagadnieniach technicznych

Moje badania w tym obszarze koncentrowały się na zastosowaniu sieci neuronowych w rozpoznawaniu i lokalizacji uszkodzeń w obwodach elektrycznych na podstawie zewnętrznych pomiarów prądów i napięć, a także diagnostyce stanu uzwojeń w 3-fazowych silnikach indukcyjnych na podstawie pomiarów przebiegów czasowych prądów i napięć stojana. Główny nacisk w tych pracach położyłem na metody opracowania cech diagnostycznych i wybór ich optymalnego zestawu.

Jedną z nich jest zastosowanie transformaty PCA do graficznej ilustracji możliwości separacyjnych różnych klas przy różnym zestawie cech diagnostycznych. Rozwiązanie to w zastosowaniu do diagnostyki obwodu elektrycznego na podstawie automatycznego pomiaru zewnętrznych prądów i napięć w obwodzie zostało opublikowane w [5,13].

Kolejnym zastosowaniem metod sztucznej inteligencji w technice było rozwiązanie problemu wykrywania liczby i pozycji uszkodzonych prętów wirnika w 3-fazowych maszynach indukcyjnych na podstawie wyłącznie sygnału prądu stojana. Do generacji cech diagnostycznych zastosowałem transformatę Fouriera prądu stojana tworzącą widmo częstotliwościowe a następnie metodę eliminacji w tył i przód dla selekcji najbardziej znaczących cech diagnostycznych. Klasyfikatorem była maszyna wektorów podtrzymujących SVM z jądrem gaussowskim. Praktycznym rezultatem badań było opracowanie automatycznego systemu lokalizacji uszkodzeń silników 3-fazowych w ramach grantu Ministerstwa Nauki i Szkolnictwa, którego byłem współwykonawcą. Teoretyczne i praktyczne rezultaty tego projektu zostały przedstawione w pracy [5].

3.2. Zastosowanie metod sztucznej inteligencji w zagadnieniach biomedycznych

Kolejnym kierunkiem moich badań w dziedzinie praktycznych zastosowań metod sztucznej inteligencji były problemy klasyfikacyjne w medycynie. Badania dotyczyły trzech różnych zadań. Pierwszym z nich było zastosowanie algorytmu genetycznego i SVM do rozpoznawania komórek krwotwórczych w szpiku kostnym dla wspomagania diagnostyki białaczki [9]. Drugim był rozwój narzędzi wspomagających diagnozowanie zaburzeń rytmu serca [3,15], a ostatnim zastosowanie metod klasyfikacji w selekcji genów na podstawie mikromacierzy ekspresji genów DNA [10].

W pracy [9] przedstawiłem zastosowanie algorytmu genetycznego (GA) i sieci SVM do klasyfikacji komórek na podstawie zdjęć mikroskopowych. Głównym zadaniem algorytmu genetycznego był optymalny wybór cech diagnostycznych dla sieci SVM. Zastosowanie GA poprawiło dokładność rozpoznawania komórek krwi o ponad 25% (w ujęciu względnym) w odniesieniu do najlepszej klasycznej metody wyboru funkcji (liniowej sieci SVM). Dodatkowo wykazałem znaczną przewagę sieci SVM nad MLP w zadaniach klasyfikacji.

W wyniku prac badawczych powstał system wspomagania diagnozy medycznej charakteryzujący się wysoką skutecznością, porównywalną ze skutecznością ludzkiego eksperta, przewyższający go znacznie szybkością działania.

Praca [3] zawiera opis i wyniki badań problemu rozpoznawania różnych typów arytmii serca. W tym rozwiązaniu algorytm genetyczny służy do obliczania wag, z jakimi bierze się wyniki klasyfikacji indywidualnych klasyfikatorów działających w zespole. Zaproponowane rozwiązanie było przetestowane na danych z bazy arytmii MIT BiH. Pozwoliło ono na obniżenie liczby błędnych klasyfikacji o ponad 20%. Ponadto współczynnik fałszywych alarmów (najbardziej niebezpiecznych błędów) został znacznie zmniejszony.

W pracy [10] została zaprezentowana nowatorska metoda selekcji genów powiązanych z danym typem choroby nowotworowej. Selekcja dotyczyła ponad 20000 genów lub sekwencji genowych uzyskanych z mikromacierzy ekspresji. W badaniach przeprowadziłem analizę różnych technik selekcji genów, w tym korelacji, hipotezy statystycznej, klasteryzacji oraz zastosowanie liniowych maszyn SVM. Transformacja PCA posłużyła do udowodnienia poprawności wyboru cech, poprzez rzutowanie ich na dwie współrzędne odpowiadające dwóm najważniejszym elementom głównym PCA. Jako finalny klasyfikator użyta była sieć SVM z jądrem gaussowskim. W pracy przedstawiono szczegółowe testy klasyfikacji siedmiu typów nowotworów. Skuteczność działania przedstawionego rozwiązania jest bardzo wysoka i może być wykorzystania przy wspomaganiu diagnostyki nowotworowej.

Zaproponowane przeze mnie metody klasyfikacji i algorytmy wykorzystujące narzędzia sztucznej inteligencji są uniwersalne i mogą być stosowane w wielu różnych aplikacjach, zarówno w zakresie inżynierii i inżynierii biomedycznej.

4. Wnioski końcowe

Badania przeprowadzone przeze mnie koncentrują się na zastosowaniu metod sztucznej inteligencji, sieci neuronowych i algorytmów zwłaszcza genetycznych do różnych zadań predykcji i klasyfikacji, w szczególności praktycznych aspektów ich zastosowania w inżynierii i medycynie. Cechą wspólną łączącą wszystkie zaprezentowane tu rozwiązania jest zastosowanie podobnych narzędzi i metod sztucznej inteligencji.

Do moich najważniejszych osiągnięć należy zaliczyć opracowanie i dogłębne przetestowanie różnych metody generacji i selekcji cech diagnostycznych. Cechy te, w zależności od rozpatrywanego zagadnienia, mają różne źródła – pochodzą z bezpośrednich pomiarów pewnych zmiennych (na przykład parametrów atmosferycznych w problemach prognozy zanieczyszczeń lub pomiarów prądów i napięć w obwodzie elektrycznym) lub są przetworzoną wersją wyników pomiarów (np. przy zastosowaniu transformaty falkowej Fouriera lub liniowej transformacji PCA danych). Opracowane przeze mnie metody selekcji optymalnego zestawu cech diagnostycznych są uniwersalne i mogą być zastosowane w różnych zadaniach praktycznych.

Innym moim osiągnięciem jest opracowanie nowych koncepcji łączenia predyktorów lub klasyfikatorów w zespół. Dzięki opracowanym przeze mnie metodom możliwe jest zastosowanie gorszych jakościowo rozwiązań dla poprawy jakości działania całego systemu. Do metod tych należą obliczanie średnich ważonych, zastosowanie transformacji PCA, algorytmów genetycznych, ślepej separacji sygnałów, jak również stosowanie dodatkowych sieci neuronowych funkcjonujących jako integratory. Nowe metody optymalnej integracji wyników działania poszczególnych podzespołów tworzących zespół zostały opracowane i przetestowane zarówno w zadaniach prognozy jak i klasyfikacji.

Wszystkie przedstawione przeze mnie zagadnienia związane są z problemami praktycznymi. Eksperymenty numeryczne zostały przeprowadzone na podstawie rzeczywistych danych z dziedziny inżynierii (dane zużycia energii w systemach energetycznych, rejestracje prądów i napięć obwodów elektrycznych oraz silników 3-fazowych, pomiary zanieczyszczenia powietrza dostępne na stronach internetowych Wojewódzkich Inspektoratów Ochrony Środowiska) lub medycyny (baza danych arytmii MIT BIH, dane z mikromacierzy ekspresji genów dostępne w sieci, własne fotografie mikroskopowe komórek krwi przy diagnostyce białaczki). We wszystkich tych dziedzinach wprowadziłem kilka nowych pomysłów i rozwiązań, które zostały zaakceptowane przez recenzentów w wielu międzynarodowych czasopismach z bazy JCR.

Należy również podkreślić, że jestem znaczącym współautorem pierwszej wersji dwóch toolboxów ICALAB dla programu Matlab, zaliczających się do najpopularniejszych pakietów komputerowej realizacji algorytmów ślepej separacji sygnałów. Pakiet ICALAB for Image Processing jest do chwili obecnej jedyną liczącą się w świecie implementacją ślepej separacji sygnałów w zastosowaniu do obrazów, również kolorowych.

3. Omówienie pozostałych osiągnięć naukowo - badawczych.

3.1 Opracowanie pakietu ICALAB

W trakcie pobytu w latach 1998-2002 w na stażach naukowych w Instytucie Badań Fizycznych i Chemicznych RIKEN w Japonii we współpracy z zespołem naukowców z Laboratory for Advanced Brain Signal Processing, a przede wszystkim z profesorami Andrzejem Cichockim i Shun-ichi Amari, zaprojektowałem i opracowałem w języku Matlab pierwszą wersję pakietu dwóch toolboxów „ICALAB” implementującą różne metody i algorytmy związane ze ślepą separacją sygnałów (BSS, BSE, ICA, SCA, MCA, NTC, NMF). Opracowany pakiet znalazł uznanie u naukowców zajmujących się tą problematyką o czym świadczy wysoka liczba cytowań (83 cytowania według Google Scholar) pozycji:

Andrzej Cichocki, Shun-ichi Amari, KRZYSZTOF SIWEK, Toshihisa Tanaka, Anh Huy Phan et al., ICALAB Toolboxes, <http://www.bsp.brain.riken.jp/ICALAB>

Pakiet ICALAB stanowią dwa niezależne toolboxy do Matlab (ICALAB for Signal Processing oraz ICALAB for Image Processing) implementujące:

- algorytmy analizy komponentów niezależnych ICA (ang. Independent Component Analysis) oparte na statystykach wyższych rzędów HOS (ang. Higher Order Statistics),
- algorytmy ślepej separacji sygnałów BSS (ang. Blind Source Separation) opartych na statystykach drugiego rzędu SOS (second order statistics) i predykcji liniowej LP (ang. Linear Prediction),
- algorytmy ślepej ekstrakcji sygnałów BSE (ang. Blind Signal Extraction) implementujących różne metody SOS i HOS.

Opracowane programy posiadają rozbudowane komponenty pre- i postprocesingu, interfejsy graficzne GUI umożliwiające łatwą pracę z sygnałami oraz obrazami, okna graficzne do prezentacji obrazów i sygnałów w postaci wykresów czasowych, transformat częstotliwościowych oraz spektrogramów. Pakiet ICALAB for Image Processing jest do chwili obecnej jedyną liczącą się w świecie implementacją ślepej separacji sygnałów w zastosowaniu do obrazów, również kolorowych.

Udział w opracowaniu pakietu: 50%

Udział w opracowaniu algorytmów: 5%

Zakres: opracowanie koncepcji pakietów (toolboxów), programowanie w Matlabie wszystkich komponentów pakietów, wkład w implementację poszczególnych algorytmów separacji, modyfikacja algorytmów separacji do współpracy z toolboxami, liczne poprawki w poszczególnych algorytmach, większość oprogramowania GUI.

3.2 Opracowanie oprogramowania komputerowego do predykcji zanieczyszczenia PM10

W wyniku współpracy z Zakładem Interdyscyplinarnych Zastosowań Fizyki Narodowego Centrum Badań Jądrowych (NCBJ Świerk, osoba kontaktowa: doc. dr Mieczysław Sowiński) samodzielnie zaprojektowałem i wykonałem aplikację komputerową do predykcji średniego dobowego zanieczyszczenia powietrza pyłem zawieszonym PM10 na jeden i dwa dni naprzód działającą w trybie on-line. Aplikacja ta jest używana przez Instytut do monitoringu jakości powietrza atmosferycznego w Warszawie.

3.3 Opracowanie modeli komputerowych sztucznych sieci neuronowych

W wyniku prowadzonych badań nad sieciami neuronowymi samodzielnie zaprojektowałem i zaimplementowałem wiele programów komputerowych do uczenia i testowania sztucznych sieci neuronowych, w tym sieci MLP, sieci Elmana, Hopfielda, Hamminga, BAM oraz sieci samoorganizujące Kohonena. Programy te używane są obecnie w pracy naukowej przez pracowników i studentów oraz w dydaktyce w Laboratorium Sztucznych Sieci Neuronowych na Wydziale Elektrycznym Politechniki Warszawskiej.

3.4 Opracowanie metod i programów komputerowych w dziedzinie Data Mining

Zajmuję się ponadto problematyką Data Mining, a przede wszystkim zagadnieniami odkrywania cech diagnostycznych i prognostycznych oraz wizualizacji danych wielowymiarowych. Moje prace badawcze dotyczą problemów automatycznego określania wymiarowości dla dekompozycji PCA i LDA. Zaimplementowałem te metody w postaci programów komputerowych.

Aktualnie prowadzę prace badawcze w problematyce rozpoznawania twarzy przy wykorzystaniu PCA i LDA. Uzyskane do tej pory wyniki pozwalają mieć nadzieję na dobre działanie końcowego systemu.

W pracy badawczej miałem często styczność z problemem niekompletnych danych pomiarowych. W polu moich zainteresowań leżą także zagadnienia automatycznego uzupełniania takich danych. Opracowałem wiele metod uzupełniania brakujących danych, w tym również bazujących na predykcji.

Interesuję się ponadto zagadnieniami predykcji szeregów czasowych w zagadnieniach ekonomicznych i finansowych. W monografii [4] (patrz Załącznik 5 punkt A) przedstawiłem rozwiązanie prognozowania wskaźnika inflacji przy użyciu sieci neuronowych.



4. KRZYSZTOF SIWEK, Stanisław Osowski, Mieczysław Sowiński, "Neural predictors and wavelet transformation for forecasting the PM10 pollution", **COMPEL**, 2011, vol. 30, No 4, pp. 1376-1388.
5. Stanisław Osowski, Jarosław Kurek, KRZYSZTOF SIWEK, "Computerized system for fault diagnosis of the rotor bars of squirrel-cage induction motor", **Problemy Eksploatacji**, 2010, vol. 79, nr 4, pp. 135-151.
6. KRZYSZTOF SIWEK, Stanisław Osowski, Bartosz Świdorski, "Trend elimination of time series of 24-hour load demand in the power system and its application in power forecasting", **Przegląd Elektrotechniczny**, 2011, vol. 87, No 3, pp. 249-253.
7. KRZYSZTOF SIWEK, "Graphical analysis of load pattern in Polish Power System using PCA and Kohonen map", **Przegląd Elektrotechniczny**, 2010, vol. 86, No 1, pp. 249-253.
8. KRZYSZTOF SIWEK, Stanisław Osowski, Ryszard Szupiluk, "Ensemble neural network approach for accurate load forecasting in the power system", **Applied Mathematics and Computer Science**, 2009, vol. 19, No 2, pp. 303-315.
9. Stanisław Osowski, Robert Siroic, Tomasz Markiewicz, KRZYSZTOF SIWEK, "Application of Support Vector Machine and Genetic Algorithm for improved blood cell recognition", **IEEE Transactions On Instrumentation And Measurement**, 2009, Vol. 58, Issue 7, pp. 2159-2168.
10. Artur Wilinski, Stanisław Osowski, KRZYSZTOF SIWEK, "Gene selection for cancer classification through ensemble of methods", **Lecture Notes in Computer Science**, 2009, LNCS 5495, pp. 507-516.
11. KRZYSZTOF SIWEK, Stanisław Osowski, "Electrical power forecasting for small energetic railroad region using SVM", **Przegląd Elektrotechniczny**, 2007, R 83, vol. 11, pp. 6-9.
12. Andrzej Cichocki, Stanisław Osowski, KRZYSZTOF SIWEK, "Comparisons of prewhitening algorithms of noisy signals", **Przegląd Elektrotechniczny**, No 2, R. 81, 2005, pp. 22-25
13. Robert Sałat, Stanisław Osowski, KRZYSZTOF SIWEK, "Principal Component Analysis for feature selection at the diagnosis of electrical circuits", **Przegląd Elektrotechniczny**, 2003, No 10, pp. 667-670
14. Stanisław Osowski, KRZYSZTOF SIWEK, "Regularization of neural networks for improved load forecasting in the power system", **IEE Proc. Generation, Transmission and Distribution**, 2002, vol. 149, No 3, pp. 340 -344
15. Stanisław Osowski, Linh Tran Hoai, KRZYSZTOF SIWEK, "Integracja klasyfikatorów neuronowych w zastosowaniu do rozpoznawania rytmów serca", **Przegląd Elektrotechniczny**, 2002, vol. 78, No 12, pp. 347-350

16. Andrzej Cichocki, KRZYSZTOF SIWEK, "Blind sources separation for noisy signals", **Przegląd Elektrotechniczny**, No 12, 2002, R. 78, pp. 344-347

b) omówienie celu naukowego/artystycznego ww. pracy/prac i osiągniętych wyników wraz z omówieniem ich ewentualnego wykorzystania.

1. Wstęp

Tematyka moich prac badawczych poczynając od rozprawy doktorskiej z roku 2001 dotyczy rozwoju metod sztucznej inteligencji w zastosowaniach praktycznych, głównie w zadaniach predykcji szeregów czasowych i klasyfikacji. W szczególności zadania te skupiają się na problematyce mającej bezpośredni związek z zastosowaniami inżynierskimi oraz biomedycznymi. Należą do nich:

- prognozowanie energii elektrycznej w skali kraju oraz małego rejonu energetycznego w szczególności kolejowego,
- prognozowanie zanieczyszczeń powietrza atmosferycznego,
- rozwój metod automatycznej diagnostyki obwodów i maszyn elektrycznych,
- opracowanie ulepszonych metod klasyfikacji wzorców biomedycznych, w tym obrazów mikrofotografii komórek krwiotwórczych szpiku kostnego,
- opracowanie metod selekcji genów ściśle powiązanych z określoną chorobą nowotworową poprzez zastosowanie podejścia klasyfikacyjnego do mikromacierzy ekspresji genów,
- opracowanie skutecznych metod klasyfikacji rytmów serca do wspomagania diagnostyki medycznej.

Uzyskane przeze mnie osiągnięcia opublikowane zostały w czasopismach z bazy JCR i obejmują szeroki zakres problemów występujących w tematyce predykcji i klasyfikacji. Główny wysiłek badawczy skupiony został na następujących zagadnieniach badawczych.

- Badanie charakteru procesu podlegającego predykcji, w tym liniowości, chaotyczności, współzależności między zmiennymi. Mają one znaczący wpływ na model predykcyjny.
- Zdefiniowanie, a następnie automatyczna selekcja optymalnych cech diagnostycznych opisujących badany proces. Cechy te w dużej mierze decydują o skuteczności rozwiązania zarówno predykcyjnego jak i klasyfikacyjnego.
- Zaprojektowanie i implementacja komputerowa systemu predykcyjnego i klasyfikatora opartego na metodach sztucznej inteligencji. W praktyce wykorzystałem takie narzędzia inteligentne jak sieci neuronowe, maszyny wektorów podtrzymujących SVM (Support Vector Machine, ze względu na możliwy sposób przedstawienia w postaci ekwiwalentnej do dwuwarstwowego perceptronu z jednym wyjściem, zwane również sieciami SVM), algorytmy genetyczne, zaawansowane metody i narzędzia statystyczne.
- Opracowanie specjalizowanych metod zwiększania dokładności działania rozwiązań predykcyjnych i klasyfikacyjnych.
- Analiza porównawcza opracowanych rozwiązań w ważnych praktycznie zadaniach predykcji i klasyfikacji.

Szczególnie duży nacisk w badaniach położyłem na opracowanie autorskich metod zwiększających dokładność działania modelu predykcyjnego i klasyfikacyjnego. Do osiągnięcia tych celów stosowałem zaawansowane techniki obliczeniowe, między innymi:

- regularyzację sieci neuronowych prowadzącą do optymalnego uproszczenia struktury i zwiększenia w ten sposób zdolności generalizacyjnych,
- zastosowanie transformacji falkowej do dekompozycji sygnałów pozwalającej podzielić optymalnie jedno trudne zadanie na szereg zadań o zmniejszonej skali trudności,
- dekompozycję sygnałów przy użyciu ślepej separacji sygnałów (BSS) pozwalającej oddzielić sygnały użyteczne od szumu i skupić się w zadaniu wyłącznie na sygnałach użytecznych,
- zastosowanie równoległe wielu predyktorów lub klasyfikatorów połączonych w zespole, pozwalające spojrzeć na rozwiązanie zadania z wielu punktów widzenia,
- integrację wielu systemów predykcyjnych lub klasyfikacyjnych w jeden system finalny, umożliwiającą wykorzystanie najlepszych cech każdego predyktora lub klasyfikatora do podwyższenia efektywności działania systemu.

Opublikowane wyniki układają się w jednotematyczny cykl 14 artykułów opublikowanych w czasopismach znajdujących się w bazie JCR i 2 w czasopismach spoza tej listy skupiający się na praktycznych zastosowaniach predykcyjnych i klasyfikacyjnych metod sztucznej inteligencji.

- Artykuły [6,7,8,11,14] poruszają problematykę predykcji obciążeń elektroenergetycznych w krajowym systemie elektroenergetycznym oraz w małym rejonie energetycznym PKP. Prace te bazują na danych rzeczywistych Krajowego Systemu Elektroenergetycznego oraz PKP Energetyka S.A.
- Publikacje [1,2,4] związane są z tematyką prognozowania zanieczyszczeń pyłowych (PM10) powietrza atmosferycznego na podstawie rzeczywistych pomiarów wielkości parametrów atmosferycznych wykonanych w Warszawie przy współpracy z NCBJ Świerk.
- Prace [3,5,9,10,13,15] dotyczą problematyki klasyfikacji, w szczególności metod eksploracji danych użytecznych w diagnostyce maszyn i obwodów elektrycznych [13], jak również w zagadnieniach biomedycznych, w tym w diagnostyce chorób serca [2,15] i chorób nowotworowych [8].
- Artykuły [1,2,3,4,8,14,15] zawierają rozwiązania metod zwiększających dokładność predykcji i klasyfikacji poprzez użycie zespołów klasyfikatorów lub predyktorów zintegrowanych w jeden system finalny przy zastosowaniu specjalizowanych nowatorskich metod opracowanych przeze mnie.
- Publikacje [12,16] dotyczą zagadnień teoretycznych ślepej separacji sygnałów, które znalazły zastosowanie w metodach integracji zespołu klasyfikatorów i predyktorów przy użyciu techniki BSS.

Uzyskane przeze mnie osiągnięcia badawcze są powiązane z aktywną współpracą krajową i międzynarodową. W latach 1998-2002 spędziłem łącznie 15 miesięcy na stażach naukowych w Instytucie Badań Fizycznych i Chemicznych RIKEN w Japonii. W tym czasie we współpracy z profesorami Andrzejem Cichockim i Shun-ichi Amari zaprojektowałem i

opracowałem w języku Matlab pierwszą wersję pakietu dwóch toolboxów ICALAB: ICALAB for Signal Processing oraz ICALAB for Image Processing, implementującą różne metody i algorytmy związane ze ślepą separacją sygnałów (BSS, BSE, ICA, SCA, MCA, NTC, NMF). Opracowany pakiet znalazł uznanie u naukowców zajmujących się tą problematyką o czym świadczy wysoka liczba cytowań tego pakietu: 83 cytowania według Google Scholar.

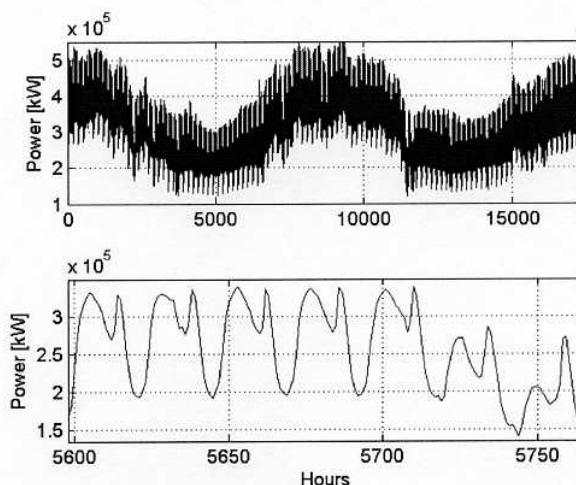
2. Zastosowanie metod sztucznej inteligencji w zadaniach predykcji

Tematyka predykcji rozwijana w pracach [6,7,8,11,14] związana była z predykcją szeregów czasowych i dotyczyła prognozowania zapotrzebowania na energię elektryczną na 24 godziny naprzód (predykcja szeregu 24 elementowego) oraz prognozowania wartości średniej dobowej zanieczyszczeń atmosferycznych PM10 na następny dzień [1,2,4].

2.1. Predykcja zapotrzebowania na energię elektryczną

Problematyka prognozowania 24-godzinnego na energię elektryczną stanowi ważny czynnik efektywności gospodarowania energią elektryczną, gdyż umożliwia precyzyjne planowanie w tej dziedzinie gospodarki zmniejszając koszty zarówno produkcji, zakupu jak i dystrybucji energii elektrycznej. Jakkolwiek problematyka ta, ze względu na jej znaczenie praktyczne jest rozwijana na świecie od wielu lat, nadal pozostaje wiele zagadnień praktycznych do rozwiązania, w szczególności dokładność prognozy. Dotychczasowe metody prognozowania, poczynając od metod liniowych, takich jak algorytmy autoregresyjne (ARX, ARMAX, ARIMA), poprzez różnego rodzaju dekompozycje (Fouriera, SVD, PCA, falkowa), klasyczne metody optymalizacyjne, w tym optymalizacji globalnej (algorytmy genetyczne, ewolucyjne, symulowanego wyżarzania) aż po najczęściej dziś stosowane sztuczne sieci neuronowe stanowią dobrze znane narzędzia, ale nie do końca wykorzystane dla polepszenia dokładności prognozy, mającej kluczowe znaczenie w ekonomice działania systemu elektroenergetycznego.

Należy podkreślić, że prognozowanie obciążeń 24-godzinnych należy do zadań trudnych, gdyż obserwuje się dużą wariancję w przebiegu godzinnych obciążeń (rys. 1). Jednocześnie przy dużej zmienności obciążeń wymaga się jak najbardziej precyzyjnej prognozy dla każdej godziny doby.



Rys. 1. Przebieg obciążenia godzinowego PSE: a) dla 2 lat, b) dla wybranego tygodnia

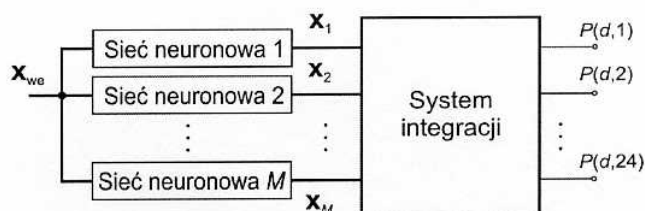
Dotychczas stosowane podejścia zakładały stosowanie różnych rozwiązań, wybierając następnie spośród nich tę, która zapewnia najlepsze działanie systemu prognostycznego na danych weryfikujących, nie biorących udziału w uczeniu. W prezentowanych rozwiązaniach zaproponowałem nowe podejścia prowadzące do zwiększenia dokładności prognozy i polegające na wykorzystaniu nawet słabszych rozwiązań, łącząc je z lepszymi w jeden zespół.

Już w swojej pracy doktorskiej zaprojektowałem i zaimplementowałem metody predykcji obciążeń energetycznych oparte na pojedynczych sieciach neuronowych MLP i samoorganizujących się przez konkurencję. Następne moje prace (po obronionej w 2001 rozprawie doktorskiej) poświęcone były opracowaniu nowych ulepszonych metod predykcji przy zastosowaniu metod sztucznej inteligencji oraz nowych podejść bazujących na zastosowaniu wielu układów prognostycznych działających jednocześnie w zespole. Warunkiem prawidłowego działania zespołu jest zastosowanie podzespołów działających niezależnie od siebie i właściwa integracja wyników pojedynczych predyktorów. Dla spełnienia tego warunku zastosowałem różne rozwiązania sieci neuronowych różniących się zarówno strukturą jak i algorytmami uczącymi. W szczególności w skład zespołu wchodzi:

- sieć wielowarstwowa jednokierunkowa MLP stosująca sigmoidalną funkcję aktywacji, będąca aproksymatorem typu globalnego,
- sieć radialna RBF o gaussowskiej funkcji aktywacji stosująca aproksymację lokalną,
- sieć (maszyna) wektorów podtrzymujących SVM charakteryzująca się specyficznym algorytmem uczenia różnym od MLP i RBF,
- sieć Kohonena tworzona przy użyciu algorytmu samoorganizacji poprzez konkurencję,
- sieć rekurencyjna Elmana charakteryzująca się strukturą ze sprzężeniami zwrotnymi.

W pracach [6,7,8,11,14] zostały przedstawione szczegóły działania zaproponowanych przeze mnie neuronowych metod prognostycznych wykorzystujących wymienione sieci. Kolejnym etapem prac było zastosowanie zaawansowanych metod prognostycznych łączących wiele pojedynczych predyktorów w jeden zespół prognostyczny.

Każdy system prognozy w postaci bądź to odpowiedniej sieci neuronowej bądź rozwiązania liniowego bazującego na statystyce, działa na tych samych danych wejściowych dostępnych w bazie danych z przeszłości. Wyniki prognozy poszczególnych systemów prognostycznych są integrowane w jednym wspólnym zespole sieci. W ten sposób w rozwiązaniu bierze udział wiele systemów prognostycznych działających równolegle. Takie rozwiązanie pozwala wykorzystać dobre cechy każdego predyktora, dla uzyskania prognozy ostatecznej. Rysunek 2 przedstawia ogólny schemat zespołu predyktorów do prognozowania wzorca obciążeń 24 godzinnych na następny dzień.



Rys. 2. Ogólny schemat zespołu predyktorów

W swoich pracach opracowałem różne podejścia do integracji wyników działania wielu predyktorów. Jedno z nich wykorzystuje ślepą separację sygnałów lub transformację PCA i stosuje liniowe podejście do integracji. Inne rozwiązanie jako integrator wykorzystuje dodatkową nieliniową sieć neuronową. Bardziej zaawansowane podejścia do integracji łączą w sobie kilka metod zarówno liniowych jak i nieliniowych.

Nowatorstwo rozwiązań zespołu prognostycznego zaproponowane przeze mnie można ująć w następujących punktach:

- optymalna redukcja wymiarowości danych poprzez zastosowanie transformaty PCA, prowadząca do zwiększenia zdolności generalizacyjnych systemu,
- zastosowanie transformaty falkowej do dekompozycji na sygnały bardziej gładkie, łatwiej poddające się prognozie,
- zastosowanie ślepej separacji sygnałów do dekompozycji wielowymiarowej danych i usunięcie komponentów szumowych,
- połączenie metod liniowych ARX i nieliniowych (sieci neuronowych) w jednym systemie,
- integracja wielu predyktorów, tworzących zespół, metodami liniowymi (zastosowanie średniej, średniej ważonej, transformacji PCA),
- integracja nieliniowa przy użyciu dodatkowej sieci neuronowej,
- jednoczesne zastosowanie na raz kilku z wymienionych powyżej metod.

Wyniki badań prowadzonych od 2001 roku zostały wszechstronnie przetestowane przeze mnie na danych rzeczywistych dotyczących Polskiego Systemu Elektroenergetycznego i małego rejonu energetycznego PKP, opublikowane w czasopismach oraz zaprezentowane na wielu krajowych i międzynarodowych konferencjach naukowych (Załącznik 4 punkt C).

2.2. Predykcja zanieczyszczenia atmosferycznego pyłem zawieszonym

Zanieczyszczenia powietrza atmosferycznego są ważnym problemem praktycznym z dziedziny ochrony środowiska jak również ochrony zdrowia. Problem zanieczyszczeń jest szczególnie istotny w dużych skupiskach ludności oraz w rejonach silnie uprzemysłowionych. Zanieczyszczeniem mającym największy niekorzystny wpływ na środowisko i zdrowie ludzkie jest pył mechaniczny zawieszony w powietrzu. Głównymi źródłami tego pyłu są: ruch pojazdów kołowych (sadza pochodząca z silników, pył unoszony przez kontakt opony z nawierzchnią) oraz efekty spalania głównie w procesach ogrzewania. Aktualnie rozpatruje się dwie frakcje pyłu zawieszonego: PM₁₀ – pył o średnicy do 10 μm oraz PM_{2,5} – pył o średnicy do 2,5 μm. Aktualnie dostępne są przede wszystkim bazy pomiarowe pyłu PM₁₀ z wielu lat. Bazy pomiarowe pyłu PM_{2,5} są tworzone dopiero od 2010 roku.

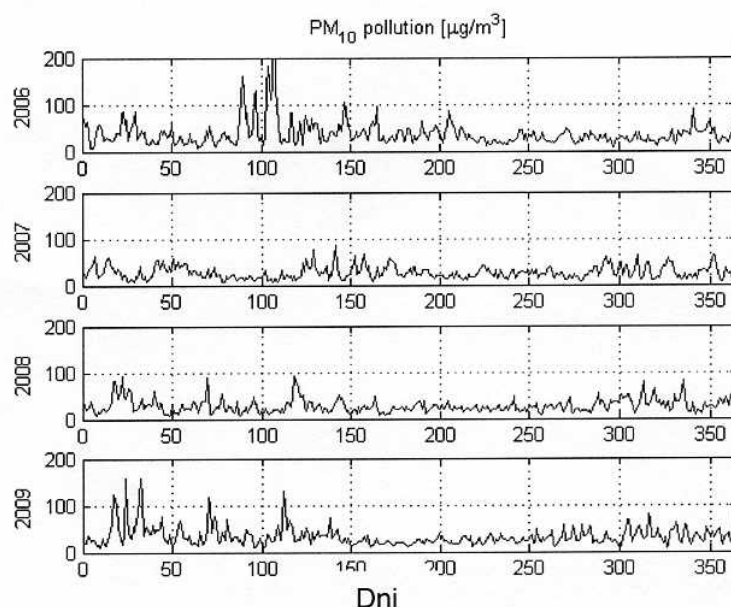
Regulacje unijne (Dyrektywa European Air Quality Directive EC/2008/50) nakładają na Polskę obowiązek oceny trendów i utrzymania dobrej jakości powietrza oraz zapewnienie, że informacje, na temat jakości powietrza, pomocne w walce z zanieczyszczeniami będą udostępnione społeczeństwu. Dyrektywa reguluje również dopuszczalne zawartości pyłu PM₁₀ w powietrzu i nakłada na państwa członkowskie obowiązek wprowadzenia kar dla samorządów, na terenie których poziomy zanieczyszczeń przekraczają wartości dopuszczalne.

Badania dotyczące prognozowania zawartości pyłu PM₁₀ w powietrzu pozwalają na stosunkowo łatwe i tanie oszacowanie poziomów zapylenia na dzień następny i podjęcie działań zmierzających do ich ograniczenia. Stąd moje zainteresowanie tą tematyką.

Rys. 3 przedstawia wykres dobowej zmienności zawartości pyłu PM10 w powietrzu atmosferycznym na terenie Warszawy w latach 2006-2009. W odróżnieniu od omawianej w poprzednim punkcie problematyki predykcji obciążeń elektroenergetycznych przebieg ten nie posiada widocznych regularności. Zadanie predykcji w tym przypadku jest więc wielokrotnie trudniejsze. Według dostępnych badań czynnikami zewnętrznymi wpływającymi na poziom zanieczyszczenia powietrza są warunki pogodowe, takie jak temperatura i wilgotność powietrza, ciśnienie atmosferyczne, siła i kierunek wiatru, nasłonecznienie itp.

W trakcie swoich prac badawczych przeprowadziłem szczegółową analizę danych pomiarowych w celu określenia charakteru modelowanego procesu. Wykluczyły one liniowość i potwierdziły nieliniowy (chaotyczny) charakter procesu o dużej zmienności obserwowanego przebiegu zanieczyszczenia.

W wyniku dogłębnej analizy danych pomiarowych zaproponowałem i szczegółowo przebadalem metody sztucznej inteligencji służące do wydobywania z nich informacji, w tym selekcji cech najbardziej wpływających na dokładność prognozy [1,2,4].



Rys. 3. Wartości średnie dobowe zanieczyszczenia powietrza atmosferycznego pyłem mechanicznym zawieszonym PM10 w latach 2006-2009 (stacja pomiarowa Warszawa-Ursynów)

Kolejnym rozwiązaniem przeze mnie zagadnieniem było opracowanie różnorodnych modeli realizujących predykcję wartości średniej zawartości pyłu PM10 z wyprzedzeniem jednego dnia. Do opracowanych i przebadanych rozwiązań zaliczają się modele liniowe ARX oraz sztuczne sieci neuronowe: MLP, RBF, maszyny SVM, sieci samoorganizujące Kohonena oraz sieci rekurencyjne Elmana [1].

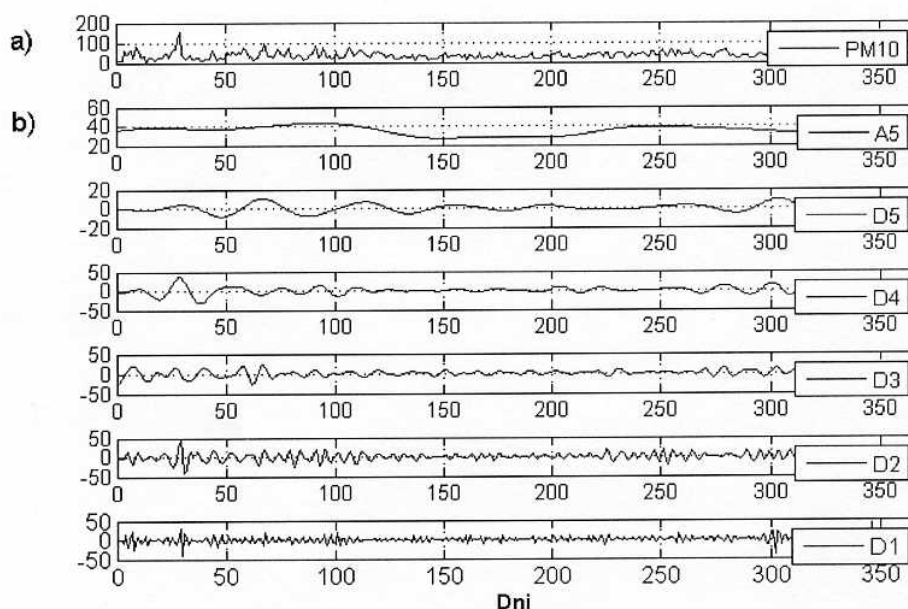
Duży nacisk w swoich badaniach położyłem na metody zwiększania dokładności działania opracowanych modeli prognostycznych. Jakkolwiek uzyskane wyniki predykcji pojedynczych modeli są na poziomie najlepszych wyników prezentowanych w międzynarodowych czasopismach, to zaproponowałem i przetestowałem nowe metody zwiększające jeszcze bardziej dokładność prognozy.

Do głównych elementów zwiększających dokładność prognozy zanieczyszczeń atmosferycznych według przeprowadzonych przeze mnie badań należy zaliczyć: dekompozycję szeregu czasowego przy użyciu transformaty falkowej, zastosowanie wielu predyktorów oraz ich właściwą integrację.

- **Dekompozycja sygnału czasowego przy użyciu transformaty falkowej**

Zaproponowane przeze mnie rozwiązanie polega na rozłożeniu przebiegu czasowego przy użyciu transformaty falkowej na przebiegi o mniejszej zmienności z kroku na krok i przeprowadzeniu oddzielnych predykcji dla sygnałów o mniejszej zmienności, które zsumowane razem tworzą prognozę końcową. Dzięki takiemu rozwiązaniu uzyskano znaczne zmniejszenie błędu końcowego prognozy.

Rysunek 4 przedstawia przykładowy rozkład szeregu czasowego PM10 przy użyciu falki db8 na 5 poziomów szczegółowości. Widać, że wyodrębnione zostały: aproksymacja A5 oraz przebiegi D1-D4 odpowiadające współczynnikom falkowym, które charakteryzują się zredukowaną zmiennością.



Rys. 4. Przykładowy rozkład falkowy przebiegu dobowego PM10 (a) na 5 poziomów szczegółowości (b) przy użyciu falki Daubechies Db4

Po przeprowadzeniu predykcji poszczególnych składowych falek wynik końcowy predykcji oblicza się jako sumę indywidualnych prognoz dla sygnałów A5 i D1-D5. W rezultacie otrzymuje się prognozę na następny dzień obciążoną znacznie mniejszym błędem niż w przypadku predykcji sygnału nierozłożonego [1,4].

W tabeli 1 przedstawiłem przykładowe zestawienie uzyskanych wyników prognozy w postaci błędów MAE, RMSE, MAPE, współczynnika korelacji R oraz indeksu zgodności IA [1] dla pojedynczego (najlepszego) rozwiązania z użyciem sieci SVM (działającej w trybie regresji) bez rozkładu falkowego (kolumna 2) i z rozkładem falkowym danych (kolumna 3).