



Instytut Badań Systemowych  
Polskiej Akademii Nauk

---

*Seria:*

**BADANIA SYSTEMOWE**

**TOM 65**

Redaktor Naukowy

Prof. dr hab. inż. Jakub Gutenbaum

## **Rada Redakcyjna:**

1. Prof. dr hab. inż. Olgierd Hryniewicz – *przewodniczący*
2. Prof. dr hab. inż. Jakub Gutenbaum – *redaktor naukowy*
3. Prof. dr hab. inż. Janusz Kacprzyk
4. Prof. dr hab. inż. Tadeusz Kaczorek
5. Prof. dr hab. inż. Roman Kulikowski
6. Doc. dr hab. inż. Marek Libura
7. Prof. dr hab. inż. Krzysztof Malinowski
8. Prof. dr hab. inż. Zbigniew Nahorski
9. Dr. hab. inż. Marek Niezgódka, prof. UW
10. Prof. dr hab. inż. Roman Słowiński
11. Doc. dr hab. inż. Jan Studziński
12. Prof. dr hab. inż. Stanisław Walukiewicz
13. Prof. dr hab. inż. Andrzej Weryński
14. Doc. dr hab. inż. Antoni Żochowski



Instytut Badań Systemowych  
Polskiej Akademii Nauk

**Antoni Wiliński**

**GMDH – metoda grupowania argumentów  
w zadaniach zautomatyzowanej predykcji  
zachowań rynków finansowych**

**Warszawa - Szczecin**

Copyright © by Instytut Badań Systemowych PAN  
Warszawa 2009

*Dr hab. inż. Antoni Wiliński – prof. nadzw.  
Zachodniopomorski Uniwersytet Technologiczny w Szczecinie  
Wydział Informatyki  
Ul. Żołnierska 49, Szczecin 71-210  
Tel. 091- 449 5660 ; fax 091-449 5540*

*Autor jest profesorem Zachodniopomorskiego Uniwersytetu Technologicznego, poprzednio Politechniki Szczecińskiej. Habilitację uzyskał na Politechnice Warszawskiej w zakresie automatyki i robotyki. W obecnej kadencji jest dziekanem Wydziału Informatyki ZUT. Od lat zajmuje się problematyką inteligencji obliczeniowej i rozpoznawania wzorców w szeregach czasowych w celach predykcyjnych lub modelowania rzeczywistości. W Katedrze Systemów Multimedialnych, której jest kierownikiem, prowadzone są badania nad webowym systemem handlu automatycznego. Rozważane są rozmaite strategie analizy technicznej specyficzne dla decyzji podejmowanych z wysoką częstotliwością przez infoboty. Wśród tych strategii istotną rolę odgrywają omawiane tu algorytmy oparte na GMDH (Group Method of Data Handling). Praca przeznaczona jest raczej dla specjalistów będących zwolennikami analizy technicznej.*

### **Recenzenci:**

Prof. dr hab. inż. Zbigniew Banaszak  
Doc. dr hab. inż. Maciej Krawczak

### **Redaktor techniczny:**

Dr inż. Anna Samborska-Owczarek

Powyższej książki w całości lub części nie wolno powielać ani przekazywać w żaden sposób, nawet za pomocą nośników mechanicznych i elektronicznych (np. zapis magnetyczny), w tym też umieszczać ani rozpowszechniać w postaci cyfrowej zarówno w Internecie, jak i w sieciach lokalnych, bez uzyskania pisemnej zgody posiadacza praw autorskich.

ISBN 9788389475237  
**ISSN 0208-8029**



*To, co znane, nie jest jeszcze przez to, że jest znane, czymś poznanym*  
Fryderyk Hegel (1770-1831)

## Wprowadzenie

Zasadniczym celem niniejszej monografii jest potwierdzenie tezy autora, że fuzja matematyki i technologii egzemplifikowana powiązaniem metod predykcji ze strategią inwestycyjną opartą na mechanizmach technologicznych platform brokerskich daje nową synergetyczną wartość. Wartości tej nie daje się uzyskać odrębnie. Przy niczym nieograniczonej fascynacji handlem<sup>1</sup> na platformach internetowych [10, 66, 81, 85, 86, 94], często dziś spotykanej wśród licznych, najczęściej młodych przedstawicieli przyszłego społeczeństwa informacyjnego, nie da się, bez odpowiedniej wiedzy, osiągnąć sukcesu. Podobnie sama wyłącznie teoretyczna wiedza, bez umiejętności jej zastosowania, pozostaje metafizyczną wartością, rzadko satysfakcjonująca twórcę. Trawestując Freddie Mercurę<sup>2</sup> można powiedzieć *The sale must go on* i coś energetycznego jest w tym aforyzmie. Jest w nim imperatyw wdrożenia odkrycia matematycznego, rozwiązania znalezionego w przestrzeni przeszłej. To zdarzenie rzadkie w świecie nauki. Z kolei szaleństwo nieprzygotowanego żywiołowego uczestnictwa w giełdowych<sup>3</sup> i foreksowych<sup>4</sup> transakcjach to zdarzenia częstotliwe, niejednokrotnie z odcieniem chorobliwego hazardu i najczęściej bez happy-endu.

---

<sup>1</sup> Przez handel na platformach internetowych rozumie się najczęściej częstotliwe zawieranie transakcji na odległość polegające na otwieraniu pozycji na rynkach udostępnianych przez daną platformę (najczęściej na rynkach walutowych). Takie zachowanie inwestora nazywane jest często, bez tłumaczenia, tradingiem.

<sup>2</sup> Freddie Mercury – słynny brytyjski piosenkarz ub. wieku, którego jednym z największych przebojów był utwór „The show must go on”.

<sup>3</sup> W pracy rozpatrywane będą predykcje walorów (akcji) giełdowych i kursów walut.

<sup>4</sup> Forex – to międzynarodowy rozproszony rynek walutowy udostępniany właśnie poprzez tysiące internetowych platform brokerskich.

Czego więc syntezą jest ostatecznie sukces na rynku<sup>5</sup> finansowym? Jest zapewne sumą wiedzy (zastosowań matematyki, statystyki, udanych modeli predykcyjnych, znajomości realiów ekonomicznych), umiejętności (znajomości technologii brokerskich, przewidywania reakcji rynków na podstawie symptomów politycznych, ekonomicznych czy społecznych, stosowania oceny percepcyjnej zdarzeń) i możliwości ciągłego uczestnictwa w grze rynkowej (inaczej mówiąc wolnego czasu po podjęciu decyzji inwestycyjnej lub użycia systemu informatycznego do monitoringu rynku i podejmowania automatycznych decyzji wg wcześniej przetestowanej strategii).

W tym aspekcie cytowane dalej motto Kanta o znaczeniu matematyki w obiektywizacji reguł uogólniających opis rzeczywistości, tworzących model wiedzy o niej, jest wyrazem przekonań autora. Matematyki nie należy jednak w tym cytacie zawęzać do jej akademickiej definicji. Dziś Kantowi chodziłoby zapewne także o szersze rozumienie tego konceptu – uwzględniającego np. także dokonania metod sztucznej inteligencji, teorii chaosu czy mechaniki kwantowej. Szerzej, zgadzając się z Włodzisławem Duchem [15], należy raczej posługiwać się tu pojęciem inteligencji obliczeniowej, która jest pojęciem szerszym niż sztuczna inteligencja. Metoda staje się bowiem wartościowa poprzez obliczenia. Przed ich przeprowadzeniem często nie ma podstaw do uznania koncepcji za metodę.

Nie odzegnując się od niczego, autor, odwołując się do klasyfikacji postaw w nauce wg Andrzeja Wierzbickiego [124], skłonny jest zaliczyć prezentowane tu poglądy do skrajnie pragmatycznych, typowych dla przedstawicieli bardziej nauk technicznych niż nauk ścisłych.

Dzisiejsze postawy inwestorów wobec zadania predykcji na rynkach finansowych w najbardziej generalny sposób różnicuje podejście zaproponowane przez Eugena Famę [20], który wprowadzając kilkadziesiąt lat temu pojęcie rynku efektywnego podzielił inwestorów na zwolenników analizy fundamentalnej, analizy technicznej, zwolenników równoczesnego wykorzystywania metod z obu obszarów wiedzy oraz sceptyków przekonanych o nieprzewidywalności absolutnie niestacjonarnych zachowań rynków. Współcześnie, uważa się (np. Hia Jong Teoh [33], Goefrey Friesen[24], Bill Cai [8]), że klasyfikacja przekonań o możliwości przewidywania rynków winna przebiegać jeszcze nieco inaczej. Powinna przebiegać mianowicie wg podziału na trzy kategorie postaw -

---

<sup>5</sup> Rynek to umowna nazwa dziedziny, w której obserwowane są zmiany ceny wybranego instrumentu finansowego np. kursu waluty, ceny akcji na giełdzie, ceny towaru itp.

że nic nie da się przewidzieć, że tylko analiza fundamentalna jest zasadna oraz, że tylko wyrafinowane metody współczesnej analizy technicznej (wykorzystujące sztuczne sieci neuronowe, rozpoznawania wzorców, algorytmy genetyczne itp.) mają sens.

Niewątpliwie do zwolenników tego ostatniego podejścia zaliczyć można większość noblistów w dziedzinie ekonomii (tych, związanych z ekonometrią, ekonomią i inżynierią finansową) ostatnich kilkunastu lat takich jak Myron Scholes, Robert Merton, Robert Engle, Harry Markowitz, Joseph Stiglitz, William Sharpe. Wszyscy oni uważali, że jest sens tworzenia złożonych modeli odwzorowania gospodarczej rzeczywistości i pomimo świadomie popełnianych błędów (na poziomie uproszczeń w założeniach koniecznych do zbudowania modelu) przyczynili się do istotnego poszerzenia wiedzy matematycznej o modelowaniu rynków.

Częste są także opinie, że prawdopodobnie najlepsze wyniki predykcyjne na poziomie zapewniającym ich praktyczną użyteczność inwestycyjną uzyskuje się poprzez kojarzenie elementów analizy technicznej i fundamentalnej. To oczywiście pewna odmiana przytoczonego wyżej poglądu o celowości poszukiwania skutecznego rozwiązania w gąszczu współczesnej inteligencji obliczeniowej. Metametodą jest tu hybrydyzacja metod o różnym charakterze i różnych technik wykonania decyzji rekomendowanych przez te modele. Bardzo dynamicznie rozwijają się ostatnio w rozpatrywanym obszarze badania chińskich analityków np. Sheng-Hsun Hsu [107] i Zhang Yudong [133] łącząc modelowanie za pomocą sztucznych sieci neuronowych z teorią zbiorów przybliżonych [88, 89] uważają, że to najbardziej perspektywiczny kierunek opanowywania niestacjonarnych procesów, a Hia Jong Teoh [33] przedstawia metodę łączącą sieci neuronowe z GMDH (Group Method of Data Handling). S. Dehuri i S. Cho [14] łączą techniki regresyjne z optymalizacją rojem cząstek. Optymalizacja rojem jest najczęściej zabiegiem pomocniczym dla określania najlepszych parametrów w wielowymiarowej przestrzeni predyktorów (indykatorów, zmiennych objaśniających, składowych portfela itp.).

Interesującym aspektem metod stosowanych dla predykcji rynków jest także umowny podział na dwie duże grupy metod (niezależnie od wykorzystywania danych o naturze fundamentalnej lub technicznej). Do pierwszej grupy można byłoby zaliczyć wszelkie metody oparte na idei pozyskiwania pewnych parametrów z przeszłości pozwalających na zbudowanie modelu wykorzystującego te parametry (czasami tylko chwilowo) i umożliwiających niejako kontynuację modelu w przyszłości. Metod tych nie należy jednak zaliczać do parametrycznych wg taksonomii stosowanej w statystyce – nie są to metody oparte na założeniu zna-



jomości rozkładu statystycznego zmiennej losowej. Do tej grupy z pewnością należy rozpatrywana tu metoda GMDH i wszelkie inne metody oparte na regresji, korelacji, średnich kroczących i ich pochodnych, modelach rodziny GARCH, algorytmach genetycznych czy sieciach neuronowych.

Druga duża rodzina to metody oparte na poszukiwaniu wzorców. Podstawowa idea przyświecająca zwolennikom tej grupy metod jest hipoteza, że jeżeli zmiany w szeregu czasowym wielokrotnie obserwowane w przeszłości następowały wg określonego powtarzalnego schematu, to prawdopodobne jest powtórzenie tego schematu także wówczas, gdy zaczyna być obserwowany jego początek. Charakter wzorców może być bardzo zróżnicowany. Wzorcami są powszechnie stosowane w analizie percepcyjnej rynków, wynikające z teorii fal Elliotta formacje graficzne (takie jak np. formacja głowy i ramion, podwójnego szczytu, spodka, diamentu itp.). Wzorcami mogą być określone sekwencje świec japońskich, może to być także zbiór jednocześnie spełnionych wielu warunków związanych z indykatorami czy parametrami rozkładu statystycznego.

Większość uznanych traderów<sup>6</sup> takich jak Larry Williams, Jack Schwager, Joe Kruttsinger czy Thomas DeMark, opisani w pracy Joe Kruttsingera [66], posługiwała się i posługuje technikami, które dziś można nazwać poszukiwaniem wzorców. Większość z opisywanych technik (spośród tych ujawnianych) to stosunkowo proste zabiegi mające jednak wspólną cechę. Są rezultatem ciężkiej wieloletniej pracy, tysięcy obserwacji i zauważenia trudnych do logicznego objaśnienia, jednak powtarzalnych sekwencji zdarzeń. Przez stosunkowo proste wzorce należy rozumieć takie, które daje się zaobserwować lub obliczyć w czasie rzeczywistym bez budowania złożonych modeli prognostycznych, zawsze jednak, jak twierdzi większość spośród tych wspomnianych zwycięzców rynku, podstawowym narzędziem wspomagającym ich decyzje jest ustawicznie pracujący komputer.

Interesującym podejściem jest kojarzenie analizy technicznej z fundamentalną. Do tej grupy można na przykład zaliczyć techniki prognostyczne oparte na obserwacji zdarzeń gospodarczych np. ogłaszanych o stałych porach dnia czy miesiąca informacjach o nastrojach lub

---

<sup>6</sup> Trader w powyższym kontekście to zawodowo trudniący się prognozowaniem rynków inwestor, często mentor dla innych graczy. Wielcy amerykańscy traderzy to ludzie sukcesu, którzy wbrew rozpowszechnionym poglądom o nieprzewidywalności rynków dowiedli, że stosowane przez nich techniki, rzadko ujawniane, są skuteczne.

wskaźnikach rozwoju gospodarczego, stopach procentowych ustalanych przez banki centralne itp. Posługiwanie się tymi technikami w sposób automatyczny umożliwia serwis [fxengines.com](http://fxengines.com) [142]. Do tej grupy metod można także zaliczyć opisywaną w niniejszej pracy fuzję GMDH z wynikami obserwacji inwestorów uczestniczących w handlu na wybranym rynku.

W Polsce problematyką GMDH bezpośrednio zainteresowało się niewielkie grono naukowców i powstało niewiele publikacji, z których najważniejsze to prace Witolda Pedrycza [2, 90] i Józefa Korbicza [62, 68]. Spośród innych autorów zainteresowanych metodą należy wymienić nazwiska Zbigniewa Banaszaka, Marcina Mrugalskiego, Janusza Kusia. Powstała także jedna (prawdopodobnie) praca doktorska wykorzystująca wprost metodę GMDH napisana przez Justynę Patalas pod kierunkiem Tadeusza Krupy.

Tytułowa metoda grupowania argumentów (atrybutów) to opracowane przez ukraińskiego uczonego Aleksieja Ivakhnenkę [34-51] podejście do interpolacji i predykcji w formie samoorganizującego się opisu rzeczywistości (modelu). W podejściu tym kojarzona jest wspomniana wiedza i umiejętności. Wiedza to zdolność do opracowania najlepszego w danych okolicznościach modelu (np. predykcyjnego), umiejętność to zdolność do zastosowania sugestii wynikających z tego modelu do podjęcia decyzji (np. inwestycyjnej, z wykorzystaniem platformy brokerskiej). Najstarsze z tych prac powstały już w latach siedemdziesiątych ub. wieku [37, 39, 47, 50].

Aleksiej Ivakhnenko [36, 38] i jego następcy Tetsuo Sawaragi [101], Hema Madala [73], Johann Mueller i Frank Lemke [40, 41, 46, 70] rozwinęli kilka algorytmów mieszczących się w obrębie tej metody.

Metoda w części dotyczącej modelu predykcyjnego polega na równoczesnym zastosowaniu dwóch zasad – zbudowania najlepszego regresyjnego modelu według kryterium opartego na informacji nieużywanej do budowy modelu, oraz drugiej zasady – ograniczenia złożoności struktury modelu regresyjnego do najmniejszej akceptowalnej przez badacza (zgodnie z zasadą brzytwy Ockhama). Istotą metody, od której pochodzi jej nazwa – grupowanie argumentów lub atrybutów, jest uwzględnianie w modelu regresyjnym tylko kilku spośród wielu argumentów – zmiennych wejściowych/zmiennych objaśniających. Oznacza to kolejne rozpatrywanie niejednokrotnie tysięcy modeli regresyjnych o różnych konfiguracjach tworzących je danych wejściowych i wyborze tylko najlepszych z nich. W tym sensie metoda jest samoorganizująca się. Na każdym kroku szeregu czasowego najlepszy w danej chwili model regresyjny ma nie tylko różne współczynniki liczbowe przy poszcze-

gólnych wyrazach równania (w ogólnym przypadku nieliniowego), ale także różną strukturę tego równania w sensie wartości stopnia wykładników wyrazów równania – modelu. Wybór kilku (trafniej, wielu) najlepszych, a nie jednego najlepszego modelu, na danym kroku złożoności algorytmu, Ivaknenko uzasadnia koniecznością pozostawienia pewnej swobody w procesie decyzyjnym. Nawiązując do prac Gabora Ivakhnenko [42] powtarza argumentację selekjonerów np. roślin czy zwierząt, którzy pozostawiają zawsze kilka-kilkanaście najlepszych par czy osobników. W ten sposób uzyskuje się niejako szerszą bazę odwzorowania rzeczywistości. Nie wszystkie najlepsze cechy mogą być przekazane przez pojedynczy egzemplarz (przez najlepszy model rzeczywistości). Ta chwilowa dominacja może być przypadkowa i w wielokrotnie powtarzanej predykcji może się nie sprawdzić.

Z punktu widzenia stopnia naszej determinacji w procesie poznania rzeczywistości metoda ta jest zapewne specyficznym konsensem pomiędzy filozoficznym ewolucjonizmem, wyrażonym równaniem regresji, a kreacjonizmem, wyrażonym poprzez odrzucenie przekonania o skuteczności prognozowania na podstawie przeszłości. Autorski kreacjonizm jest tu wyrażony poprzez zwątpienie w możliwość precyzyjnej prognozy. Sceptycyzm ten wyrażany jest wielokrotnie – myślą przewodnią tej pracy jest przecież odrzucenie wiary w gaussowski rozkład ceny walorów względem średniej (np. rozdział 3) i ciągłe stosowanie rozbudowanej przestrzeni zmiennych decyzyjnych związanych z technologiami platform brokerskich. Sceptycyzm ten wyraża się także w rozpatrywaniu wyłącznie prognoz krótko- i bardzo krótkoterminowych, całkowicie wynikających z matematycznej czy statystycznej wiedzy pozyskanej z przeszłości, ignorującej „chiromancje” ekonomiczne czy polityczne. Postawa taka rzutuje zresztą na przyszłe zastosowania metody – jako automatycznie realizowanego scalpingu<sup>7</sup>.

Metoda z pewnością należy do grupy wielu współczesnych metod opartych na myśleniu indukcyjnym takich jak sieci neuronowe, RBF (*radial basis functions*), algorytmy ewolucyjne, teoria zbiorów przybliżonych. W aspekcie rozwiązywanych zadań predykcyjnych i decyzyjnych metoda łączy podejścia charakterystyczne dla ekstrakcji wiedzy i data mining'u (eksploracji danych) z rozpoznawaniem wzorców i klasyfikacją. Ta część metody z pewnością nawiązuje do postawy przeświadczenia o możliwości ekstrakcji wiedzy z bazy danych. Ta postawa

---

<sup>7</sup> Scalping – to specyficzna dla foreksowych platform brokerskich technika inwestowania polegająca na częstotliwym otwieraniu i zamykaniu pozycji, co kilka-kilkanaście minut, a nawet częściej.

występuje wyraźnie u protoplasty metody – prof. Ivaknenki, który pisał [39, 49], cytując Gabora, że świat realny jest bardzo złożony, bardziej, niż to się powszechnie uważa, nie oznacza to jednak, że należy zrezygnować z możliwości rozpoznania tej złożoności. Mało tego, bardzo często złożony model jest niespodziewanie precyzyjny i nie przeuczony<sup>8</sup>.

Zastosowanie indukcji w poszukiwaniu reguły rządzącej rzeczywistością, czyli wyprowadzenie nieznanego do tej pory prawa na podstawie obserwacji, zaczęło być równie ważne jak dedukcyjny sposób myślenia dopiero w epoce informacyjnej i towarzyszącej jej burzliwemu wzrostowi możliwości obliczeniowych maszyn. Pomimo tej technologicznej rewolucji autor konstatuje ciągle duży dyskomfort w zderzeniu potrzeb obliczeniowych rozpatrywanej tytułowej metody GMDH (Group Method of Data Handling)<sup>9</sup> i możliwości aktualnych komputerów. Dodatkową obliczeniową komplikacją podejścia proponowanego przez Aleksieja Ivakhnenkę i naśladowców – Hema Madalę [73], Johanna Mullera [70], jest rozszerzenie regresyjno-neuronowej asocjacji metody GMDH o kryterium zewnętrzne (oparte na informacji spoza rozpatrywanego zbioru danych uczących) związane z uwarunkowaniami technologicznymi. Te uwarunkowania technologiczne to możliwość uwzględnienia ograniczeń brokerskich.

Istotną wnoszoną w niniejszej pracy innowacją jest zastosowanie ograniczeń w strategii inwestycyjnej związanych bezpośrednio z technologiami stosowanymi na platformach internetowych. Innowacja ta, to specyficzna przestrzeń  $SL^{10} \times TP^{11} \times Tdop$  (poziom akceptowalnej straty/poziom satysfakcjonującego zysku/ustalony maksymalny okres

---

<sup>8</sup> To twierdzenie o krok od teorii chaosu, której Ivaknenko nie przywołuje.

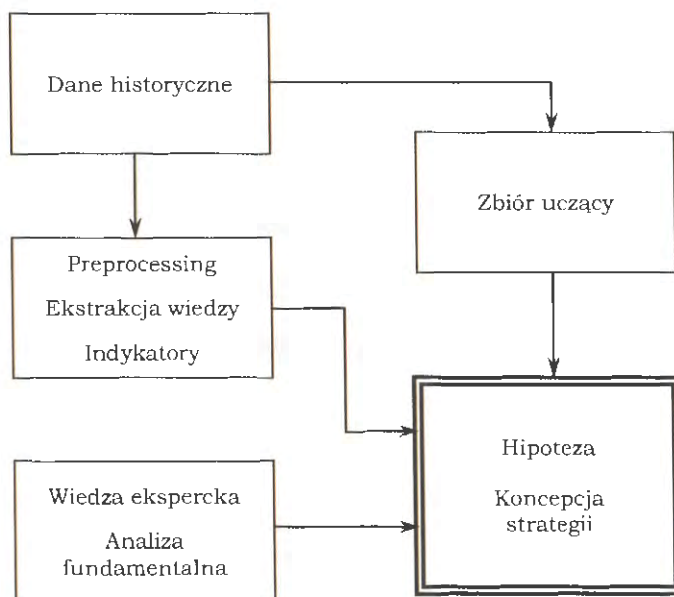
<sup>9</sup> Group Method of Data Handling – angielskie tłumaczenie metody opracowanej przez Ivakhnenkę w oryginale zwane *Metod Gruppovogo Uceta Argumetov* (ros.). W języku polskim nie ma dotąd jednoznacznego przyjętego powszechnie tłumaczenia. Właściwe wydaje się być – Metoda Grupowania Atrybutów czy Metoda Grupowego Uwzględniania Argumentów.

<sup>10</sup> SL – Stop Loss – stosowany na internetowych platformach brokerskich mechanizm półautomatycznego ograniczania starty. Polega on na ręcznym ustawieniu dopuszczalnej przez inwestora ponoszonej przez niego straty, w przypadku zmian rynku w kierunku niezgodnym z oczekiwaniami inwestora i automatycznym wykonaniu tego ograniczenia przez platformę, jeżeli rynek rzeczywiście zmieni się na niekorzyść gracza.

<sup>11</sup> TP – *Take Profit* – inny mechanizm półautomatycznego handlu na platformie brokerskiej. Jeżeli rynek zmieni się zgodnie z oczekiwaniami inwestora, to mechanizm TP dyskontuje zysk. To wyraz obawy inwestora przed utratą chwilowych korzyści, które już osiągnął.

otwarcia pozycji), charakterystyczna dla współczesnych platform internetowych w handlu (tradingu) automatycznym. Tu właśnie technologia miesza się z modelem teoretycznym – modelem wiedzy o obserwowanej rzeczywistości. Rzeczywistością tą dla inwestora jest szereg czasowy danych obserwowanych na rynkach finansowych. W pracy, w dalszej jej części, przedstawione zostaną przykłady zastosowania metody GMDH (dokładniej, niektórych, opracowanych na jej podstawie autorskich algorytmów) na różnych rynkach finansowych. Podstawą do oceny skuteczności algorytmu będzie najczęściej kryterium Sharpe’a (lub Sortino) [106, 109, 140] uwzględniające jednocześnie zysk i ryzyko (średnią stopy zwrotu i jej wariancję).

Ogólna koncepcja kojarzenia dwóch systemów – generowania koncepcji predykcyjnej i systemu podejmowania decyzji w warunkach konkretnej platformy technologicznej przedstawiona jest na dwu schematach.



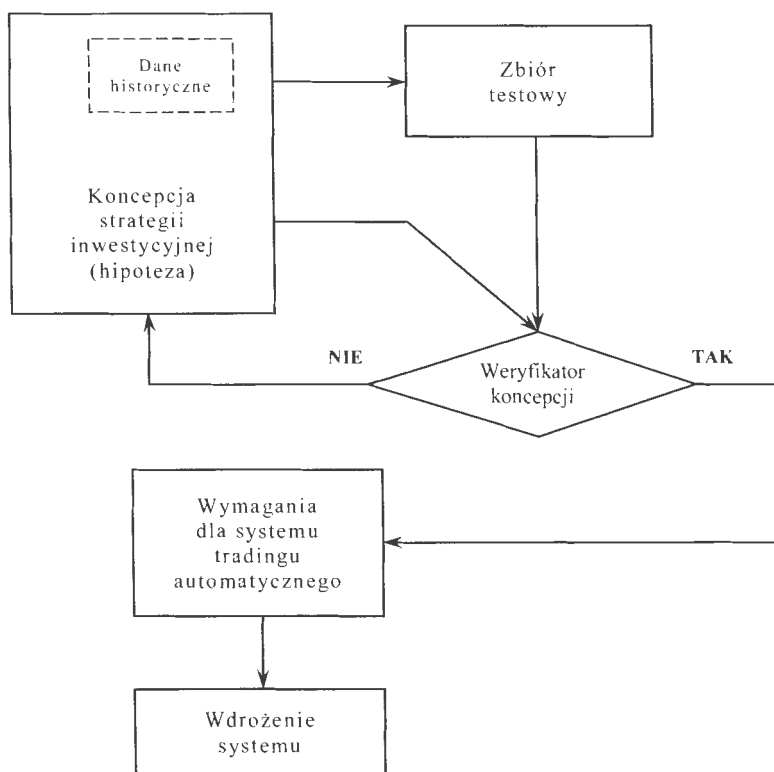
Rys. 1. Schemat blokowy procesu powstawania hipotez badawczych dotyczących koncepcji strategii inwestycyjnych.

Rys. 1 przedstawia schemat poszukiwania koncepcji strategii predykcyjnej. Rodzi się ona w przestrzeni danych historycznych w wyniku ustawicznej penetracji różnymi (z założenia heurystycznymi)

## Wprowadzenie

algorytmami tej przestrzeni, stosowaniu wyrafinowanej filtracji, uznanych i nowych indyktorów wysyłających sygnały do otwarcia lub zamknięcia pozycji. Do tych zabiegów, należących do analizy technicznej, dodać można ekspercką wiedzę o przyszłości rozpatrywanego rynku, wynikającą z przewidywanych skutków zdarzeń politycznych, gospodarczych czy społecznych. Symboliczną sumę tej wiedzy (model wiedzy) przedstawia rys. 1. Jej rezultatem jest pewna strategia, z którą w danej chwili utożsamia się inwestor w swym przekonaniu o jej skuteczności.

Gdy jest już o jej wartości przekonany, to warto podjąć decyzję o jej wdrożeniu. Schemat tego postępowania ilustruje rys. 2.



Rys. 2. Schemat blokowy procesu opracowywania wymagań dla systemu tradingu automatycznego.

Inwestor (gracz giełdowy, użytkownik platformy foreksowej) rozważa decyzje o testowaniu strategii. Racjonalne i logiczne jest, by dokonać tego na nowych, nieuwzględnianych wcześniej danych. To ty-

powy zabieg zwany cross-walidacją lub sprawdzeniem krzyżowym. Może on polegać na sekwencyjnym potraktowaniu przestrzeni danych historycznych (najpierw uczenie, potem testowanie) lub wykonywaniu testów co pewien czas w miarę przemieszczania się wzdłuż szeregu czasowego.

Jeżeli weryfikacja strategii jest pomyślna, to może nastąpić faza wdrożenia systemu z uwzględnieniem wymagań stawianych przez daną platformę inwestycyjną. Są to najczęściej wymagania natury technologicznej, np. uwzględnienie API<sup>12</sup> serwera, jeżeli jest udostępniane, uwzględnienie ograniczeń brokera, np. dopuszczalnej częstotliwości tradingu, dopuszczalnej odległości pomiędzy otwarciem pozycji a zleceniem zamknięcia, wielkości prowizji zależnej od zainwestowanej kwoty (najczęściej nieliniowej) itp.

Ograniczenia te odniesione do konkretnej rozpatrywanej strategii, o obiecujących wynikach uzyskanych w fazie testowania, mogą uczynić ją całkowicie nieprzydatną w praktyce. To właśnie istota rozpatrywanego tu zderzenia teoretycznej sprawności strategii z twardą rzeczywistością.

Twórca metody Group Method of Data Handling – prof. Ivakhnenko, początkowo nie rozpatrywał możliwości jej zastosowania na rynkach finansowych. Rozważał częściej przypadki predykcji ekonomicznej, ekologicznej, czy w systemach technicznych [37, 39, 48, 51]. Zrobił to dopiero pod koniec ubiegłego wieku wspólnie z niemieckimi badaczami Muellerem i Lemke [40, 41]. W tym czasie burzliwie rozwijały się statystyczne metody predykcyjne oparte najczęściej na ułomnych założeniach, np. o normalnym rozkładzie stopy zwrotu lub stacjonarności rynku. Pomimo kilku nagród Nobla dotyczących tego obszaru aktywności człowieka (Markowitz, Sharp, Merton, Scholes, McFadden, Engle) modele budowane przez te znakomitości nie doprowadziły, na szczęście, do ostatecznego zdefiniowania wiarygodnego modelu predykcyjnego. Nie doprowadza do niego, także na szczęście, ta praca. Doprowadza do lepszego zrozumienia aforyzmu Hegla: „To, co znane, nie jest jeszcze przez to, że jest znane, czymś poznanym”.

W odróżnieniu od wielu prac ekonometrycznych praca koncentruje się na pragmatycznym powiązaniu predykcji z możliwościami jej

---

<sup>12</sup> API – Application Programming Interface – interfejs programowania aplikacji, system umożliwiający, w danym rozumieniu, komunikację pomiędzy serwerem, na którym umieszczona jest aplikacja brokerska a klientem, który dzięki API ma możliwość oprogramowania swoich aplikacji w sposób umożliwiający komunikację z serwerem.

## *Wprowadzenie*

internetowej (automatycznej) realizacji. Jak wspomniano, wymaga to uwzględnienia szeregu niespodziewanych ograniczeń natury technologicznej. By udowodnić ten związek w pracy przedstawiono wyniki wielu eksperymentów w przestrzeniach danych historycznych różnych instrumentów finansowych. Wnioski są interesujące – zwłaszcza w wymiarze praktycznym.

Praca składa się z 6 rozdziałów. Pierwsze dwa poświęcone są szczegółowemu badaniu różnych odmian algorytmu kombinatorycznego zbudowanych na koncepcji algorytmu liniowego rozpatrywanego przez Ivakhnenkę [42, 50]. Zmodyfikowane algorytmy GMDH są doprowadzone do takiego stopnia szczegółowości, który umożliwia ich implementację w dowolnie wybranym środowisku programowym. Autor nie publikuje własnych skryptów opracowanych w środowisku Matlab z uwagi na ich rozmiary, być może także ze względu na niedoskonałość kodowania, ale także mając prawo do ochrony tych utworów jako dodatkowej wartości. Skrypty programowe towarzyszące tej pracy są niejednokrotnie bardzo rozbudowane i są często rezultatem wielu miesięcy intensywnych eksperymentów. W dodatku do pracy przedstawiony jest fragment najbardziej złożonego algorytmu GMDH pozwalający Czytelnikowi na opracowanie własnego programu symulacyjnego.

W rozdziale trzecim zbadane zostały wszelkie możliwe fragmenty wielowymiarowej przestrzeni predyktorów, które użyte były w pierwszych rozdziałach do badania skuteczności algorytmów GMDH. Trzy ostatnie rozdziały to prezentacja wyników rozmaitych badań z zastosowaniem tytułowej metody. Badania przeprowadzane są w różnych przestrzeniach danych, także wstępnie przetworzonych i w zaskakująco różnych, odmiennych stylach handlu. Ich wspólnym mianownikiem jest zastosowanie krocząco zmienianej regresji opartej na GMDH – ogólnie podsumowując z wyjątkowo dobrymi rezultatami.

Praca jest przygotowaniem do handlu automatycznego. Wyników takich eksperymentów, z wykorzystaniem systemu informatycznego umożliwiającego handel automatyczny, autor tu nie zamieszcza. Wszystkie symulacje wykonane są na różnych danych historycznych ze starannym odcięciem danych uczących od danych testowych. Celem zasadniczym, jest przygotowanie strategii do zastosowania w kompletnych systemach internetowych – zawierających moduły pozyskiwania bieżących danych, moduły dostępu do zasobów historycznych i moduły realizacji handlu wg przetestowanych i adaptowanych na bieżąco strategii inwestycyjnych.





# 2

## Badania algorytmów kombinatorycznych

### GMDH

*Nie wystarczy mieć sprawny umysł, trzeba go jeszcze dobrze używać*  
Kartezjusz (1596-1650)

#### 2.1. Badania algorytmu estymacji potencjału

Przez algorytm estymacji potencjału rozumie się tu będzie algorytm nazwany w poprzednim rozdziale CL1. Badania symulacyjne tego algorytmu przeprowadzono w środowisku Matlab. Celem badań było, jak wcześniej wspomniano, ustalenie optymalnej w danej chwili (kroku) struktury modelu i możliwej do uzyskania dokładności predykcji.

W dotychczasowych rozważaniach na temat metody GMDH nie została scharakteryzowana przestrzeń argumentów  $X_i$ . Argumenty, które wybiera się jako zmienne niezależne opisujące w szeregu czasowym zmienną wyjściową  $y$ , mają kluczowe znaczenie dla jakości prognozy. Wymagają one jednak tak szczegółowego przedstawienia i uzasadnienia, że poświęcony im będzie odrębny rozdział.

Na potrzeby testowania algorytmu CL1 wybrano  $M=26$  argumentów, w tym także argumenty tej samej zmiennej, ale przesunięte w czasie. Najogólniej je charakteryzując, były to rozmaite przekształcenia podstawowej świecy OHLC<sup>16</sup>, takie jak średnie kroczące, przesunięte

---

<sup>16</sup> Open, High, Low, Close – wartości zmiennej w wybranym przedziale czasowym

wartości poszczególnych składników świecy, momenty statystyczne zamknięcia świecy (np. wariancja, skośność i kurtoza), kwantyle rozkładu normalnego zbudowanego na najbliższej historii szeregu czasowego i inne.

Zmienną wyjściową (obserwowaną) w przedstawianych badaniach była wartość zamknięcia świecy 15 min<sup>17</sup> pary walutowej EUR/USD. Predykcji dokonywano najczęściej na odcinku  $h_p=1$  do 3, co oznaczało prognozowanie od dwóch do czterech punktów (wartość zamknięcia świecy EUR/USD) wprzód. Istotnym parametrem symulacji była także liczność wektorów argumentów i wektora wyników, którą po kilku testach przyjęto  $L_w=20$  (patrz rys. 13 i 14). Rozpatrywano także inne wielkości okna szeregu czasowego np.  $L_w=30, 50, 100$ .

Ze względu na czas obliczeń ograniczono także maksymalną złożoność modelu predykcyjnego do wartości  $S_{max}=6$ . Oznacza to, że najwyższy stopień (liczba składników) równania regresyjnego był równy  $S_{max}$ . W danym przypadku rozpatrywanego tu algorytmu oznaczało to  $6 \times 2 + 1 = 13$  współczynników liczbowych tego równania – w każdym stopniu (warstwie sieci neuronowej) dokładane były dwa nowe wyrazy. Wraz z wyrazem wolnym daje to powyższą liczbę współczynników.

Kolejnym istotnym parametrem symulacji była liczba powtórzeń całego cyklu prognostycznego, przez który należy rozumieć jeden krok w sieci jednokierunkowej polegający na obliczeniu prognozy dla  $h_p$  obszarów wprzód. Cykl ten można powtarzać w przestrzeni danych historycznych wiele razy, zarówno inicjując serie kroków w różnych punktach początkowych, stosując różne liczby kolejno wykorzystywanych kroków lub wykonując symulacje co ileś określonych okresów.

W prezentowanych tu wynikach symulacji wykonano  $k=10$  obliczeń inicjowanych co 2 świece w przestrzeni zbudowanej na danych związanych z parą EUR/USD w 2006 roku. Ponieważ horyzont predykcyjny  $h_p=3$  był większy niż odległość pomiędzy kolejno inicjowanymi obliczeniami, to odcinki predykcyjne częściowo zachodziły na siebie. Dawało to szansę (na razie nierozpatrywaną) wykorzystania obu kolejnych prognoz np. w aspekcie ich wzajemnego potwierdzania się lub ich sprzeczności. Niewielki fragment macierzy zmiennych w postaci świec 15 minutowych (obserwowanych co 15 minut) przedstawiono w listin-

---

<sup>17</sup> Świeca 15 min (czyt. świeca piętnastominutowa) oznacza wektor czterech wartości OHLC zaobserwowanych w przedziale 15 minut na wybranym rynku. Przyjętym standardem pomiarowym jest pobranie pomiaru otwarcia świecy 15 min (wartości O) jako pierwszej zaobserwowanej zmiany kursu po rozpoczęciu pełnego kwadransa.

gu oznaczonym jako Listing 1. Dane te dotyczą wartości kursu EUR/USD w 2006 roku.

Wprowadzenie wielu powtórzeń całego cyklu obliczenia modelu prognostycznego (tu  $K=10$  razy) pozwala na sprawdzenie jego wiarygodności w różnych miejscach szeregu czasowego, a także przeprowadzenie analizy ryzyka w systemie tradingowym zbudowanym (i dalej przedstawionym), skojarzonym z systemem prognostycznym.

1.2590	1.2595	1.2588	1.2591
1.2592	1.2592	1.2587	1.2589
1.2590	1.2593	1.2590	1.2593
1.2594	1.2601	1.2594	1.2598
1.2599	1.2600	1.2595	1.2597
1.2598	1.2602	1.2596	1.2601
1.2600	1.2600	1.2598	1.2598
1.2598	1.2599	1.2595	1.2598
1.2598	1.2603	1.2596	1.2603
1.2602	1.2608	1.2601	1.2605
1.2606	1.2607	1.2602	1.2606
1.2605	1.2606	1.2602	1.2605
1.2606	1.2607	1.2603	1.2603
1.2602	1.2604	1.2600	1.2602
1.2601	1.2603	1.2600	1.2601
1.2602	1.2607	1.2602	1.2606
1.2605	1.2612	1.2601	1.2603
1.2602	1.2605	1.2601	1.2604
1.2603	1.2612	1.2603	1.2610

Listing 1. Przykład danych ilustrujący zmienność kursu pary walutowej EUR/USD w formie macierzy, której wiersze oznaczają tzw. świece 15 min., a kolumny kolejno wartości O (Open – wartość kursu na początku okresu 15 min), H (High – wartość najwyższą w tym okresie), L (Low – wartość najniższą) i C (Close – wartość w chwili zakończenia okresu).

W każdym z  $K=10$  kroków symulacji obliczane były modele opisujące wektor zmiennej wyjściowej  $y$  kolejno w  $S=1, 2 \dots S_{\max}$  warstwach złożoności modelu. Istotnym problemem było racjonalne (heurystyczne) ograniczenie liczby wykonanych w każdym stopniu pętli, tak by osiągnąć akceptowalny czas obliczeń. W aspekcie implementacji problem ten rozwiązano poprzez określenie dopuszczalnej maksymalnej liczby modeli wybieranych w każdym stopniu (liczby  $F_w, F_b, F_v \dots$  rys. 8). W pierwszej warstwie sieci neuronowej z rys.10 rozpatrywano pary argumentów wejściowych ze zbioru wektorów  $X_i$  o długości każdego wektora  $L_w=|x_i|= 20^{18}$  i liczności zbioru argumentów wejściowych  $M=|X|= 26$ . Liczba możliwych kombinacji par argumentów wynosiła więc  $26 \times (26-1)/2$  i wszystkie te kombinacje były rozpatrywane.

Dla każdej pary argumentów  $(x_i, x_j)$  był obliczany liniowy model regresyjny postaci:

$$Z_{ij} = a_0 + a_1 X_i + a_2 X_j$$

W modelu tym wykorzystywano wartość wektorów  $x_i$  i  $x_j$ , ale rozpoczynające się  $h_p$  wyrazów wstecz od chwili bieżącej. Zgodnie z poprzednimi wyjaśnieniami wykorzystywano do porównań poprawności prognozy  $h_p$  ostatnich wartości zmiennej obserwowanej  $y$  – znane w chwili bieżącej. W ten sposób wykorzystując kryterium  $C_R$  (2) można było oszacować jakość predykcji dla każdego modelu (każdej pary argumentów  $x_i$  i  $x_j$ ).

Każda para pierwotnych argumentów  $(x_i$  i  $x_j)$  to nowy model  $\hat{Z}_{ij}$  i określający go wektor współczynników liczbowych równania regresji  $A_{ij}^z = [a_0 \ a_1 \ a_2]$ . Współczynniki te powinny być oznaczone jako  $a_{0(ij)^z}$ ,  $a_{1(ij)^z}$  itd. lecz z uwagi na rosnącą złożoność algorytmu zrezygnowano ze stosowania wszystkich indeksów (często stosowano to również w cytowanej tu literaturze specjalistycznej), pamiętać jednak należy, że np. każdy współczynnik liczbowy  $a_0, a_1, \dots$ , ma inną wartość w kolejnych równaniach regresji.

Ponieważ celem algorytmu CL1 jest ustalenie możliwości prognostycznych (maksymalnych korzyści teoretycznie możliwych do uzyskania) przy założeniu, że znany jest z danych historycznych rzeczywisty przebieg prognostycznego odcinka zmiennej wyjściowej, to w rozpatrywanej pierwszej warstwie dokonano (w każdym kroku  $k=1,2,\dots, K$ ) prognozy tego przebiegu z wykorzystaniem wiedzy o przestrzeni danych z bezpośredniej przeszłości.

---

<sup>18</sup> Symbolem  $|x|$  oznaczono liczbę wyrazów wektora  $x$ .

```

for i=1:M-1 //M - liczba argumentów - zmiennych wejściowych
pierwotnych
for j=i+1:M
  Xij'=[1 xi' dla i=t-hp-Lw...t-hp xj' dla j=t-hp-Lw...t-hp ]
  // macierz danych rozszerzona o wektor jedynek, przesunięta hp
  // kroków wstecz od chwili bieżącej złożona z trzech wektorów,
  // w tym wektora jedynek reprezentujących wyraz wolny//
  Aij'=Xij' \ ywk dla k=t-Lw...t //wektor współczynników liczbowych
  ŷwij=Xij'hp * Aij' //równanie regresji - model wektora zmiennej
  // wyjściowej prognozujący zmienna y hp kroków wprzód//
  ŷpij=ŷwij(max yp-min yp)+min yp
  //model ŷwij skonwertowany do wartości rzeczywistych (tu wartość kursu
  // pary walutowej), gdzie yp to wektor Lw wartości zmiennej wyjściowej
  // od chwili bieżącej plus hp kroków wprzód poczynając - w tym modelu
  // zakłada się, że przyszłość jest znana z danych historycznych//

  s11 = ∑Lw (ywk-ŷwij)2
  s21 = ∑Lw (ywk-ywk)2
  Δ1 = s11 / s21
  //kryterium selekcji par argumentów w pierwszej warstwie sieci Δ1wg (2)//

  S1a = ∑hp (ywk-ŷwij)2
  S2a = ∑hp (ywk-ywk)2
  Δ1a = S1a1 / S2a1
  //kryterium selekcji par zbudowane na ostatnich hp różnicach
  // pomiędzy rzeczywistymi a estymowanymi wartościami zmiennej
  // wyjściowej//
  if Δ1a ≤ CR1 then
  //selekcja w warstwie pierwszej//
  k1=k1+1;
  if k1 ≥ Fy then <wyjście z pętli i,j>
  // licznik par (i,j) spełniających warunek selekcji//
  B(k1)=[Δ1a ij Aij' ŷwij] ;
  // wektor cech charakteryzujących model spełniający warunek
  // selekcji//
  end //koniec selekcji//
end //koniec pętli j//
end //koniec pętli i//

```

Listing 2. Operacje realizowane w pierwszej warstwie sieci neuronowej.

Formalnie tę część algorytmu (pierwszą warstwę) przedstawić można w postaci kolejnych operacji przedstawionych jako Listing 2. W formie pseudokodu przedstawiono tu wszystkie istotne operacje wykonywane w pierwszej warstwie sieci neuronowej w pętłach  $i=1, 2, \dots, M-1$  oraz  $j=i+1, i+2, \dots, M$ .

Pętle te generowały kolejne pary argumentów  $(x_i \text{ i } x_j)$ .

Każda para umożliwiała przygotowanie modelu regresyjnego z udziałem dwóch argumentów (tych właśnie dwóch wybranych) w postaci równania liniowego z wyrazem wolnym.

Na ten lokalny algorytm obliczenia modelu w formie wielomianu opisującego w przybliżeniu zmienną obserwowaną  $y$  składały się zawsze następujące kroki:

- utworzenie rozszerzenia macierzy złożonej z dwóch wybranych kolumn  $x_i$  i  $x_j$  o wektor jedynek w celu przygotowanie odpowiedniej postaci modelu – takiej, by występował w niej wyraz wolny;
- obliczenie współczynników równania regresji  $A$ , takich, które na rozpatrywanej długości wektorów danych dają najmniejszy błąd wg metody najmniejszych kwadratów;
- obliczenie przebiegu zmiennej modelowanej  $\hat{y}$  wg obliczonych współczynników  $A$ ;
- analiza przebiegów zmiennej obserwowanej  $y$  i jej modelu  $\hat{y}$  w celu określenia rozbieżności pomiędzy modelem i rzeczywistością.

Schemat ten będzie wielokrotnie powtarzany także w kolejnych warstwach sieci.

Ta część obliczeń modeli pierwszej warstwy kończyła się więc przygotowaniem zbioru modeli o różnych wskaźnikach rozbieżności. Jeżeli stosowane były dwa kryteria to naturalnie powstawały dwa zbiory modeli o różnych preferencjach.

W charakterystyce algorytmu przedstawionej w pseudokodzie Listing 2. przedstawione są np. dwa kryteria jakości modelu. Pierwsze  $\Delta_1$  bazuje na zgodności przebiegów obserwowanej wzdłuż całego wektora o długości  $L_w$ , drugie  $\Delta_1_a$  na badaniu zbieżności wyłącznie na odcinku predykcyjnym  $h_p$ .

Na podstawie tych kryteriów (ewentualnie także innych, o ile będą logicznie uzasadnione) zbiory tych samych modeli można posortować wg stopnia spełnienia oczekiwania zbieżności z wynikami rzeczywistymi  $y$ . Sortowanie to formalnie zapisano w schemacie zapisanym jako Listing 3.

```

B1opt = sortowanie B(k1) wg pierwszej cechy Δ1a
//sortowanie przeprowadzone tak, by na pierwszym miejscu -
wierszu w macierzy B1opt był model z najmniejszą wartością kry-
terium Δ1a//
Δśr 1a =  $\sum_k^{\text{k1max}} \Delta 1k / k1 \max$ 
Δśr 1b =  $\sum_k^{F_y} \Delta 1k / F_y$ 
//obliczenie średnich wartości wskaźników rozbieżności najlep-
szych modeli - dla wszystkich spełniających kryterium i dla Fy
najlepszych spośród nich//
//koniec obliczeń pierwszej warstwy
    
```

Listing 3. Operacje sortowania wykonane po wyjściu z pierwszej warstwy sieci.

W powyższym pseudokodach zastosowano oznaczenia:

$t$  – oznacza chwilę bieżącą (jej indeks);

$X_{ij}'$  – macierz o wymiarach  $[3 \times L_w]$ , gdzie pierwsza kolumna zawiera jedynki ( dla obliczeń wyrazów wolnych), a dwie następne – wektory wybranych w danej pętli zmiennych wejściowych  $x_i$  i  $x_j$ . Apostrof  $x'$  oznacza, że zmienne zostały skonwertowane do postaci znormalizowanej (jak w (3));

$A_{ij}'$  – wektor współczynników liczbowych równania regresji  $[1 \times 3]$ ;

$h_p$  – horyzont prognostyczny (liczba kroków predykcji);

$y_{wk}$  – wektor zmiennej wyjściowej  $[1 \times L_w]$  do chwili bieżącej (znormalizowany);

$\hat{y}_{wij}$  – model zmiennej wyjściowej  $[1 \times L_w]$  zbudowany na podstawie równania regresji dwóch przesuniętych o  $h_p$  wstecz wektorów zmiennych wejściowych  $x_i$  i  $x_j'$ ,

$X_{ij}^{hp}$  – macierz zmiennych wejściowych (jak  $X_{ij}'$ ) wykorzystująca także  $h_p$  ostatnich wartości wektorów  $x_i$  i  $x_j$ ;

$\hat{y}_{pij}$  – wektor zmiennej wyjściowej utworzony na podstawie równania regresji ze zmiennymi  $x_i$  i  $x_j$  skonwertowany na wartości rzeczywiste (nie względne) zmiennej.

$\Delta 1$  – kryterium selekcji modeli w pierwszej warstwie zbudowane na różnicach całych wektorów – rzeczywistych modelowanych wartości zmiennej wyjściowej  $y$ ;



$\Delta 1a$  – kryterium selekcji modeli pierwszej warstwy zbudowane na różnicach pomiędzy rzeczywistym i prognozowanym przebiegiem zmiennej wyjściowej obserwowanym wyłącznie w odcinku predykcyjnym  $h_p$ .

$k1$  – licznik modeli warstwy pierwszej spełniających zadane kryterium selekcji  $C_{RI}$ ;

**$B1(k1)$**  – macierz zawierająca informacje o wszelkich niezbędnych cechach modelu predykcyjnego

**$B1_{opt}$**  – macierz zawierająca wektory  $B1$  uporządkowane (posortowane) wg rosnącego kryterium dokładności predykcji  $\Delta 1$ .

Pozostałe zmienne są dostatecznie objaśnione w komentarzach //...// umieszczonych w pseudokodzie.

Druga i kolejne warstwy tego algorytmu różnią się od pierwszej istotnymi własnościami. Po pierwsze w każdej warstwie następuje przyrost liczby kolumn w macierzy zmiennych wejściowych o dwie kolumny w każdej kolejnej warstwie. Po drugie – modele stanowiące owe kolumny przesunięte są w stosunku do wektorów pierwotnych danych  $x_i$ ,  $x_j$  o  $h_p$  wyrazów do przodu, są bowiem rezultatem prognozy (oceny) przeprowadzonej w pierwszym stopniu. Pseudokod (Listing 4) dotyczący drugiej warstwy sieci obserwowanej to, w związku z tym, już cztery zagnieżdżone pętle (dwie pierwsze sięgają jak poprzednio zbioru argumentów pierwotnych  $X_i$  a dwie następne dotyczą rezultatów modelowania – wyników pierwszej warstwy).

Skrypt zaimplementowany na podstawie algorytmów przedstawionych kolejno w postaci pseudokodu w Listing 4 i Listing 5 dotyczy drugiej warstwy sieci przedstawia naturalną powtarzalność algorytmu – dla istniejącego na podstawie danych historycznych wektora zmiennej wyjściowej obliczane są opisujące go modele, z nich wybierane są najlepsze wg pewnego zewnętrznego kryterium. Kolejne modele o wyższej złożoności (z wielomianami wyższych stopni) budowane są tak długo, aż nastąpi pogorszenie kryterium.

Dla uzmysłowienia różnic w jakości tej prognozy przeprowadzono badania w różnych miejscach szeregu czasowego i w różnych warstwach sieci.

## Rozdział 2. Badania algorytmów kombinatorycznych GMDH

```

for i =1: M-1
  for j=i+1 : M
    for ii=1 : Fy-1
      for jj=ii+1 : Fy
//Fy to heurystyczne ograniczenie liczby najlepszych modeli pierw-
szej warstwy //
      Zi,j,ii,jj=[1 x1' :i=t-Lw...t j':j=t-Lw...t B1opt(ii) B1opt(jj)];
      Az = Zi,j,ii,jj \ ywk : k= t-Lw+hp...t+hp
//wektor współczynników liczbowych równania regresji - tu [1x5]//
      Zw = Zi,j,ii,jj * Az
// model wektora zmiennej wyjściowej y prognozowany przez wymie-
szanie czterech argumentów - powinien mieć takie cztery indeksy,
które dla przejrzystości , opuszczono //
      s12 = ∑Lw (ywk - ZΛ)2
      s22 = ∑Lw (ywk - ywk-)2
      Δ 2 = s12 / s22
// kryterium jakości aproksymacji wg (2) dla drugiej warstwy ob-
liczone na podstawie regularności całego wektora o liczności Lw
//
      s1a2 = ∑hp (ywk - ZΛ)2
      s2a2 = ∑hp (ywk - ywk-)2
      Δ 2a = s1a2 / s2a2
//kryterium selekcji najlepszych kombinacji czterech argumentów
obliczane na podstawie rozbieżności ostatnich hp wartości wektora
rzeczywistej zmiennej ywk i modelowanej ZΛ//
      if Δ 2a ≤ CR2 then
//selekcja w warstwie drugiej przy progu selekcji CR2 //
      k2 = k2+1;
      if k2 ≥ Fz then <wyjście z pętli>
//licznik modeli spełniających kryterium selekcji oraz sprawdzenie,
czy tych modeli jest wystarczająco dużo- Fz //
      B (k2) = [Δ 2a i j ii jj Az ZΛ ]
// wektor cech modelu spełniającego warunki selekcji//

      end // koniec pętli jj//
      end //koniec pętli ii//
      end // koniec pętli j//
      end // koniec pętli i//

```

Listing 4. Operacje realizowane w drugiej warstwie sieci neuronowej.

```

B2opt = sortowanie B (k2) //wg pierwszej cechy Δ 2a w wektorze B,
tak by na pierwszym miejscu – wiersza macierzy B2opt był mo-
del z najmniejszą wartością Δ 2a//

Δsr2a = ∑kk2max Δ2k/k2max

Δs2b = ∑kFz Δ2k/ Fz

//obliczenie średnich wartości jakości modeli drugiej warstwy dla
wszystkich wyselekcjonowanych przypadków oraz dla Fz przypadków
najlepszych spośród nich//
//koniec obliczeń drugiej warstwy sieci//
    
```

Listing 5. Operacje sortowania wykonane po wyjściu z drugiej warstwy sieci.

Skuteczność predykcji w warstwie pierwszej można ocenić (bardzo wstępnie) na przykładzie wykresu rzeczywistego przebiegu<sup>19</sup> (wewnętrzne dwie krzywe rys. 13) – oznaczonego w cytowanych pseudokodach  $y_{wk}$  oraz przebiegu symulowanego przez najlepsze pary argumentów (zewnętrzne krzywe na rys. 15 oznaczone dodatkowymi punktami).

Wykres sporządzony jest w dziedzinie zmiennych znormalizowanych (stąd indeks „w” – względny) i widać tu dużą różnicę pomiędzy rzeczywistą i prognozowaną wartością dla trzech ostatnich okresów  $h_p=3$ .

Wykres sporządzony został dla kryterium obliczanego na całej długości porównywanych wektorów (tu dla  $y_{wk}$ , który liczy  $L_w=26$  składników), na wykresie natomiast pokazany jest końcowy, najbardziej interesujący fragment rozbieżności dla  $h_p$  ostatnich wyrazów.

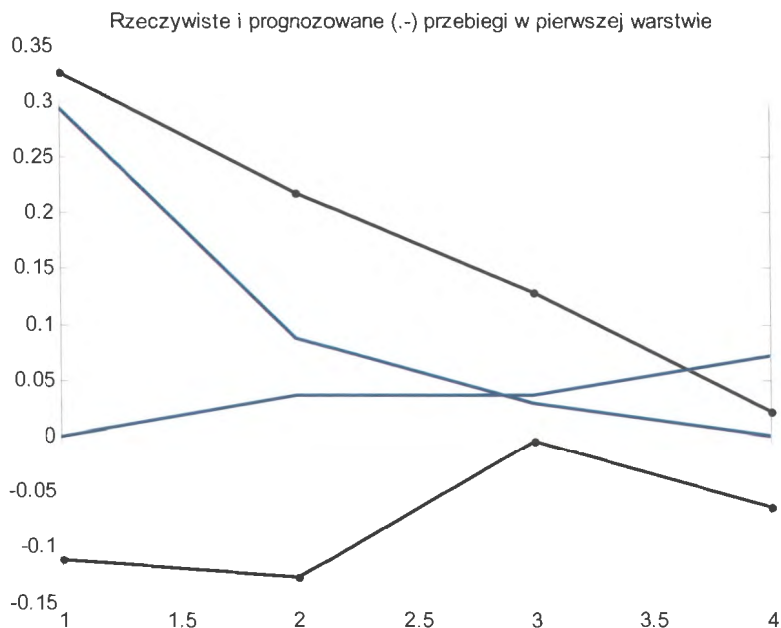
Istotne jest także jaka jest względna dokładność prognozy w układzie skonwertowanym do zmiennych rzeczywistych ( $w_g$  (3)) do postaci:

$$y_{pi} = y_{wi}(\max_i y_{pi} - \min_i y_{pi}) + \min_i y_{pi} \quad (10)$$

Zależności te, dla dwóch przypadkowo wybranych symulacji na odcinku  $h_p$  przedstawiono na rys. 16.

<sup>19</sup> Dane dotyczą pary EUR/USD od świecy nr 800 z roku 2006 poczynając, licząc kolejno od początku roku (bez wypełniania weekendów). Druga para wykresów sporządzona jest dla świecy nr 805.

## Rozdział 2. Badania algorytmów kombinatorycznych GMDH



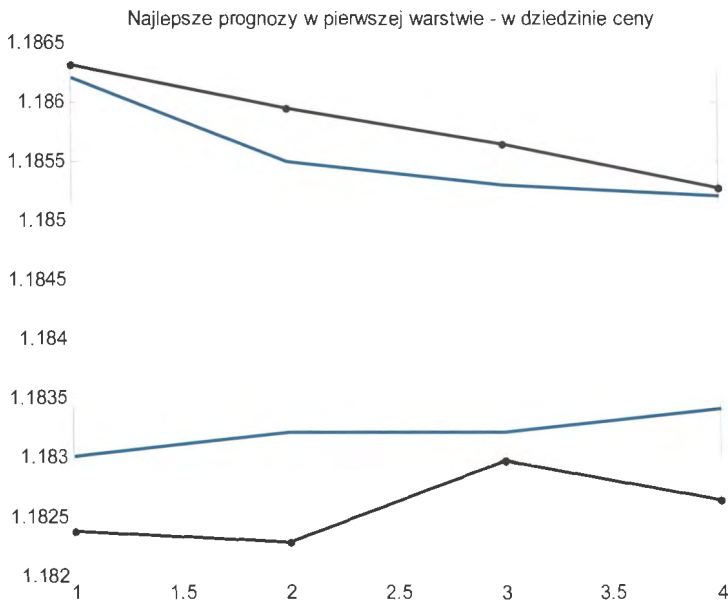
Rys. 15. Rzeczywiste (krzywe ciągłe) i prognozowane (krzywe z punktami) przebiegi zmiennej wyjściowej dla horyzontu prognozy  $h_p = 3$  w pierwszej warstwie obliczeniowej.

Na wykresie rys. 16 widać podobieństwo przebiegów do tych, przedstawionych w układzie zmiennych znormalizowanych rys. 15. Wykres ten niesie jednak podstawową informację o skuteczności predykcji - widać na nim, że odchylenia pomiędzy rzeczywistą i prognozowaną krzywą dla obu przedstawionych przypadków są rzędu 2-7 pipsów<sup>20</sup>. To dość obiecujący, choć nie rewelacyjny rezultat biorąc pod uwagę 15 minutowy horyzont prognozowania i obligatoryjny spread<sup>21</sup>

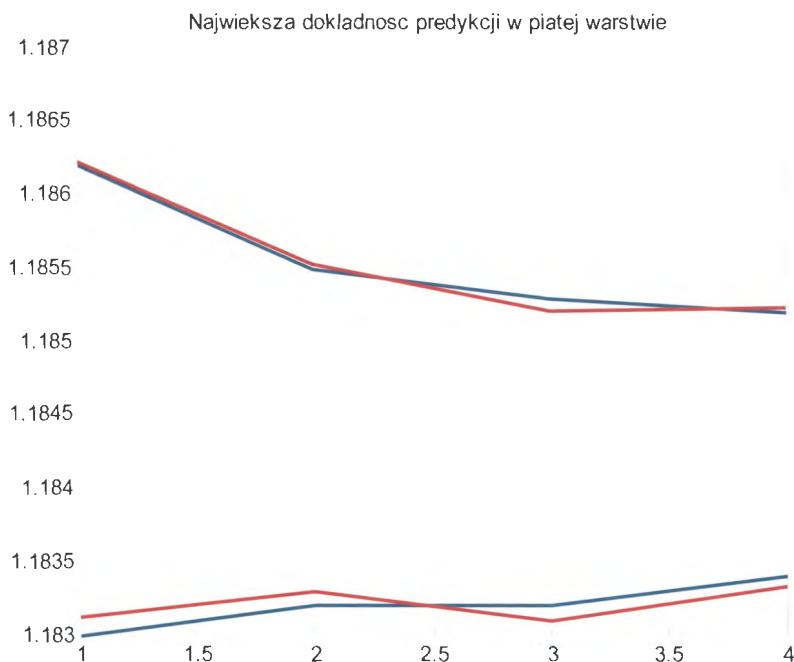
<sup>20</sup> Jednostka zmiany wartości kursu (w dziedzinie dyskretnej) wynosząca jedną dziesięciotysięczną. Miara ta dotyczy wszystkich popularnych par walutowych takich jak EUR/USD, GBP/USD, EUR/CHF, GBP/EUR i innych, których wartości są rzędu ok. 1-2. Pary walutowe związane z jenem np. USD/JPY mają inną miarę - pips to zmiana kursu o jedną tysięczną. Określenie pips jest powszechnie używanym pojęciem w żargonie środowiskowym, od liczby mnogiej w języku angielskim. Przyjętego tłumaczenia na język polski nie ma.

<sup>21</sup> Spread (ang.) to powszechnie stosowana nazwa różnicy ceny zakupu i sprzedaży na rynkach walutowych (na platformach brokerskich) równoznaczna prowizji brokera.

brokera rzędu 2-3 pipsów. Należy jednak pamiętać, że to rezultat najlepszej pary pierwszej zaledwie warstwy sieci. Ta sama konfrontacja rzeczywistego i prognozowanego przebiegu przeprowadzona w warstwie piątej rozpatrywanej sieci, w układzie zmiennych skonwertowanych do rzeczywistych zmian kursu daje obraz jak na rys. 17. Tu największe odchylenie pomiędzy rzeczywistymi i prognozowanymi wartościami kursu wynosi 2 pipsy. To odchylenie pomiędzy przebiegami może, lecz nie musi być inspiracją do poszukiwania strategii inwestycyjnej. Pierwszy, dość oczywisty wniosek o możliwości wykorzystania predykcji o dokładności ilustrowanej na rys. 17 jest taki, że dokładność ta (rzędu kilku pipsów) może być bardzo atrakcyjna, lub całkowicie bezużyteczna. Zależy to przede wszystkim od lokalnej zmienności  $y$ , zgodnie zresztą z (3). Jeżeli zmienność ta jest relatywnie duża to przy dokładności rzędu kilku pipsów (często nawet 1-2 pipsów), metoda może być bazą do opracowania skutecznej strategii. Jeżeli jednak w rozpatrywanym okresie symulacji zaobserwowany byłby płaski, horyzontalny trend zmiennej wyjściowej, to błąd rzędu kilku pipsów byłby istotnym zakłóceniem strategii.



Rys. 16. Rzeczywiste (krzywe ciągłe wewnętrzne) i prognozowane (krzywe z punktami) przebiegi zmiennej wyjściowej przedstawione w układzie rzeczywistych (nieznormalizowanych) wartości kursu EUR/USD dla pierwszej warstwy sieci neuronowej.

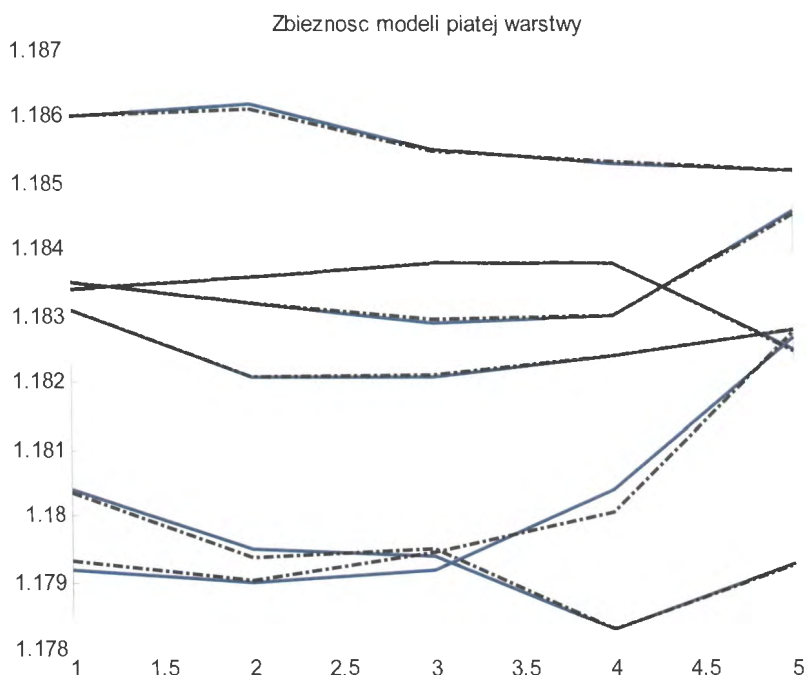


Rys. 17. Rzeczywiste i modelowane przebiegi zmiennej wyjściowej na odcinku predykcyjnym. Widoczne są niewielkie rozbieżności w porównaniu do rys. 16.

By dokonać pełnej analizy możliwości prognostycznych zbadano wiele innych aspektów modelu.

Przedstawioną na rys. 17 zbieżność rzeczywistego i prognozowanego przebiegu zmiennej wyjściowej sprawdzono wielokrotnie. Poglądowe wyniki przedstawiono na rys. 18. Na wykresie tym sporządzonym w dziedzinie rzeczywistych wartości kursu walutowego przedstawiono 10 obok siebie położonych (rozpoczynanych co 2 świece) przypadków porównania rzeczywistych i modelowanych przebiegów zmiennej wyjściowej. Każda para tych przebiegów dotyczy 4 odcinków horyzontu predykcyjnego  $h_p=4$  (stąd wyznaczone są wartości w punktach 1, 2, ..., 5).

Nie wdając się w szczegółowe rozważania dotyczące poszczególnych par zauważyć można dużą zbieżność rzeczywistych (krzywe łamane ciągłe) i modelowanych (krzywe przerywane) przebiegów. Symulacje przeprowadzono do szóstego stopnia złożoności (szóstej warstwy sieci), a rys. 18 przedstawia dla przykładu porównanie przebiegów rzeczywistych z symulacjami najlepszych wyników piątej warstwy.



Rys. 18. Rzeczywiste (ciągłe) i symulowane (przerywane) przebiegi zmiennej wyjściowej w piątej warstwie sieci neuronowej. Widać dużą zbieżność krzywych odzwierciedlających modelowane zmiany kursu z przebiegami rzeczywistymi.

Charakterystyczna jest dość duża łatwość, z którą można wyróżnić w gąszczu krzywych przebiegi dotyczące tego samego okresu predykcyjnego. Są zazwyczaj położone bardzo blisko siebie.

Można ocenić maksymalne błędy w estymacji kursu dla różnych przypadków na kilka zaledwie pipsów – np. w dalszej części wykresu pośrodku widoczne jest odchylenie rzędu 3-4 pipsów i jest to chyba największy błąd predykcji dla wszystkich sześciu symulacji przedstawionych na wykresie.

Należy przy tym pamiętać, że są to badania wykonane dla algorytmu CL, którego istotą jest założenie, że znane są dane historyczne pozwalające na bezpośrednią i natychmiastową ocenę dokładności modelu (kolejnych modeli). Celem nadrzędnym tego algorytmu jest bowiem oszacowanie możliwej do uzyskania dokładności poprzez zwiększenie złożoności modelu oraz badanie jego skuteczności predykcyjnej na znanym przebiegu końcowego odcinka zmiennej wyjściowej. Innymi

słowy, przebiegi te wskazują, jak doskonała mogłaby być prognoza na odcinku aż 4 okresów wprzód, gdyby zastosowano najlepszy w danej chwili model o złożoności określonej przez 5 warstwę rozpatrywanej sieci obliczeniowej.

Zakładając (także opierając się na wnioskach z rys. 18), że dość dobre rezultaty uzyskuje się w piątej warstwie (być może także nieco wcześniej, jednak piąta jest z pewnością wystarczająco złożona), zbadano zmienność modeli tej warstwy.

To jeden z najistotniejszych wyników badanego algorytmu CL1. Gdyby zmienność modeli, co do ich struktury polegającej na wykorzystywaniu określonych argumentów była mała, umożliwiałoby to postawienie hipotezy o pewnej stabilności modelu (a także całego modelowanego środowiska – modelu fizycznego) i budowanie koncepcji predykcji na tej stabilności.

Rezultaty te przedstawiono na wykresie 19. Poszczególne linie poziome odpowiadają użytym w modelu argumentom pierwotnym, o ile poziom ten oznaczony jest okrągłym markerem. Linie pionowe odpowiadają kolejnym przypadkom prognozy inicjowanym w różnych punktach szeregu czasowego (dla różnych numerów świec). W danym przypadku są to modele dla tych samych chwil czasowych, dla których sporządzone zostały badania zbieżności (rys. 16), czyli co 10 okresów.

Na rys. 19 widoczne jest wielokrotne powtarzanie niektórych argumentów w modelu. Przez strukturę modelu rozumie się tu liczbę i indeksy argumentów ułożonych na jednej pionowej, czyli w tym samym przypadku predykcyjnym.

Najistotniejszym wnioskiem tego najważniejszego badania jest konstatacja ogromnej zmienności modeli. Na wykresie nie ma dwóch identycznych pod względem użytych argumentów modeli. Zmieniają się, często bardzo znacznie, w każdym kolejno rozpatrywanym przypadku (tu w sześciu przypadkach, zbadanych co 10 świec). W dalszej kolejności wyjaśniony zostanie powód wielokrotnego użycia tego samego argumentu pierwotnego w poszczególnych warstwach sieci.

Wniosek o korzystnym z punktu widzenia kryterium zawężeniu zbioru argumentów jest wielokrotnie podkreślony w literaturze naukowej poświęconej problematyce klasyfikacji i predykcji [4, 16, 28]. Z zestawu modeli przedstawionych na rys. 19 wynika, że niektóre z argumentów opisujących zmienną wyjściową są szczególnie (w danym momencie) atrakcyjne inne zaś są ignorowane. Te szczególnie atrakcyjne są wielokrotnie używane w poszczególnych warstwach sieci co skutkuje kumulacją wielokrotnego ich użycia w rozpatrywanej tu piątej warstwie.

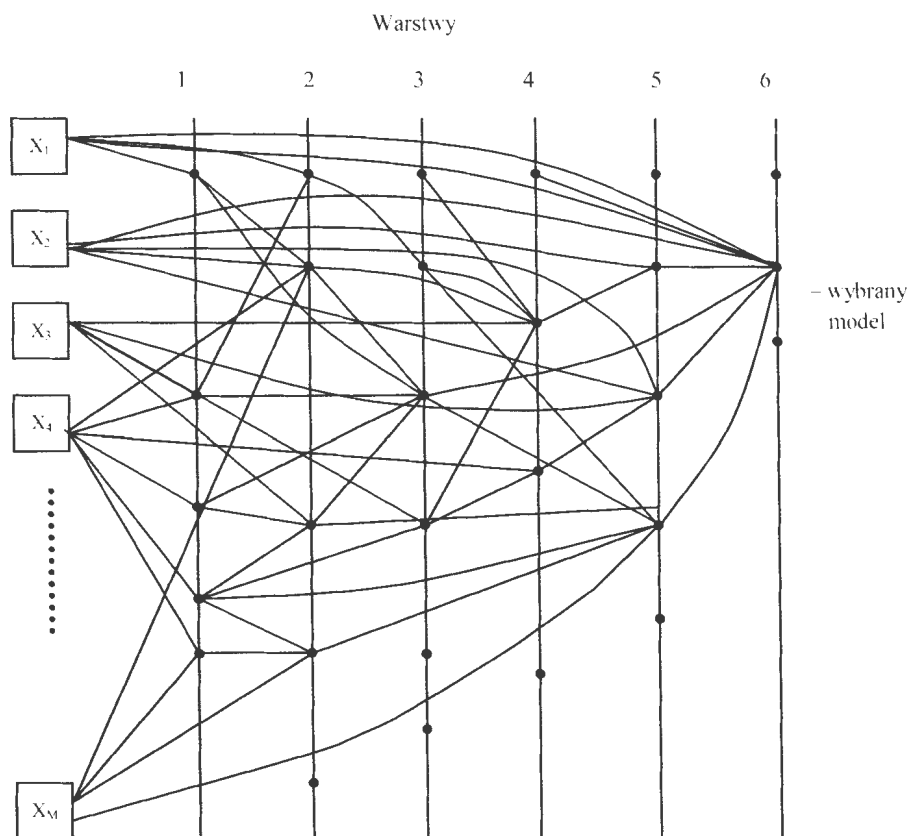




Rys. 19. Struktura najlepszych modeli piątej warstwy sieci neuronowej dla sześciu kolejnych przypadków predykcji wykonanej co 10 okresów w szeregu czasowym świec 15 min dla pary EUR/USD.

Za chwilę zaś, te które były całkowicie ignorowane okazują się bardzo przydatne. To sytuacja charakterystyczna dla środowiska niestacjonarnego, jednak poddaje się one pewnemu uogólnieniu – nie jest konieczna cała dostępna informacja, by dokonać dobrej predykcji (przynajmniej w sensie lokalnym). Być może nawet nadmiar tej informacji przeszkadza i wywołuje pogorszenie prognozy [15, 16, 126].

Sytuację tę dobrze ilustruje schemat sieci neuronowej przedstawiony na rys. 20. Przy założeniu, że wybranym modelem jest ten, wskazany jako jedyny w szóstej warstwie, pokazano symbolicznie, jak do tego wybranego neuronu zbiegają sygnały wejścia kolejno po 2 ze wszystkich poprzednich warstw w tym także z warstwy wejściowej  $x_1, x_2 \dots x_m$ . Stąd liczba wejść do neuronu w 6-tej warstwie wynosi 12, a do neuronu w niżej położonych warstwach wynosi  $2 \cdot S$ , gdzie  $S$  – numer warstwy.

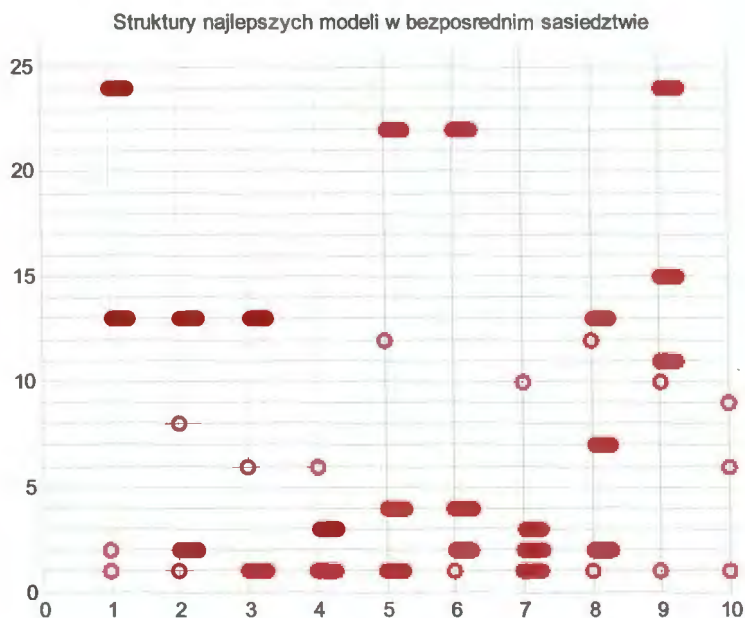


Rys. 20. Schemat sieci obliczeniowej z ujawnionymi wejściami do szóstej warstwy z neuronów poprzednich warstw i także z warstwy wejściowej.

Na schemacie tym widoczne jest wykorzystanie tylko niektórych neuronów, niekiedy wielokrotnie, co zgodne jest z wyżej przedstawioną analizą a także z rezultatami symulacji zaprezentowanymi na rys. 19.

Na rys. 21 przedstawiono struktury modeli badanych co 10 okresów. Można przypuszczać, że to dość znaczna odległość w szeregu czasowym i może być ona przyczyną tak znacznego zróżnicowania struktur. Dlatego zbadane też zostały struktury modeli co jeden krok (jedną świecę).

Wyniki tych badań przedstawiono na rys. 21.

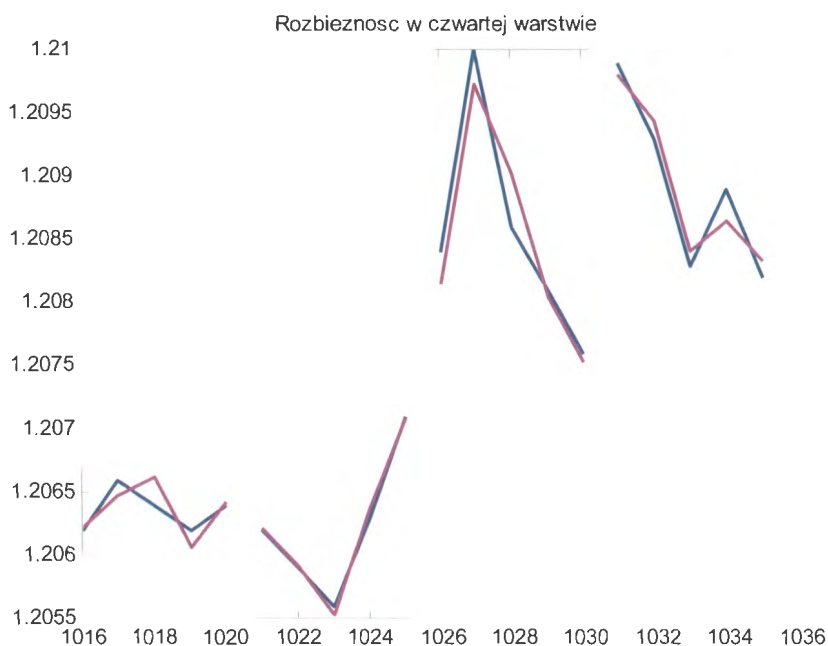


Rys. 21. Zróżnicowanie struktur najlepszych modeli piątej warstwy badane co jeden krok.

Widoczne jest także duże zróżnicowanie struktur. Nie da się zauważyć dwóch identycznych obok siebie położonych modeli. To dowód znacznej zmienności tego rynku (tu pary walutowej EUR/USD) utrudniającej bezpośrednią prognozę. To co jest interesującym rezultatem tych badań, to także wniosek o wykorzystywaniu, nierzadko częstotliwym, tych samych argumentów (na rys. 21 są to argumenty o numerach 1, 2, 3, 4, 13, 22, 24) i całkowitym pomijaniu innych (np. 5, 17, 18, 23). Nie ma jednak na tym etapie badań żadnej gwarancji, że w innym miejscu szeregu czasowego (np. 1000 świec dalej) numery aktywnie stosowanych i nieaktywnych argumentów będą zupełnie inne.

Istotnym także rezultatem badań algorytmu CL1 była ocena dokładności przedstawiona nie tak jak na wykresach rys. 15-18, lecz sekwencyjnie, nałożona w postaci krótkich odcinków prognozowanych przebiegów na rzeczywiste wartości zmiennej wyjściowej.

## Rozdział 2. Badania algorytmów kombinatorycznych GMDH



Rys. 22. Badanie rozbieżności pomiędzy rzeczywistym przebiegiem zmiennej wyjściowej i prognozowanymi odcinkami o długości  $hp=4$  okresy – w czwartej warstwie sieci neuronowej.

Badania takie wykonano dla przykładu w czwartej i szóstej warstwie sieci neuronowej. Wykresy dla czwartej warstwy sporządzono w dziedzinie wartości rzeczywistych (nie poddanych normalizacji)<sup>22</sup>.

Na rysunku 22 przedstawiono wyniki symulacji obserwowane w czwartej warstwie sieci. Wybrano najlepsze wyniki aproksymacji zapisane po sortowaniu na pierwszym miejscu w zbiorze B4opt (przez analogię do omówionego w przedstawionym pseudokodzie B1opt (Listing 3) – to taki sam zbiór wyników odnoszący się do czwartej warstwy sieci neuronowej). Na rysunkach widoczne są czteroodcinkowe fragmenty prognozowanych przebiegów zmiennej wyjściowej dla symulacji wykonywanej co 5 kroków.

<sup>22</sup> dla świec piętnastominutowych pary EUR/USD od numeru 1000 w 2006 roku poczynając.

Różnice pomiędzy krzywymi łamanymi, z których jedne symbolizują rzeczywiste przebiegi obserwowanej zmiennej, a drugie – przebiegi modelowane, są niewielkie. Wykres ten pozwala na bezpośrednią ocenę przydatności metody, lub raczej, upoważnia do dalszego jej rozważania, gdyż wskazuje na interesujący potencjał. Rozbieżności maksymalne na wykresie rys. 22 to 4-5 pipsów. Są jednak takie przebiegi (np. drugi z kolei odcinek  $h_p$ ), gdzie zbieżność jest niemal doskonała.

Podobne, lecz o jeszcze mniejszych rozbieżnościach przebiegi odcinków prognozowanych zaobserwowano po zakończeniu obliczeń w szóstej warstwie. Na wykresie (rys. 23) widoczne są także czteroodcinkowe fragmenty prognozowanych przebiegów zmiennej wyjściowej.

Wykres ten sporządzony jest dla odmiany w dziedzinie względnych (znormalizowanych) wartości zmiennej obserwowanej. Widać tu dużo większą zgodność niż w prezentowanej przed chwilą regularnością w warstwie czwartej.



Rys. 23. Badanie rozbieżności w szóstej warstwie w dziedzinie zmiennej wyjściowej znormalizowanej (przebiegi rzeczywiste – krzywe ciągłe, przebiegi symulowane – krzywe przerywane).

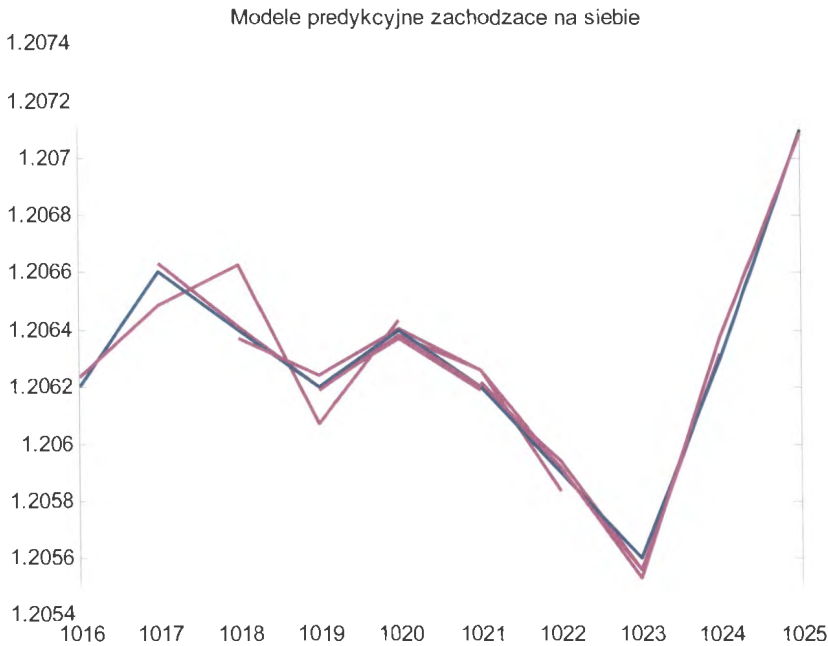
Zagadnieniem wymagającym odrębnego rozpatrzenia jest, czy celowa jest kontynuacja zwiększenia złożoności modelu poprzez dodawanie kolejnych warstw sieci obliczającej coraz bardziej skomplikowane modele predykcyjne. Czy nie okaże się, że wzrostowi złożoności przestaje towarzyszyć wzrost dokładności predykcji.

Ogólne wrażenie wynikające z pobieżnej nawet analizy obu ostatnich wykresów może być bardzo pozytywne w kontekście przyszłych praktycznych zastosowań predykcji w systemie transakcyjnym. Otóż, przy dużych nawet zmiennościach rzeczywistego przebiegu zmiennej wyjściowej widać dość ciasno towarzyszące tym zmianom przebiegi odcinków symulowanych zmian. Oznacza to, że jeżeli będzie podjęta próba zbudowania systemu transakcyjnego wykorzystująca wyraźne zmiany kursu, to zmianom tym mogą towarzyszyć stosunkowo dobre prognozy. Na wykresie rys. 22 widać, jak przy zmianach kursu rzędu 30-40 pipsów w czasie kilku świec, różnice pomiędzy krzywą rzeczywistych przebiegów i krzywymi prognostycznymi są rzędu zaledwie kilku pipsów.

Dodatkowo, prognozy te są wzajemnie potwierdzane poprzez nakładające się odcinki predykcji o długości  $h_p$ , wówczas, gdy badania wykonano z większą częstotliwością w tym samym szeregu czasowym (nie co 5 okresów jak wcześniej, lecz częściej).

Badania takie wykonano poprzez umieszczenie kilku symulacji o długości odcinka prognostycznego  $h_p=4$  okresy tuż obok siebie – co jeden okres. W ten sposób, tak jak na rys. 24 uzyskano nakładanie się kolejnych prognoz – i tym samym kolejne potwierdzenie (lub zaprzeczenie – gdyby przebieg kolejnej prognozy pozostawał w sprzeczności z poprzednimi).

Pozwala to już w tej chwili na rozważanie koncepcji strategii bazującej na wzajemnym potwierdzaniu predykcji przez kolejne symulacje  $h_p$  – odcinkowe. Logiczny jest np. następujący schemat podejmowania decyzji – jeżeli kolejne prognozy potwierdzą się – to będą wykorzystywane, a jeżeli są w sprzeczności, to decyzja o otwarciu pozycji będzie zaniechana. Oczywiście, jedynie statystyczne potwierdzenie tych rekomendacji będzie podstawą do pozytywnej oceny strategii.



Rys. 24. Wzajemnie potwierdzające się kolejne prognozy wykonywane sekwencja jedna tuż za drugą, w czwartej warstwie, na tle przebiegu rzeczywistego.

Jak wspomniano na wstępie rozdziału zasadniczym celem badania rozpatrywanego tu algorytmu CL1 było sprawdzenie możliwości uzyskania dokładności prognozy. Naturalnym oczekiwanym kierunkiem wzrostu dokładności prognozy winna być rosnąca złożoność modelu opisującego zmienną wyjściową.

Miarą dokładności prognozy użytą w rozpatrywanych pseudokodach jest wielkość  $\Delta 1_a, \Delta 2_a \dots \Delta S_a \dots \Delta S_{max_a}$ , gdzie S to stopień złożoności modelu (warstwa sieci). Na wyjściu z każdej warstwy otrzymuje się wiele modeli, można więc  $\Delta S_a$  traktować jako zbiór wyników  $\Delta S_a^1, \Delta S_a^2 \dots \Delta S_a^{F_s}$  gdzie  $F_s$  oznacza liczbę dopuszczalną modeli w danej warstwie.

W każdej warstwie można obliczyć średnią wartość dokładności  $\Delta S_a$ , przy czym istotne będzie to, które z modeli będą zaliczone do zbioru wpływającego na wartość średnią. Jeżeli np. pod uwagę weźmie się  $F_s$  wyników (modeli) to:

$$\bar{\Delta S}_a = \frac{1}{F_s} \sum_{k=1}^{F_s} \Delta S_a^i \quad (11)$$

gdzie:

$\bar{\Delta S}_a$  – średnia wartość dokładności predykcji w S-tym stopniu (S-tej warstwie sieci) (a – oznacza sposób obliczania średniej);

$F_s$  – liczba branych pod uwagę modeli w s-tej warstwie, przyjętej arbitralnie, tak by uwzględnić ograniczenia techniczne komputera (czas obliczeń);

$\Delta S_a^i$  – pojedynczy i-ty wynik – dokładność i-go modelu w S-tej warstwie.

$\Delta S_a^i$  zgodnie z (2) jest obliczane wg wzoru:

$$\Delta S_a^i = \frac{\sum_k (y_{wk} - \hat{y}_k)^2}{\sum_k (y_{wk} - \bar{y}_{wk})^2} \quad (12)$$

W powyższym wzorze indeks  $k$  związany jest z długością wektora zmiennej wyjściowej lub jej modelu i najczęściej ma wartość równą maksymalnej liczbie wyrazów  $L_w$  w tych wektorach, a indeks  $i$  jest związany z kolejnym rozpatrywanym modelem – jednym spośród wielu rozpatrywanych w kombinacjach danej warstwy.

Poza tym badano także dokładność predykcji nie dla wszystkich  $F_s$  lecz tylko dla wyselekcjonowanej najlepszej części spośród nich. Np. dla  $H_s$  najlepszych (w sensie dokładności  $\Delta S_a$ ) obliczono średnią:

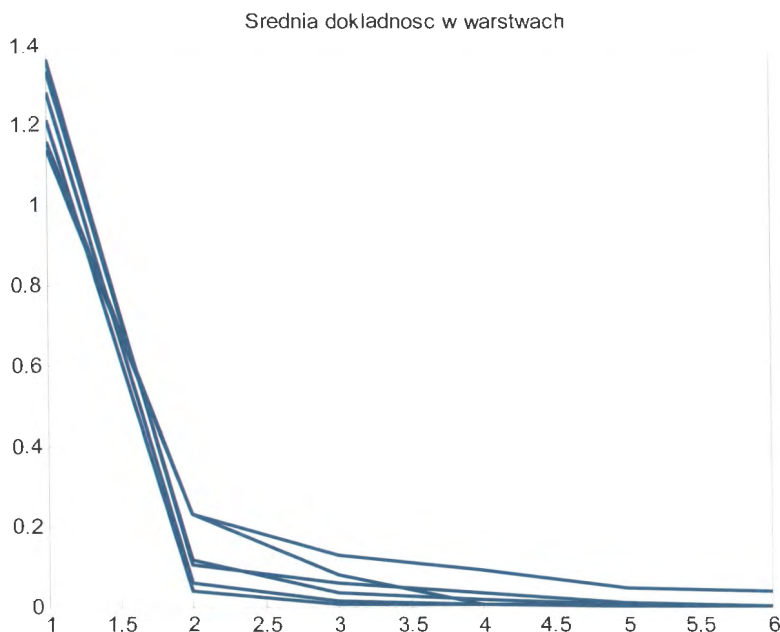
$$\bar{\Delta S}_b = \frac{1}{H_s} \sum_{i=1}^{H_s} \Delta S_a^i \quad (13)$$

gdzie  $H_s$  – ograniczenie liczby modeli w S-tej warstwie (najczęściej przyjmowano  $H=10$ , niezależnie od numeru warstwy);

Jeżeli  $H_s < F_s$  to oczywiście kryterium z lepszych wyników będzie lepsze (niższe) niż wynik dla wszystkich modeli, czyli  $\bar{\Delta S}_b < \bar{\Delta S}_a$ .

Wyniki tych średnich wartości dokładności predykcji przedstawiono na rys. 25.





Rys. 25. Przebiegi średnich dokładności  $\Delta S_s$  dla  $S=1,2,\dots,6$  obliczone w różnych miejscach przestrzeni danych. Każda krzywa oznacza wyniki obliczone dla innej świecy początkowej.

Ważnym wynikiem tych badań było także określenie maksymalnej możliwej do uzyskania dokładności w danym punkcie przestrzeni danych. Przez dokładność tę (2 i 13) rozumiano absolutne minimum kryterium niezależnie od złożoności modelu, czy warstwy sieci, w której zostało osiągnięte.

Poszukiwano więc:

$$\Delta_{\min}^l = \min_s \min_i \Delta S_a^i \quad (14)$$

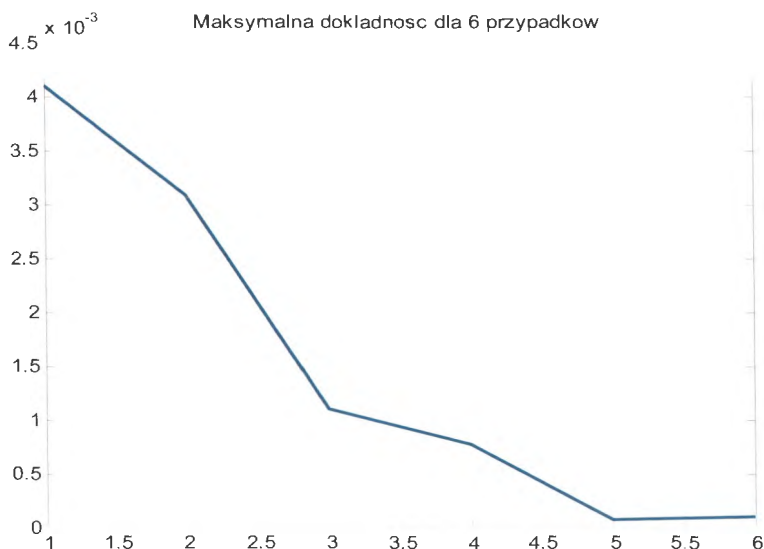
dla każdego miejsca  $l$  (świecy startowej) w przestrzeni danych, dla każdej  $s$  warstwy sieci neuronowej i dla każdego  $i$  – go modelu.

To minimum w kolejnych zadaniach predykcyjnych może być osiągnięte zarówno dla różnej w danej warstwie struktury modelu jak i w różnych warstwach.

Wyniki okazały się wielce interesujące. Po pierwsze, okazało się, że możliwe jest uzyskanie znacznie lepszych rezultatów niż średnie przedstawione w różnych warstwach dla różnych punktów startowych w przestrzeni danych, przedstawione na rys. 22. Po drugie, zgodnie z literaturowymi wskazówkami [50, 80] niekoniecznie najlepsze wyniki uzyskano w wyższych warstwach (o wyższej złożoności modelu).

Dla przykładu – w serii 6 symulacji co 2 świece od świecy nr 1000 w 2006 roku poczynając, uzyskano najlepsze wyniki  $\Delta'_{\min}$ ,  $l=1, 2, \dots, 10$ , raz w piątej warstwie, raz w czwartej warstwie i trzykrotnie w szóstej warstwie.

Te najlepsze wyniki przedstawiono na rys. 26.



Rys. 26. Najlepsze wyniki dokładności predykcji w różnych miejscach przestrzeni danych wg kryterium (2).

To wyniki dla 6 kolejnych przypadków przedstawione w dziedzinie znormalizowanej zmiennej wyjściowej (tzn. takiej, dla której w rozpatrywanym zakresie zmian wartość minimalna wynosiła 0, a maksymalna 1). Wyniki prezentowane na rys. 26 wyrażone są w tysięcznych (wszystkie poniżej 1% zakresu).

Oczywiście tak doskonale wyniki nie mają bezpośredniego przełożenia na jakość prognozy, gdy zgodnie z założeniami do przedstawi-

nego tu algorytmu CL1 poszukiwane tu są maksymalne możliwe do uzyskania dokładności. Towarzyszy temu sztuczne całkowicie założenie, że znana jest najbliższa przyszłość. Inaczej mówiąc, algorytm ten pozwala na znalezienie najlepszego modelu prognostycznego i oszacowanie możliwej do uzyskania dokładności, gdyby znana była przewidywana struktura tego modelu.

W rzeczywistości jest jednak inaczej. Stąd podstawowym założeniem algorytmu CL2 będzie próba wykorzystania najlepszego, określonego w danej chwili modelu w najbliższej przyszłości – w najbliższym kroku (ewentualnie kilku, lecz nie wiadomo z góry ile kroków predykcyjnych warto brać pod uwagę). To założenie, że właściwości przestrzeni danych mają pewną inercję umożliwiającą skorzystanie z najbliższych położonych cech (w sensie bliskości w szeregu czasowym), może być iluzoryczne. Nie można a priori przewidzieć słuszności lub niepoprawności tej tezy – zależy to wyłącznie od skłonności środowiska do ewolucji.

Wyniki przedstawione na rys. 26 są innowacyjne w sensie precyzji przewidywania. Ivakhnenko [50] przewidywał, że jeżeli wartości kryterium (2), także zbudowane w konsekwencji przekształcenia kryterium (13) będą rzędu 0.2-0.3 to będzie to wynik świadczący o uzyskaniu zadowalającej prognozy. Z wykresu rys. 25 wynika, że dla wyników od drugiej warstwy w górę wynik ten jest uzyskiwany. Skoro konstatuje się ten wynik dla wielu modeli (danej warstwy, danego kroku początkowego), to uprawnione jest przewidywanie, że wybrany model, doskonały przed chwilą, będzie jeszcze, może już nie tak doskonały, jednak satysfakcjonujący w kolejnym kroku. Należy sobie wyobrazić ogromną, rosnącą liczbę modeli w miarę wzrostu stopnia złożoności modeli (warstwy sieci neuronowej). Gdyby nie ograniczenia natury technicznej wynikające z mizernej obecnie mocy przerobowej współczesnych komputerów, możliwe byłoby pełniejsze wykorzystywanie kolejnych przypadków szeregu Gabora-Kołmogorowa. Ograniczenie tych rozważań do  $H=10$  najlepszych przypadków jest jedynie gestem wskazującym drogę, pomimo tego, że to 10 najlepszych rozwiązań. Prawdziwy model rzeczywistości, „fizyczny” jak wyrażał się o nim Ivakhnenko, jest jednak o wiele bardziej skomplikowany. Świat w ogóle jest złożony, ale sokratejska świadomość niewiedzy o jego złożoności winna być czynnikiem motywującym do poszukiwania prawdy – modelu (niefizycznego) najlepiej opisującego rzeczywistość. Podejście polegające na permanentnym wdrażaniu modelu w celu predykcyjnym jest chyba najbardziej bezwzględny, znanym nauce, sposobem weryfikacji adekwatności modelu do rzeczywistości.

## 2.2. Badania algorytmu predykcyjnego

Wracając do przedstawionej wcześniej różnicy pomiędzy algorytmami CL1 i CL2 należy podkreślić, że algorytm CL1 służy do oszacowania możliwości (ekstremalnych) predykcji w sensie jej dokładności, a algorytm CL2 służy po prostu do wykonania działania prognostycznego. Algorytm CL2 może być (będzie) więc elementem jeszcze nie zdefiniowanego algorytmu nadrzędnego – strategii inwestycyjnej polegającej na wykorzystaniu prognozy realizowanej na zmienności pary walutowej (w rozpatrywanym przykładzie EUR/USD).

Zasadniczo pierwsze części algorytmu CL2 (nazwa – autorska – ma związek z przymiotnikami „kombinatoryczny” (ang. *combinatorial*) i „liniowy”, gdyż te cechy są charakterystyczne dla rozpatrywanego tu algorytmu z rodziny GMDH), nie różni się od CL1 w wielu fragmentach obliczeniowych.

Zdecydowana różnica pojawia się w drugiej fazie odpowiadającej za zbudowanie prognozy. Może ona być wykonana na kilka sposobów, które w istocie rzeczy mogą na tyle różnicować rozważany algorytm, że mówić można o zupełnie odrębnych rozwiązaniach. Podstawą większości tych odmiennych sposobów drugiej, prognostycznej fazy algorytmu CL2 będzie jednak zawsze dekompozycja złożonego równania regresji wyższego rzędu (wygenerowanego w wyższej warstwie sieci neuronowej) na ekwiwalentne równanie zawierające tylko dane pierwotne z przestrzeni  $X_i$ . Zadanie to okazało się dość złożone od strony implementacyjnej i zostanie mu poświęcona odrębna część pracy (załącznik).

Ogólnie, **algorytm CL2** można podzielić na fazę uczenia sieci, w której ustalone są najlepsze modele w każdej z warstw, oraz fazę prognozy.

### Faza uczenia sieci

1. Obliczenie parametrów (współczynników liczbowych wielomianów) modeli dla okresu poprzedzającego chwilę bieżącą we wszystkich warstwach;
2. Wykorzystanie obliczonych w p. 1 współczynników liczbowych modeli regresyjnych do wyznaczania prognozowanych przebiegów zmiennej wyjściowej w każdej warstwie sieci;
3. Obliczenie rozbieżności pomiędzy prognozowanymi i rzeczywistymi wartościami zmiennej wyjściowej wg ustalonego kryte-

rium dla każdej warstwy (w warunkach dostępu do danych historycznych – to wszakże faza uczenia sieci).

4. Wybór złożoności modelu (warstwy) do wykonania prognozy – wybrana jest struktura modelu, który w kilku ostatnich krokach daje najlepsze rezultaty wg kryterium z punktu 3 i której złożoność jest akceptowalna w aspekcie czasochłonności obliczeń (nie jest zbyt duża);

### **Faza prognozy**

5. Na podstawie wybranego w punkcie 4 modelu wykonywana jest prognoza i ta prognoza jest wdrażana w ewentualnej strategii inwestycyjnej;
6. Podobnie jak w p.1 i 2 wykonywana jest prognoza w innych warstwach;
7. Po upływie stosowanego horyzontu prognostycznego obliczania jest dokładność wszystkich modeli;
8. Struktura modelu wyselekcjonowanego na podstawie p.4 jest weryfikowana na podstawie p. 7 (lub pozostawiona bez zmian).

W opisanym wyżej algorytmie występuje szereg parametrów różnicujących jego przebieg i wyniki. Do najważniejszych należałoby zaliczyć:

- długość wektorów zmiennych wejściowych, branych pod uwagę przy tworzeniu modelu regresyjnego;
- wielkość horyzontu predykcyjnego;
- rodzaje kryterium rozbieżności wykorzystywanego w p.3 algorytmu;
- wielkość kryterium selekcji CR (innego dla każdej warstwy i innego dla każdego zastosowanego kryterium);
- liczba modeli warstwy niższej zaliczona w wyniku selekcji do warstwy wyższej (tzn. ile spośród najlepszych modeli w danej warstwie będzie rozpatrywanych jako wejścia do warstwy wyższej; oznacza to praktyczne przeniesienie zasady ewolucji z otaczającego nas fizycznego świata do rozpatrywanego tu algorytmu);
- liczba warstw sieci od p.1 algorytmu poczynając.

Zmienna i ważna może być także sama przestrzeń zmiennych wejściowych, a także sposób ich przetworzenia (preprocessingu) zastosowany do przygotowania bezpośrednich wejść  $X_i$  (w tym także z uwzględnieniem wielkości przesunięć wstecz dla poszczególnych wektorów zmiennych wejściowych).

W pierwszej kolejności przeprowadzono badania dla różnych kryteriów z pkt.3.

Rozważano algorytm A, w którym w fazie uczenia stosowane były następujące cechy charakterystyczne:

- kryterium:

$$\Delta S_1 = \frac{\sum_{k=i-L_w}^{k=i} (y_k - \tilde{y}_k)^2}{\sum_{k=i-L_w}^{k=i} (y_k - \bar{y}_k)^2} \quad (15)$$

- $L_w = 30$  - taką długość wektora (-ów) zmiennych wejściowych przyjęto jako zasadniczą, (po kilku testach dla innych wartości  $L_w$ ).
- $h_p = 2$  - taką długość horyzontu predykcyjnego przyjęto, także po próbach nieco dłuższych okresów prognostycznych.

W charakterze zmiennej obserwowanej  $y$  przyjęto względną wartość w każdej warstwie (oznaczaną wcześniej  $y_w$ ). Pod znakiem sumy znajdują się wartości  $y$  od pozycji  $L_w$  wstecz od bieżącej  $i$ -tej pozycji aż do końca dostępnej informacji, czyli do chwili  $i$ -tej (bieżącej).

Następnie rozważany był algorytm B, w którym kryterium przyjęto postać:

$$\Delta S_2 = \frac{\sum_{k=i-h_p}^{k=i} (y_k - \tilde{y}_k)^2}{\sum_{k=i-h_p}^{k=1} (y_k - \bar{y}_k)^2} \quad (16)$$

$$L_w = 6;$$

$$h_p = 2;$$

Te parametry mają to samo znaczenie co w poprzednim algorytmie A. Zwrócić tu należy uwagę na koncentrację uwagi na ostatnich  $h_p$  okresach, zamiast na całości modelowanego odcinka  $L_w$ .

Trzecią z kolei odmianą algorytmu oznaczonego C wyróżniało kryterium minimaksowe:

$$\Delta S_3 = \max_k |y_k - \tilde{y}_k| \quad \text{dla } k=i-L_w, \dots, i \quad (17)$$

Stosowano tu parametry:

$$L_w = 20$$

$$h_p = 2;$$

Podsumowując, przeprowadzone zostaną badania trzech odmian algorytmu CL2 (tab. 1).

Tab. 1. Plan eksperymentów algorytmu CL2.

Odmiana algorytmu CL2	Kryterium	Długość wektora danych $L_w$	Horyzont predykcyjny $h_p$
CL2 A	$\Delta S_1$	30	2
CL2 B	$\Delta S_2$	6	2
CL2 C	$\Delta S_3$	20	2

Powyższe parametry ( $L_w$ ,  $h_p$ ) nie były jedynymi badanymi, jednak te wymienione w tabeli są parametrami, wg których wykonano zaprezentowane dalej symulacje.

Przeprowadzono badania wszystkich odmian algorytmu CL2 w przestrzeni danych świec 15 minut pary (EUR/USD) z 2006 roku (od świecy nr 1000 poczynając), co 15 minut (czyli co jedną świecę). W każdym kroku wykonywano predykcje na  $h_p$  świec wprzód wykorzystując historię  $L_w$  ostatnich wartości zmiennych wejściowych  $X$ . Nie był to jednak odcinek ostatnich  $L_w$  świec, gdyż wiele spośród wybieranych do modeli regresyjnych zmiennych było przesuniętych o różne okresy wstecz.

Badania takie powtarzane były w różnych miejscach przestrzeni danych w obrębie całego 2006 roku i wszędzie uzyskiwano zbliżony do przedstawionego w kolejnym rozdziale obraz skuteczności algorytmu.

### 2.3. Badania algorytmu z kryterium rozbieżności opartym na całym obserwowanym wektorze zmiennych wejściowych

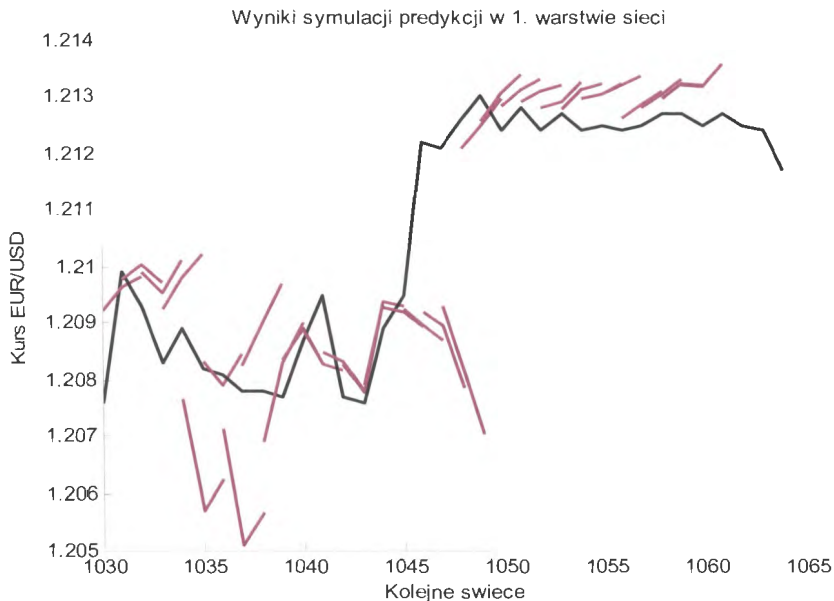
Kryterium rozbieżności pomiędzy rzeczywistym i modelowanym przebiegiem zmiennej obserwowanej wg (15) nakazuje sprawdzenie zbieżności modelu  $\hat{y}$  oraz rzeczywistych zaobserwowanych wartości  $y$  na całym odcinku o długości  $L_w$  okresów (świec) od chwili bieżącej wstecz.

Badania przeprowadzono dla okresu predykcyjnego wprzód  $h_p=2$ . Na niektórych wykresach przedstawiono jednak tylko jeden (pierwszy) odcinek predykcyjny. Uczyniono to dla uzyskania większej przejrzystości wykresów, a także z powodu gorszych wyników prognozy na tym drugim odcinku.

Na kolejnych wykresach przedstawiono proponowane w kolejnych warstwach sieci obliczeniowej odcinki zmiennej wyjściowej złożone z dwóch okresów predykcyjnych nakładane na rzeczywisty przebieg obserwowanej wartości  $y$ . Wszystkie wykresy sporządzone są po skonwertowaniu wyników ze zmiennych znormalizowanych do rzeczywistych. Taka wizualizacja wyników pozwala na natychmiastowe odniesienie się do przydatności metody w praktycznym tradingu.

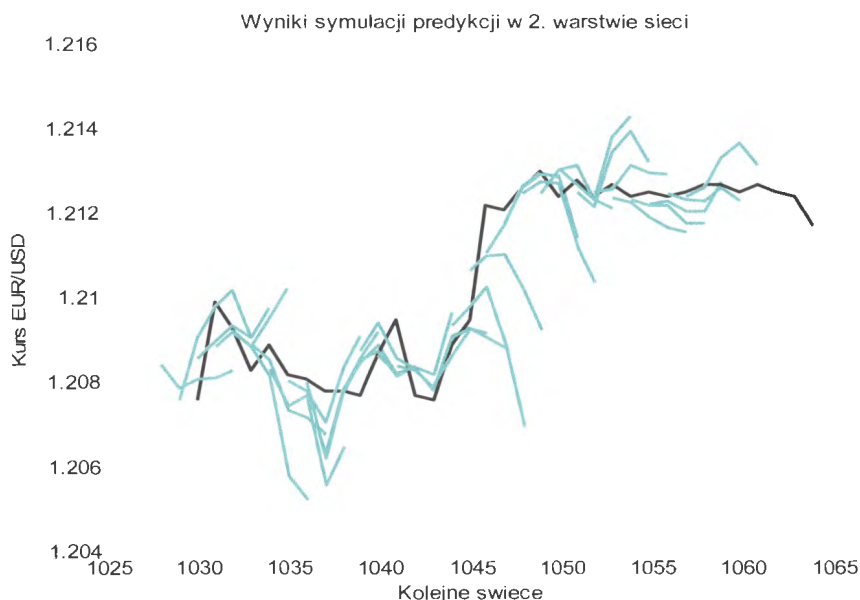
Mając bowiem wiedzę o ograniczeniach wynikających z mechanizmów transakcji stosowanych przez platformy brokerskie, w tym także z marginesem prowizyjnym (spreadem), można niemal natychmiast odrzucić metodę, gdyby błąd predykcji był większy od naturalnej zmienności zmiennej obserwowanej.

Na wykresach przedstawiono najlepsze symulacje prognozy w kolejnych warstwach sieci neuronowej. Na obu wykresach 27 i 28 widać dużą rozbieżność pomiędzy zmienną wyjściową (czarny kolor, bez nieciągłości) a kolejno nakładanymi odcinkami predykcyjnymi (dwuokresowymi na rys. 27 i czterookresowymi na rys. 28).



Rys. 27. Rezultaty prognozowania przebiegu zmiennej EUR/USD wykonanego za pomocą najprostszego modelu liniowego zawierającego tylko dwa najlepsze wyselekcjonowane czynniki modelu (zmienne wejściowe).





Rys. 28. Rezultaty modelowania w drugiej warstwie sieci neuronowej (za pomocą modeli zawierających 4 wyselekcjonowane czynniki). Rzeczywisty przebieg zmiennej wyjściowej to krzywa bez nieciągłości.

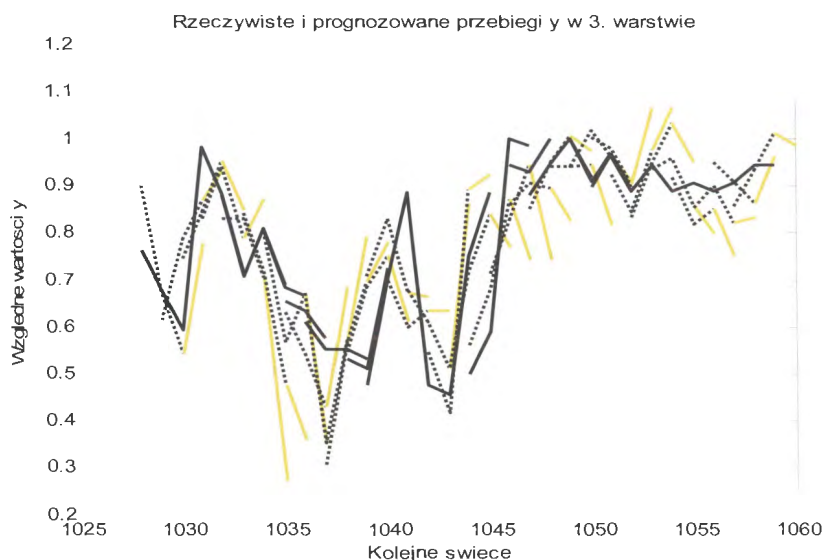
Odcinki te generowane są co jeden krok, tak więc w rozwiniętej fazie prognozowania może wystąpić równocześnie wiele modeli predykcyjnych – każdy z nich w innej odległości od początku prognozy. Być może jest to okoliczność umożliwiająca stworzenie innego algorytmu – z wzajemnie potwierdzającymi się prognozami. Na razie warto zauważyć, że w dziedzinie rzeczywistych zmian wartości funkcji obserwowanej (kursu EUR/USD) przedstawione przebiegi predykcyjne w pierwszej warstwie mogą pojawić się nawet w dużym oddaleniu od rzeczywistego przebiegu  $y$ , a w drugiej warstwie leżą wprawdzie blisko, lecz niekoniecznie pozwolą na wyciąganie praktycznych wniosków do zastosowania w strategii tradingowej.

Z tych dwu dotychczasowych obserwacji można wyciągnąć wstępny wniosek, że interesująca byłaby obserwacja pierwszej pochodnej tych prognoz. Ta zmienna ma ogromne praktyczne znaczenie – informuje bowiem, czy przewidywany jest wzrost, czy spadek kursu. Możliwość prognozowania tych zmian ma w istocie „strategiczne” znaczenie.

W kolejnej trzeciej warstwie obraz zgodności pomiędzy modelami prognostycznymi, a rzeczywistym przebiegiem (czarne krzywe na rys. 29) jest zgoła inny. Wykres ten przedstawiono dla odmiany w dzie-

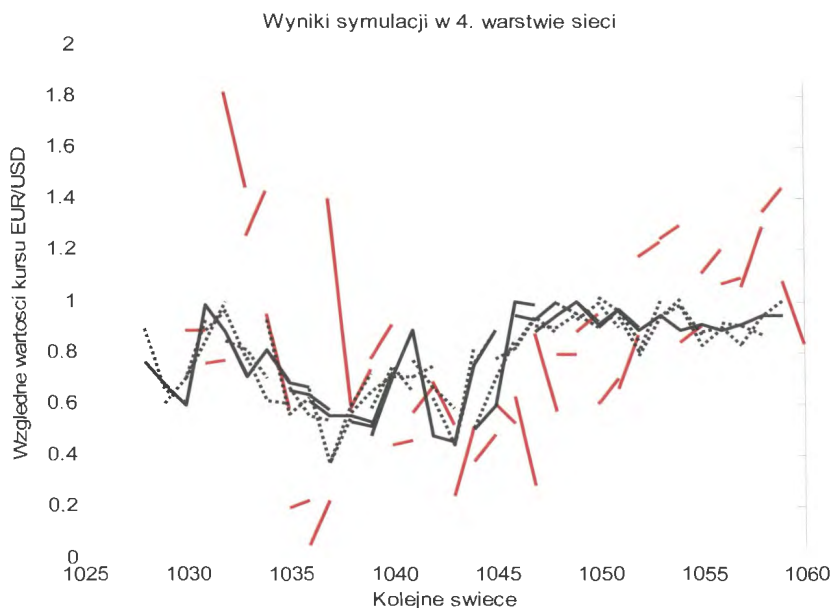
dzinie względnych wartości kursu. Stąd krzywa łamana jest wielokrotnie przerywana, gdyż dla każdego prognozowanego odcinka tworzony jest odrębny układ odniesienia i w nim umieszczone są przebiegi rzeczywistych i prognozowanych zmiennych.

Na tym wykresie widać już częstą równoległość odcinków predykcyjnych i rzeczywistych. W większości przypadków ich pochodne mają ten sam znak. Widać wobec tego perspektywę ich wykorzystania w stosunkowo prostym systemie tradingowym.



Rys. 29. Rezultaty modelowania w trzeciej warstwie sieci neuronowej. Rzeczywisty przebieg zmiennej wyjściowej – to czarna ciągła krzywa, aproksymacja kilku okresów – krzywe przerywane, prognoza – krzywe żółte.

Na kolejnym wykresie przedstawiono wyniki badań w czwartej warstwie. Tu widać także wspomnianą równoległość odcinków predykcyjnych. To dość charakterystyczne wykresy umieszczone w przestrzeni względnych wartości zmiennej  $y$ . Widać tu wyraźnie znaczenie pierwszej pochodnej progностycznych odcinków – nie jest istotna klasycznie rozumiana dokładność odtworzenia przebiegu krzywej, a trafność przewidywania kierunku zmian.



Rys. 30. Rezultaty modelowania w czwartej warstwie. Odcinki ciągłe to rzeczywiste przebiegi, przerywane to rezultaty najlepszych aproksymacji, pojedyncze oddalone odcinki, to predykcje jeden odcinek wprzód.

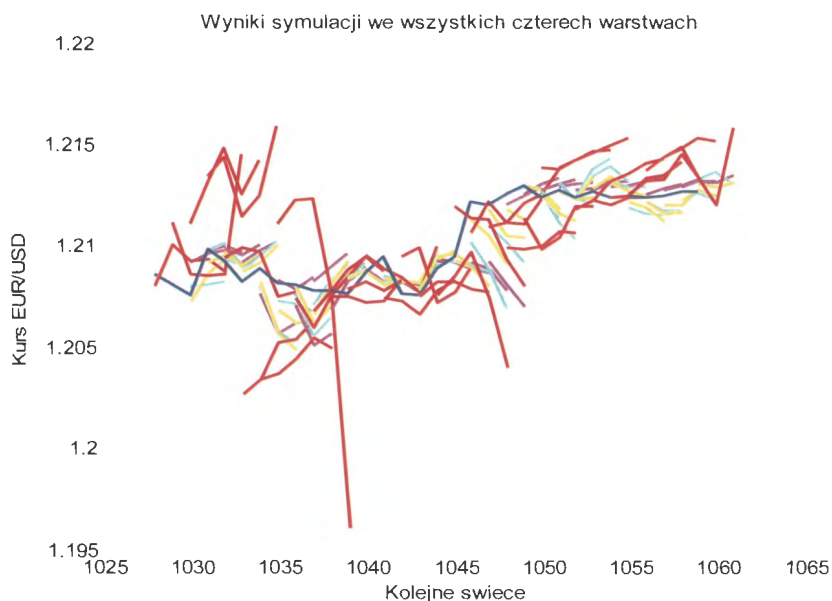
Na wykresie 30 występują trzy rodzaje krzywych łamanych. Ciągłe – to rzeczywiste przebiegi zmiennej obserwowanej (tu kursu EUR/USD skonwertowanego do wartości względnych). Przebiegi te są umieszczane w kolejnych układach wartości względnych, stąd ich wzajemne przesunięcia. Podobnie jak w przypadku prognozy w trzeciej warstwie, także i tu widać znaczenie pierwszej pochodnej prognozowanych odcinków. Na wykresie umieszczono prognozowany przebieg tylko pierwszego z dwu okresów  $h_p$ . W wyniku dalszych analiz okazało się, że trafność prognozy w pierwszym odcinku znacznie przekraczała trafność w kolejnym – stąd prezentacja tylko pierwszego z nich. Sposób obliczenia prognozy i przygotowania jej implementacji będzie przedstawiony nieco później.

Na wykresie rys. 31 przedstawiono prognozy wykonane w różnych warstwach (rozpatrywano cztery pierwsze) wzajemnie nałożone na rzeczywisty przebieg zmiennej obserwowanej  $y$ . Widać tu „rozmycie” prognoz pozostających jednak w pobliżu wartości rzeczywiście obserwowanych. Wykres ten, w odróżnieniu od dwóch poprzednich pozwala na obserwowanie prognoz w rzeczywistym oddaleniu od poprawnego

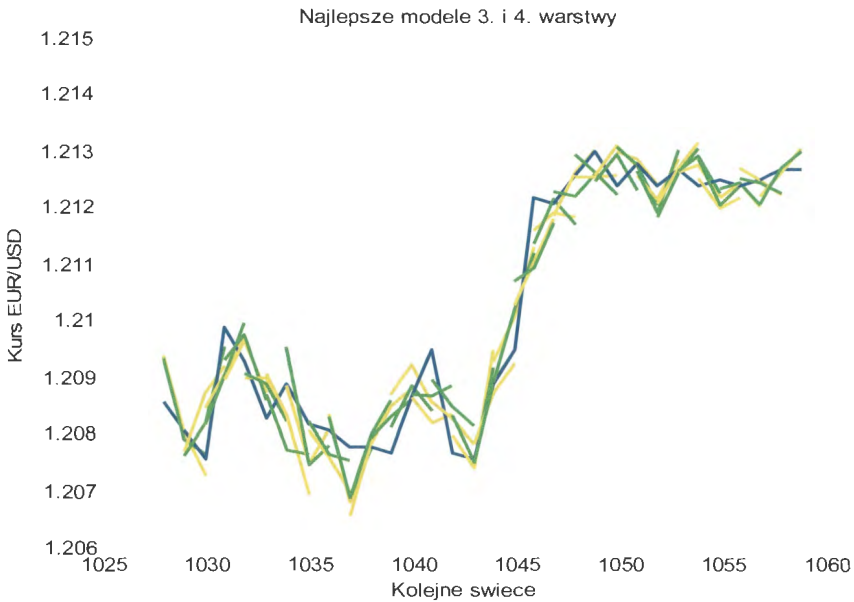
wyniku. Pozwala więc na wstępną intuicyjną ocenę możliwości zysków i strat.

Gdyby wokół wszystkich krzywych wykresu (rys. 31) zamieścić ich obwiednię, to po odrzuceniu szczególnie skrajnych przypadków powstałby korytarz o szerokości rzędu 20-30 pipsów. Nie jest to najgorszy wynik biorąc pod uwagę brak sprecyzowanej w tej chwili strategii prognostycznej. Podsumowując ten syntetyczny obraz działania całej sieci, można powiedzieć, że jest logiczny i stwarza przesłanki efektywnego wykorzystania tych wyników.

Podobnie obiecujący rezultat widoczny jest na rys. 32. Umieszczono tu najlepsze dwuodcinkowe modele trzeciej i czwartej warstwy. Dwuodcinkowe – oznacza, że na wykres naniesiono tylko dwa ostatnie odcinki modelu aproksymującego. Są to ostatnie odcinki modeli aproksymacyjnych, a więc odtwarzających przebieg zmiennej  $y$ , nie zaś modeli prognostycznych (jak na rys. 31). Warto jednak zwrócić uwagę, że tym razem szerokość umownej wstęgi to zaledwie ok. 5-10 pipsów.



Rys. 31. Wyniki symulacji w czterech warstwach sieci naniesione równocześnie na wspólny wykres.

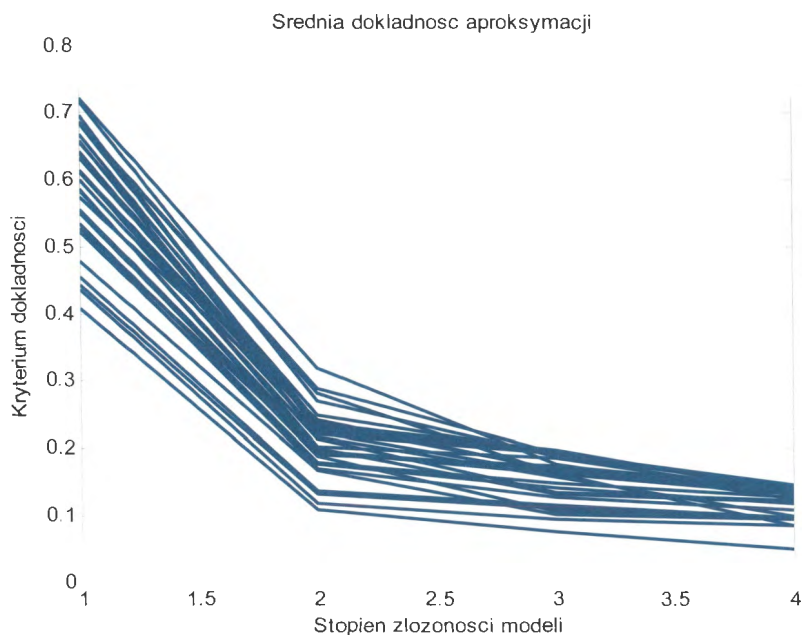


Rys. 32. Najlepsze dwuodcinkowe przebiegi modeli trzeciej i czwartej warstwy.

Nawiązując do dokładności aproksymacji określonej wcześniej dla algorytmu CL1 i przedstawionej na wykresie 23, tu zbadano wg tego samego kryterium (13) dokładność w rozpatrywanych czterech warstwach. Na rys. 32 widać dużą regularność przebiegu krzywych  $\bar{\Delta}S_u(S)$ , a więc średnich odchyłek na odcinku predykcyjnym o długości  $h_p$  w przestrzeni względnych wartości  $y_w$  wyrażonych w funkcji złożoności  $S$ .  $S$  zmienia się tu od 1 do 4 (warstwy sieci) a  $\bar{\Delta}S_u$  – średnie odchylenie w najlepszych modelach warstwy pierwszej, zmienia się od 0.4-0.7. Ivakhnenko [50] i Mueller, Lemke [80] za dobrą dokładność (miarę rozbieżności) uważali wynik 0.1-0.3. Wyniki takie uzyskiwane są już od drugiej warstwy i systematycznie i przewidywalnie zmieniają się (maleją) do warstwy czwartej.

Każda krzywa na wykresie 32 ilustruje dokładność w poszczególnych warstwach (od 1 do 4) dla innego punktu początkowego w przestrzeni danych.

Zgodnie z (12) i (13) średnia ta obliczana jest dla  $H$  najlepszych wyników spośród wyselekcjonowanych modeli o kryterium dokładności  $CR$  mniejszym od ustalonego progu  $CR_s^b$ ,  $S=1, 2, \dots, 4$  (przedtem, w CL2 badano 6 warstw).



Rys. 33. Średnia dokładność aproksymacji w poszczególnych warstwach sieci (1 do 4). Każda krzywa to średnia dokładność najlepszych modeli rozpoczynających się w różnych punktach przestrzeni danych.

W przedstawianych tu badaniach wektor  $CR_s^b$ ,  $s=1 \dots 4$  zastosowano, po kilku próbach następujące progi selekcji:

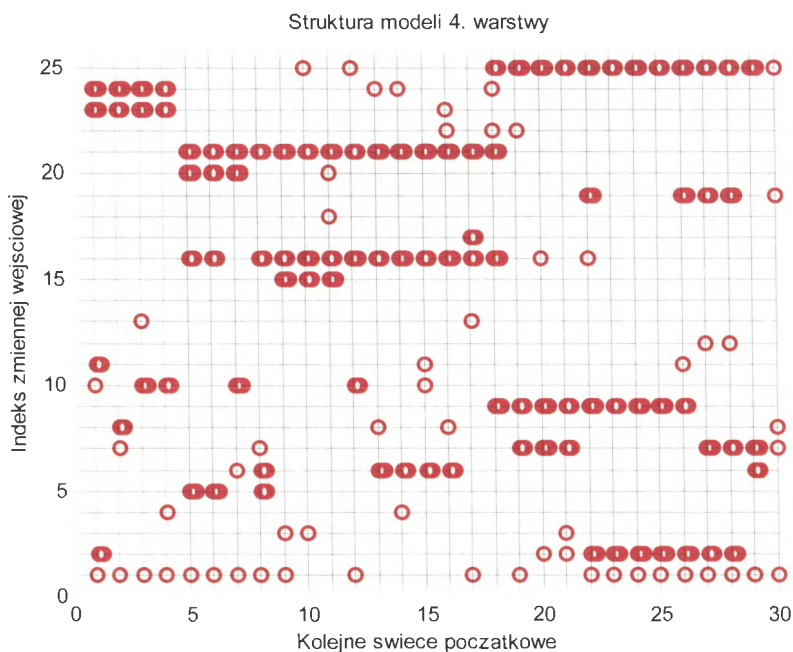
$$CR^b = [4.0 \ 1.0 \ 0.2 \ 0.15] \quad (18)$$

Między wcześniej określonym  $CR^a$  i określonym teraz  $CR^b$  występuje różnica w sposobie obliczania kryterium dokładności.  $CR^a$  i  $\Delta S^b$  dotyczą rozbieżności obliczanej na prognostycznym odcinku porównywanych wektorów (o długości  $hp$ ), a  $CR^b$  i  $\Delta S^b$  dotyczą rozbieżności obliczanej na całej długości wektorów  $y$  i  $\hat{y}$  (i dalszych  $\hat{z}$ ,  $\hat{v}$  i  $\hat{w}$ ).

Z wektora progów selekcji (18) i wyników na wykresie 31 wynika, że dla poprawienia jakości modeli można byłoby obniżyć wartości  $CR^b$  w pierwszych dwóch warstwach. Na wykresie widać jak wszystkie krzywe średniej dokładności mieszczą się tuż poniżej progu  $CR_3^b = 0.20$  (dla 3 warstwy) i  $CR_4^b = 0.15$  (dla czwartej warstwy) i przebiegają z dużym zapasem poniżej  $CR_1^b = 4.0$  i  $CR_2^b = 1.0$ .

Lepszymi ograniczeniami byłyby więc np. wartości  $CR_1^h=0.75$  i  $CR_2^h=0.35$ .

Bardzo ważnym rezultatem badań jest stwierdzenie zmienności struktury modeli sąsiadujących z sobą w czasie. Stwierdzono, że każde dwa modele sąsiadujące w sensie rozpatrywanej chwili bieżącej są różne, i to różne w każdej z czterech rozpatrywanych warstw.



Rys. 34. Diagram ilustrujący zmienność struktury modeli w czwartej warstwie.

W aspekcie rozpatrywanej tu prognozy to informacja niekorzystna. Najważniejszym bowiem założeniem przeprowadzonej prognozy jest przypuszczenie, że przed chwilą obliczony optymalny model aproksymacyjny, jeszcze przez chwilę (przez jakiś czas wprzód) nie zmieni swojej struktury i po wstawieniu doń nowych danych udanie „przedłuży” dotychczasowy przebieg obserwowanej zmiennej  $y$ .

Na rys. 34 widać ogromne zróżnicowanie modeli. Przez model należy tu rozumieć wektor zawierający indeksy zmiennych wejściowych ( $X_i$ ), które znaleźć można na każdej pionowej, oznaczającej kolejny rozpatrywany w szeregu czasowym przypadek. Zmienne oznaczone pogrubionym markerem sygnalizują wielokrotne użycie danej zmiennej

w modelach warstw pośrednich składających się na model końcowy warstwy wyższej (na wykresie – warstwy czwartej).

Obraz tego zróżnicowania modeli nie pozbawia sensu stawianej tu tezy o wykonaniu prognozy na podstawie najlepszego bieżącego modelu. Przedstawione tu są bowiem tylko najlepsze modele (czwartej warstwy) wg określonego kryterium (na rynku – wg minimum (15)). Tymczasem modele te wybrane są ze zbioru wielu modeli o dużym często zróżnicowaniu strukturalnym, ale niewielkim zróżnicowaniu, jeżeli chodzi o dokładność predykcji. Sytuacja taka stwarza szansę, że zastosowanie danego najlepszego modelu pozwoli w kolejnym kroku (lub kilku kolejnych) dokonać udanej prognozy, mimo że będzie tam, ten jeden krok wprzód, obowiązywał już zupełnie inny model.

Teza ta możliwa jest do weryfikacji jedynie w konkretnej rozpatrywanej przestrzeni danych. Należy więc zbadać jak w tej przestrzeni prognoza, o której później, rozmija się, lub nie, z rzeczywistym przebiegiem krzywej. Wprowadzono następujące wskaźniki jakości prognozy:

– **Zwrot** (w literaturze przedmiotu często używa się terminu „stopa zwrotu” w zbliżonym kontekście) po jednym okresie, rozumiany jako różnica pomiędzy wartościami  $y_{L+1}$  a wartością  $y_L$  (znajdącą się na końcu wektora  $L_w$ ). Zwrot należy dosłownie rozumieć jako zwrot z zainwestowanego kapitału. Jeżeli bowiem inwestor postawił na wzrost kursu (otworzył pozycję długą), to przyrost zmiennej obserwowanej oznacza przyrost (w najprostszym przypadku liniowy) zainwestowanego kapitału. Oczywiście zwrot z kapitału może mieć, niestety, także wartość ujemną. W rozpatrywanym zadaniu będzie to:

$$dy = y_{L+1} - y_L \quad (19)$$

gdzie:

$y_L$  – ostatnia dostępna wartość zmiennej obserwowanej (czyli obserwowana w chwili bieżącej), znajdująca się na końcu wektora wartość obserwowana o długości  $L_w$ ;

$y_{L+1}$  – pierwsza predykcyjna wartość zmiennej  $y$ , znajdująca się krok wprzód po zmiennej  $y_L$ ; jest ona dostępna dopiero po upływie czasu potrzebnego do wykonania kroku wprzód (lub pobrana jest z danych historycznych).

– **Zwrot modelowany** – podobnie dla rozpatrywanego modelu danej warstwy np.  $\hat{y}$ ,  $z$ ,  $v$ , w itd.

$$d\hat{y} = \hat{y}_{L+1} - \hat{y}_L \quad (20)$$



gdzie:

$\hat{y}_L$  – odpowiednia ostatnia wartość wektora  $\hat{y}$  o długości  $L_w$ ;

$\hat{y}_{L+1}$  – pierwsza symulowana w danym modelu  $\hat{y}$  wartość predycyjna.

– **Różnica zwrotów (błąd predykcji):**

$$e_{\hat{y}} = | dy - d\hat{y} | \quad (21)$$

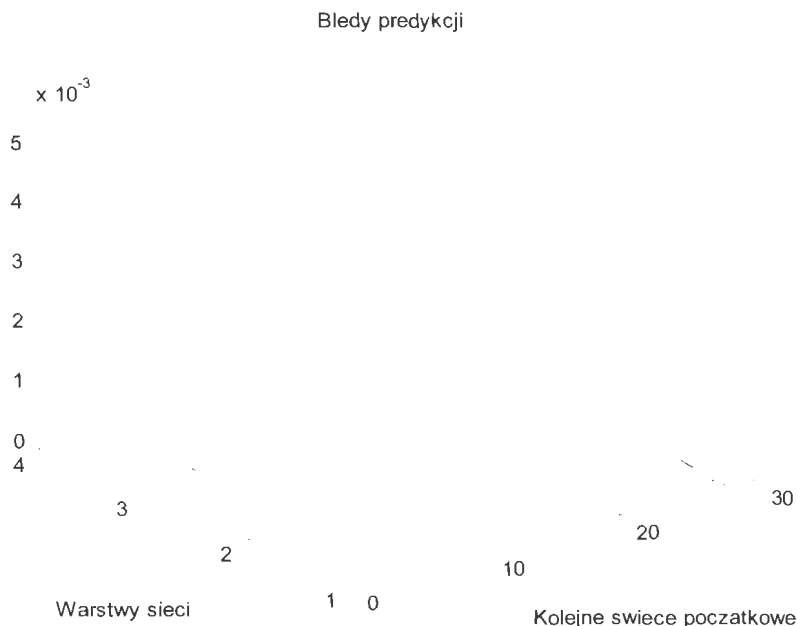
– **Znak błędu predykcji:**

$$s_{\hat{y}} = \begin{cases} 1 & \text{gdy } \text{sign}(dy) = \text{sign}(d\hat{y}) \\ 0 & \text{w przeciwnym wypadku} \end{cases} \quad (21)$$

Znak błędu predykcji ma duże praktyczne znaczenie dla strategii inwestycyjnej. Jeżeli znak rzeczywistego odchylenia (stopy zwrotu) jest taki sam jak znak odchylenia (stopy zwrotu) symulowanego przez model  $\hat{y}$ , to ewentualny błąd  $e_{\hat{y}}$  wg (21) nie niesie dużego zagrożenia. Z punktu widzenia zarządzania ryzykiem efekt tego błędu sprowadza się do konstatacji – mieliśmy wygrać to, co wynika z  $\hat{y}$ , a wygraliśmy to, co wynika z  $y$ . Wygraliśmy mniej lub więcej, ale wygraliśmy.

Gorzej, gdy znak (22) jest ujemny. Oznacza to, że model sugerował np. pozycję długą, a kurs w prognozowanym okresie obniżył swoją wartość. To porażka, której wielkość zależy od różnicy (21) i tu niestety ma znaczenie fakt, czy przegrana jest duża czy mała. By ją ograniczyć, będzie konieczne opracowanie specjalnych technik zawierających się w formułach strategii inwestycyjnej.

Na rys. 35 przedstawione są błędy popełniane w poszczególnych warstwach sieci obliczeniowej dla  $S=1, 2, 3, 4$  oraz dla kolejno wykonywanych kroków w przestrzeni danych, dla  $L_w=30$  (długość wektora danych branych pod uwagę) oraz przy horyzoncie predycyjnym  $h_p=2$ , przy czym błąd wg (21) mierzony jest tylko jeden krok wprzód. Widać duże zróżnicowanie błędów, w skrajnych przypadkach osiągających wartość kilkudziesięciu pipsów ( $1 \cdot 10^{-3}$  to 10 pipsów). Występuje jednak także wiele błędów bliskich zera. Spośród wszystkich błędów zagrożenie stanowią wyłącznie te, dla których prognoza ma znak przeciwny do rzeczywistej wartości zmiennej  $y$ , czyli te, dla których wyrażenie (21) przyjmuje wartość 0.

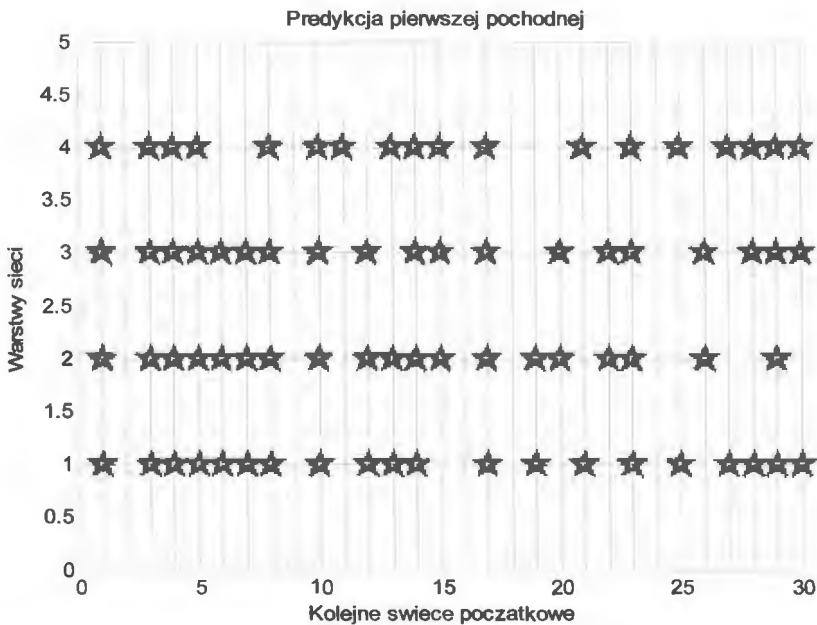


Rys. 35. Wykres błędów predykcyjnych popełnionych w poszczególnych warstwach sieci (wartości absolutne).

Na rys. 36 przedstawiono wyniki badań częstotliwości występowania zgodności znaku (zgodność pierwszej pochodnej)  $\hat{y}$  i  $y$  zgodnie z określeniem znaku wg (22). Zmienna  $s\hat{y}$  przyjmująca wartość 1 oznaczona jest gwiazdką. Puste miejsce na diagramie oznacza wartość 0 tej zmiennej.

W poszczególnych warstwach, na 30 kolejnych inicjowanych co jeden krok symulacji, uzyskano następujące liczby poprawnych prognoz  $L_s = [20 \ 20 \ 19 \ 18]$ . To w przybliżeniu skuteczność poprawnej prognozy na poziomie  $2/3$ . Tym samym algorytmem sprawdzono wiele innych miejsc w przestrzeni danych uzyskując podobne rezultaty.

Znaczenie tego rezultatu już teraz można ocenić jako interesujące i obiecujące z punktu widzenia strategii inwestycyjnej. Wyjaśnienia wymaga sposób realizacji predykcji zgodny z podstawowymi założeniami rozpatrywanego algorytmu kombinatorycznego GMDH.



Rys. 36. Diagram poprawności predykcji w aspekcie zgodności pierwszej pochodnej zmiennej obserwowanej i modeli  $\hat{y}$  w każdej z 4 warstw.

W nawiązaniu do ogólnego schematu algorytmu polegającego na wykonaniu ośmiu wymienionych wcześniej kroków, już pierwszy z nich zaleca obliczenie współczynników liczbowych modelu (dla różnych struktur wynikających z kolejnych warstw sieci), które to współczynniki pozwolą na przeprowadzenie predykcji w dalszych krokach. Zalecenie to dotyczy każdej warstwy.

W pierwszej warstwie przyjmuje się, że model będzie miał prostą postać liniową:

$$\hat{y} = [1 \quad x_w^{i-h_p} \quad x_w^{j-h_p}] * A_1^T \quad (23)$$

$$[L_w \times 1] \quad [L_w \times 3] \quad [3 \times 1]$$

gdzie:

1 – jest wektorem jedynek o wymiarze  $[L_w \times 1]$  niezbędnym do wykorzystania wyrazu wolnego z wektora współczynników  $A_{1opt}^T$ ;

$x_w^{i-h_p}$  – jest wektorem i-tej zmiennej wejściowej  $x_w$  o długości  $L_w$  kończącym się  $h_p$  wyrazów przed chwilą bieżącą;

$x_w^{j-hp}$  – jest wektorem  $j$ -tej zmiennej wejściowej, takiej, że  $j \neq i$ , o długości  $L_w$ , także kończącym na  $h_p$  przed chwilą bieżącą;

$A_1^T$  – transponowany wektor współczynników liczbowych równania regresji, obliczony wcześniej jako wynik równania  $A_1^T = X_w \setminus y_w$ , gdzie  $X_w = [1 \ x_w^{i-hp} \ x_w^{j-hp}]$ , a  $y_w$  – jest wektorem rzeczywistych wartości  $y$  skonwertowanym do wartości względnych  $y_w$  tak jak w (10).

Równanie (23) wielokrotnie rozwiązywane dla różnych par zmiennych wejściowych ( $x_i, x_j$ ) pozwala na znajdowanie różnych modeli  $\hat{y}$ , opisujących lepiej lub gorzej obserwowany wektor zmiennej wyjściowej  $y$ . Każdy z tych modeli ma swoje współczynniki  $A_1^T$  (w pierwszej warstwie o postaci  $(a_0, a_1, a_2)$ ). Każdy model ma także swoje stopnie dopasowania do zmiennej obserwowanej  $y$  wg różnych kryteriów. W rozpatrywanym w tym rozdziale przypadku ten stopień dopasowania ustalany jest wg kryterium (15) dotyczącego zgodności wzdłuż całego wektora  $\hat{y}$  z przebiegiem rzeczywistym  $y$ . Wszystkie te dane i wyniki obliczeń zapisywane są w pewnej, wygodniej z punktu widzenia implementacji, strukturze (macierzy komórkowej) o postaci:

$$A_k = [K_1 \ i \ j \ A_{1ij}^T \ \hat{y}_{ij}] \quad (24)$$

gdzie:

$K_1 \in \{ \Delta s_1, \Delta s_2, \Delta s_3 \}$ , jest kryterium uogólnionym, jednym wcześniej zdefiniowanych w (15)÷(17);

$i, j$  – to indeksy zmiennych wejściowych w postaci przed transformacją do modeli w wyższych warstwach. Indeksy te mogą być licznikami iteracji w pętłach uwzględniających wszelkie możliwe kombinacje par;

$A_{ij}^T$  – wektor (transponowany) współczynników liczbowych równania regresji pierwszej warstwy, będący wynikiem rozwiązania  $A_{ij}^T = X_{wij} \setminus y_w$ , gdzie  $X_{wij}$  jest macierzą jak w (23), lecz dla rozpatrywanych w danej chwili pary zmiennych  $(i, j)$ ;

$\hat{y}_{ij}$  – model zmiennej wyjściowej będący rezultatem przetwarzania w pierwszej warstwie.

Tych modeli  $\hat{y}_{ij}$  w pierwszej warstwie jest tyle, ile w wyniku implementacji tego algorytmu zbadanych będzie par zmiennych wejściowych ( $x_i, x_j$ ). Jeżeli zbiór zmiennych wejściowych jest o liczności  $M$  to kombinacji par będzie najwyżej  $M(M-1)/2$ .

Dla dużego  $M$  może to być duża liczba w aspekcie zagrożenia dla rozsądnego czasu obliczeń, zwłaszcza w perspektywie wielokrotnego zwiększenia tego czasu w kolejnych warstwach sieci obliczeniowej.

Z tego powodu, jedną z najistotniejszych zasad GMDH, nawiązującą do idei ewolucyjnego rozwoju świata ożywionego [37, 42, 50], jest zasada wykorzystywania w następnej warstwie sieci (w następnym pokoleniu ewolucyjnym) tylko najlepszych rozwiązań poprzedniej warstwy (poprzedniego pokolenia).

W celu uporządkowania wszystkich wygenerowanych rozwiązań (modeli) pierwszej warstwy dokonano sortowania wszystkich rozwiązań (struktur)  $A_k$  wg (24) w kolejności rosnącej wg  $K_1$ , tworząc uporządkowane struktury  $B_{1opt}$ :

$$B_{1opt} = \text{sort } A_k (K_1) \quad (25)$$

Oznacza to uporządkowanie struktur  $A_k$  wg rosnącego parametru  $K_1$ . A więc na pierwszych miejscach ciągu  $B_{1opt}$  znajdują się struktury o najniższych wartościach  $K_1$ . Podobnie postępowano w kolejnych warstwach.

Struktura optymalnych wybranych w poszczególnych warstwach modeli (tablic komórkowych) – w kolejnych pierwszych trzech warstwach przedstawiała się następująco:

$$\begin{aligned} B_{1opt} &= [K_1 \ i \ j \ A_1 \ \hat{y}] \\ B_{2opt} &= [K_2 \ i \ j \ ii \ jj \ A_2 \ \hat{z}] \\ B_{3opt} &= [K_3 \ i \ j \ ii \ jj \ iii \ jjj \ A_3 \ \hat{w}] \end{aligned} \quad (26)$$

gdzie:

- $i, j$  – oznaczają indeksy zmiennych wejściowych  $i, j = 1, 2, \dots, n, i \neq j$ ;
- $ii, jj, iii, jjj$  – oznaczają indeksy zmiennych kolejnej (drugiej i trzeciej) warstwy oznaczające także kolejność w przygotowanym w poprzedniej warstwie zbiorze modeli posortowanym wg kryterium  $K$ ;
- $K_1, K_2, \dots$  – kryteria jakości poszczególnych modeli w kolejnych warstwach. Po posortowaniu modele ułożone są w porządku rosnącym wg  $K_i$ .

Tak więc  $B_{1opt}, B_{2opt}, B_{3opt}$ , to pewne struktury zawierające niemal pełną (na potrzeby algorytmu wystarczającą) informację o modelach danej warstwy. Informacja ta uporządkowana jest kolejno w formie połączonych obok siebie wektorów (lub liczb) zawierających:

$K$  – kryterium jakości modelu – najczęściej jest to wynik predykcji w formie uprzednio zdefiniowanej jako  $\Delta S_{1max}, \Delta S_{1sr}, \Delta S_{2max} \dots$

$i, j, ii, jj$  itd. – opisane wcześniej indeksy;

$A_k$  – wektory współczynników, liczbowych równań regresji w poszczególnych warstwach  $k = 1, 2 \dots$ . Wektory te mają różną długość np. w warstwie pierwszej 3 współczynniki wystarczają do przedstawienia równania regresji z dwiema zmiennymi, w warstwie drugiej 5 współ-

czynników, w trzeciej 7 itd. Liczba ta wynika z przyjętej w danym algorytmie koncepcji narastającej złożoności równań regresji w miarę wzrostu liczby warstw sieci obliczeniowej;

$\hat{y}$ ,  $z^{\wedge}$ ,  $\hat{w}$  – wektory modelowanej zmiennej wyjściowej umożliwiające ocenę jakości prognozy (w końcowym odcinku równym długości horyzontu prognostycznego lub ocenę rozbieżności na całej długości rzeczywistego i modelowanego wektora opisywanej zmiennej wyjściowej).

Najbardziej złożona strukturą w powyższym zestawieniu jest model trzeciej warstwy  $B_{3opt}$ . Przez  $B_{3opt}$  należy w zasadzie rozumieć zbiór kilku modeli wyselekcjonowanych i uporządkowanych wg rosnącego  $K_3$ , z których ten, znajdujący się na pierwszym miejscu może być np. wykorzystany do przedstawienia (wizualizacji) prognozy zaś wszystkie razem mogą być wykorzystane w kolejnej czwartej warstwie.

W konkretnej rozpatrywanej tu implementacji struktura trzeciej warstwy:

$$B_{3opt} = [K_3 \text{ i j ii jj iii jjj } A_3 \hat{w}] \quad (27)$$

proceedzi do równania regresji postaci:

$$v^{\wedge} = [1 \ x_i \ x_j \ z_{ii} \ z_{jj} \ w_{iii} \ w_{jjj}] * A_3^T \quad (28)$$

które może być elementem (ten model) kolejnego zbioru  $B_{4opt}$  w czwartej warstwie.

Wszystkie składniki  $x_i$ ,  $x_j$ ,  $z_{ii}$  ... macierzy w nawiasie kwadratowym to wektory o zadanej długości (wynikającej z konieczności uczenia modelu – w badaniach przyjmowano najczęściej  $L_w = 20, 30$  lub  $50$  wyrazów). Tak więc  $v$  (28) stawało się wektorem  $[L_w \times 1] = [L_w \times 7] \times [7 \times 1]$ , z którego najbardziej interesujące były zawsze ostatnie  $h_p$  wyrazy (horyzont predykcyjny).

Pamiętając o strukturze kolejnych modeli (26) można przedstawić sposób pozyskiwania informacji dla równania predykcyjnego (28) kolejno jak poniżej:

$x_i$  – będzie to wektor zmiennej wejściowej (surowej, nie przetworzonej) o indeksie, który określony jest w strukturze optymalnego modelu trzeciej warstwy jako  $B_{3opt}(1,2)$ . Indeks „1” w nawiasie oznacza najlepszy model wśród wszystkich  $B_{3opt}(:, :)$  a „2” oznacza drugą pozycję w strukturze (25), a więc komórkę przechowującą wartość indeksu „i” dla  $x_i$ .

$x_j$  – podobnie dla drugiej zmiennej  $x_j$ - wg  $B_{3opt}(1,3)$ .

$z_{ii}$  – modele drugiej warstwy  $z_{ii}$  i  $z_{jj}$  wprowadzone są do powyższego równania, nie jako najlepsze z drugiej warstwy, lecz jako ułożone na

pozycjach  $ii$  oraz  $jj$  w kolejności rosnącego kryterium. Stąd, by określić dane dla modelu o strukturze:

$$\hat{z}^1 = [\mathbf{1} x_{iz}^1 x_{jz}^1] * A_{1z}^1$$

należy cofnąć się do warstwy pierwszej. W niej można znaleźć indeksy zmiennych  $x_{iz}^1$  i  $x_{jz}^1$  ukryte w polach  $B_{1opt}(ii,2)$  oraz  $B_{1opt}(ii,3)$ , gdzie  $ii$  to indeks w modelu trzeciej warstwy  $B_{3opt}(1,4)$ .

$$\hat{z}^2 = [\mathbf{1} x_{iz}^2 x_{jz}^2] * A_{1z}^2$$

$z_{jj}$  – podobnie dla drugiego modelu:

dane są w polach  $B_{1opt}(jj,2)$  oraz  $B_{1opt}(jj,3)$ , gdzie  $jj$  to  $B_{3opt}(1,5)$ , a  $A_{1z}^2$  to  $B_{1opt}(jj, 4:6)$ .

$w_{iii}$  – wywołuje już zdecydowanie większą złożoność w procesie poszukiwania danych. Związany jest z modelem:

$$\hat{w}^1 = [\mathbf{1} x_{iw}^1 x_{jw}^1 z_{iw}^1 z_{jw}^1] * A_{1z}^1 \quad (29)$$

gdzie:

$$z_{iw}^1 [\mathbf{1} x_{i wz}^1 x_{j wz}^1] * A_{1 wz}^1 \quad (30)$$

$$z_{jw}^1 [\mathbf{1} x_{i wz}^2 x_{j wz}^2] * A_{1 wz}^2 \quad (31)$$

gdzie:  $A_{2w}^1 = B_{2opt}(B_{3opt}(1,6), 6:10)$

czyli są to współczynniki umieszczone na pozycjach od 6 do 10 w wierszu (modelu) sklasyfikowanym na pozycji  $B_{3opt}(1,6)$ ;

$A_{1wz}^1 = B_{1opt}(B_{2opt}(B_{3opt}(1,6), 4), 4:6)$

czyli są to współczynniki na pozycjach od 4 do 6 w wierszu (modelu) pierwszej warstwy określonym przez 4 pozycję modelu drugiej warstwy, takim z kolei, który jest wyznaczony przez szóstą wartość w najlepszym modelu trzeciej warstwy  $B_{3opt}(1,6)$ .

$A_{1wz}^2 = B_{1opt}(B_{2opt}(B_{3opt}(1,6), 5), 4:6)$

podobnie  $jw$ ., lecz model pierwszej warstwy wyznaczony jest przez 5 pozycję modelu drugiej war-

stwy określonego przez trzecią warstwę jak poprzednio.

$x_{iwz^1}$  z równania (30) to wektor danych wejściowych o podobnie pośrednio określonym indeksie. Indeks ten można określić jako pole  $B_{1opt} (iwz1, 2)$  gdzie  $iwz1 = B_{2opt} (B_{3opt} (1,6), 4)$ .

$x_{jwz^1}$  – podobnie, lecz będzie to pole  $B_{1opt} (iwz1, 3)$ .

$x_{iwz^2}$  – z równań (30) to także wektor zmiennej wejściowej w indeksie  $D_{1opt} (iwz2, 2)$  gdzie  $iwz2 = B_{2opt} ((B_{3opt} (1,6), 5)$ .

$x_{jwz^2}$  – podobnie, lecz będzie to indeks określony przez pole  $B_{1opt} (iwz2, 3)$ .

$x_{iw^1}$  – z równania (29) to wektor zmiennej wejściowej występujący w modelu nadrzędnym (28) wynikającym z (27). Indeks tej zmiennej można pośrednio określić jako  $B_{2opt} (iii, 2)$ , gdzie  $iii = D_{3opt} (1, 6)$ .

$x_{jw^1}$  – to samo, lecz będzie wykorzystane trzecie pole w modelu drugiej warstwy, czyli  $B_{2opt} (iii, 3)$ .

$w_{jjj}$  – jest ostatnim składnikiem równania (28).

Wyraz ten generuje model:

$$\hat{w}^2 = [1 \ x_{iw^2} \ x_{jw^2} \ z_{iw^2} \ z_{jw^2}] * A_{zw^2} \quad (32)$$

Model ten wywołuje podobne zależności implementacyjne jak wyżej przedstawione dla (29) z tą różnicą, że wszystkie one odnoszą się do siódmego pola najlepszego modelu trzeciej warstwy  $B_{3opt} (1, 7)$ , a nie do szóstego jak w (29).

Siódme pole wg (28) w  $B_{3opt}$  jest bowiem odpowiedzialne za przechowywanie indeksu drugiego modelu drugiej warstwy wchodzącego w skład najlepszego modelu predykcyjnego trzeciej warstwy.

Implementacja prognostycznego tego fragmentu algorytmu jest już tak złożona, że wraz z objaśnieniami dotyczącymi implementacji w czwartej warstwie przeniesiona została do załącznika do niniejszej pracy.



## 2.4. Badania algorytmu z kryterium opartym na rozbieżności w zakresie horyzontu predykcyjnego

W drugiej odmianie algorytmu CL2 B zgodnie z planem eksperymentu z tab.1 zmienione zostało kryterium. Kryterium  $\Delta S_1$  (15) zostało zmienione na kryterium  $\Delta S_2$  wg (16).

To kryterium oznacza, że rozpatrywana była zgodność wyłącznie na odcinku predykcyjnym, w danych badaniach wynoszącym  $h_p=2$ .

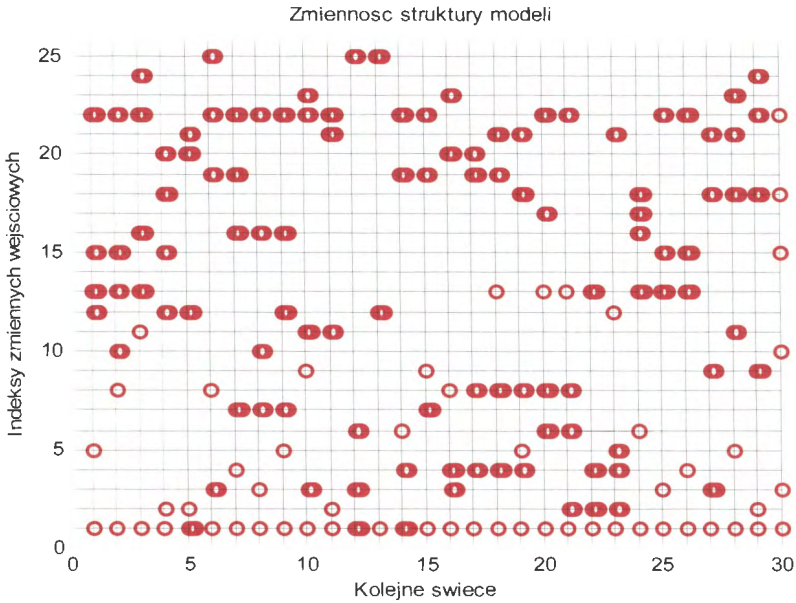
Po kilku próbach wykonywanych w różnych miejscach szeregu czasowego przedstawia się ostatecznie wyniki tego samego odcinka przestrzeni danych pary EUR/USD z 2006 roku.

W związku z tak istotną zmianą sposobu pomiaru rozbieżności (przedtem, dla przykładu, na odcinku rzędu kilkudziesięciu punktów pomiarowych, teraz tylko kilku, a w tym konkretnym przypadku tylko dwóch) celowe okazało się skrócenie odcinka (wektora) zmiennych (każdej zmiennej) wejściowych i zmiennej obserwowanej  $y$ . Zastosowano  $L_w = 6$ , czyli długość wektora zmiennych, na podstawie którego wyznaczano wektor  $A^T$  współczynników liczbowych równania regresji była równa zaledwie 6. Na tej podstawie budowano przebieg predykcyjny  $\hat{y}$ , z którego zaledwie dwa odcinki były podstawą do oceny zgodności wg (16).

Tak zdecydowane skrócenie wektorów branych pod uwagę może być interpretowane jako wykorzystanie niewielkiej liczby danych. Wielokrotnie w praktyce badawczej zdarzają się sytuacje, gdy prognozy dokonuje się na podstawie „płytkiego” zaplecza danych. W tym przypadku, gdyby rezultaty zbliżone były do poprzednich, można byłoby mówić o takim właśnie stosowaniu niewielkiej liczby danych – do udanego prognozowania.

Z przeprowadzonych rezultatów symulacji przedstawia się zmienność modeli w każdym kroku wzdłuż szeregu czasowego. W każdym kroku, czyli co jedną świecę (tu 15 minutową w szeregu EUR/USD w 2006 roku). Wyniki tej zmienności przedstawiono na rys. 37.

Podobnie jak poprzednio (na rys. 34) badana była struktura modeli w czwartej warstwie sieci obliczeniowej. Obraz zmienności przedstawiony na rys. 37 jest zdecydowanie mniej korzystny niż na rys. 34, przy czym jest to ocena powierzchowna, bez wprowadzania jakiegoś dodatkowego szczególnego kryterium. Przez zmienność modeli w kolejnych krokach symulacji można rozumieć ich skłonność do zmiany zestawu zmiennych wejściowych symbolizowanych przez punkty na kolejnych pionowych.



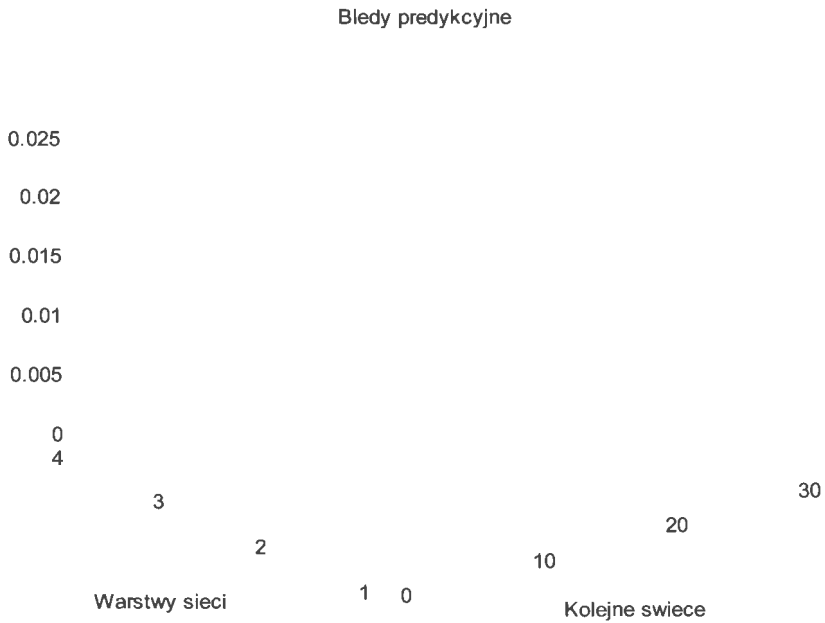
Rys. 37. Diagram zmienności struktury modeli prognostycznych czwartej warstwy w badaniach z zastosowaniem kryterium opartego na odcinku predykcyjnym.

Jeżeli więc na dwóch kolejnych pionowych, symbolizujących dwie obok siebie położone świece, widniałyby te same punkty, to oznaczałoby to, że struktura modelu nie uległa zmianie. Na rys. 37 takiej sytuacji nie ma – w każdym z 30 rozpatrywanych kroków występowała inna struktura (najlepszego modelu czwartej warstwy).

Z tego też powodu modele generowane według rozpatrywanego kryterium  $\Delta S_2$  są przez swoją zmienność mniej obiecujące w aspekcie strategii inwestycyjnej niż modele otrzymane wg  $\Delta S_1$ .

Błędy predykcyjne w rozpatrywanej tu serii badań przedstawiono na rys. 38. W porównaniu z błędami poprzednimi (z rys. 35) widać, że błędy te są większe (dochodzą do wartości 0.025) i niezbyt zróżnicowane w warstwach. To kolejny powód, dla którego można ocenić, że modele generowane wg rozpatrywanego kryterium będą gorsze.

Dla zbadania, jak poprzednio, przydatności tych modeli w strategii bazującej na zgodności znaków zmiennej rzeczywistej  $y$  i jej modelu  $\hat{y}$ , sporządzono diagram 39. Na wykresie tym, jak poprzednio na rys. 36 przedstawiono rezultaty zbieżności lub rozbieżności tych znaków.



Rys. 38. Wykres błędów predykcyjnych popełnionych w symulacjach z zastosowaniem kryterium  $\Delta S_2$ .



Rys. 39. Diagram poprawności predykcji wg zgodności pierwszych pochodnych  $y$  i  $\hat{y}$  dla kryterium  $\Delta S_2$ .

Zgodnie z dotychczasowymi ocenami wyniki przedstawione na tym diagramie także są nieco gorsze niż analogiczne dla poprzedniego kryterium. Najwyższą zgodność znaków pierwszych pochodnych uzyskano tu w pierwszej warstwie (18 na 30 kroków), w kolejnych warstwach było to 16, 17 i 17 przypadków zgodności. To raczej niewiele ponad 50%.

Podsumowując – modele wygenerowane z zastosowaniem kryterium opartego na zgodności wyłącznie odcinka predykcyjnego okazały się gorsze od modeli przygotowanych wg  $\Delta S_1$ .

Zbadano jeszcze jedno kryterium selekcji – oparte na zasadzie minimalizacji maksymalnego odchylenia w odcinku uczącym.

## 2.5. Badania algorytmu z kryterium opartym na zasadzie minimaksowej

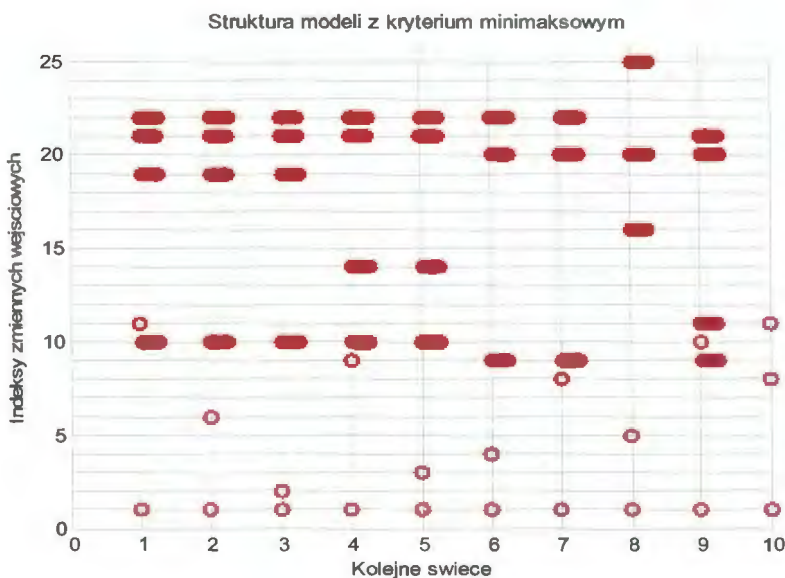
W podrozdziale tym przedstawia się wyniki badań algorytmu predykcyjnego opartego na kryterium (17):  $\Delta S_3 = \max_k |y_k - \hat{y}_k|$  dla  $k=i-L_w, \dots, i$ , gdzie  $k$  oznacza indeks każdej kolejnej wartości zmiennej rzeczywistej  $y$  i modelowanej  $\hat{y}$  wzdłuż całej długości wektora uczącego  $L_w$ . Indeks  $i$  oznacza pozycją chwili bieżącej w szeregu czasowym  $y$  i  $\hat{y}$ .

Kryterium  $\Delta S_3$  jest oczywiście minimalizowane stąd potoczne określenie tego kryterium jako minimaksowe. W zbiorze rozpatrywanych modeli określa się więc dla każdego z nich pewną cechę (maksymalne odchylenie wg (17)) i wybiera się ten model, który to odchylenie ma najmniejsze w zbiorze.

Podobnie jak dla poprzednich dwóch kryteriów w algorytmie określonym wcześniej jako CL2 C przeprowadzono badania dla 30 kolejnych szeregów czasowych, a każdym z nich badając  $L_w = 20$  wartości zmiennych wstecz od chwili bieżącej (zgodnie z tab. 1). Stosowano horyzont predykcyjny o wielkości dwóch odcinków wprzód  $h_p = 2$ . Badane szeregi czasowe następowały bezpośrednio po sobie (co jedną świecę). Dzięki temu możliwe było wyciągnięcie ewentualnych wniosków dotyczących zmienności struktury najlepszych modeli w czasie.

Wyniki tych badań przedstawiono na rys. 38. Widać tu zdecydowanie mniejszą zmienność struktur niż w poprzednio prezentowanym algorytmie CL2 B, którego odpowiednie wyniki przedstawione są na rys. 37. Tu struktura modeli, przy zauważalnej powtarzalności występowania niektórych zmiennych wejściowych w kolejnych modelach regresji sugeruje ewentualność poprawności tezy o niezmienności (przez chwilę)

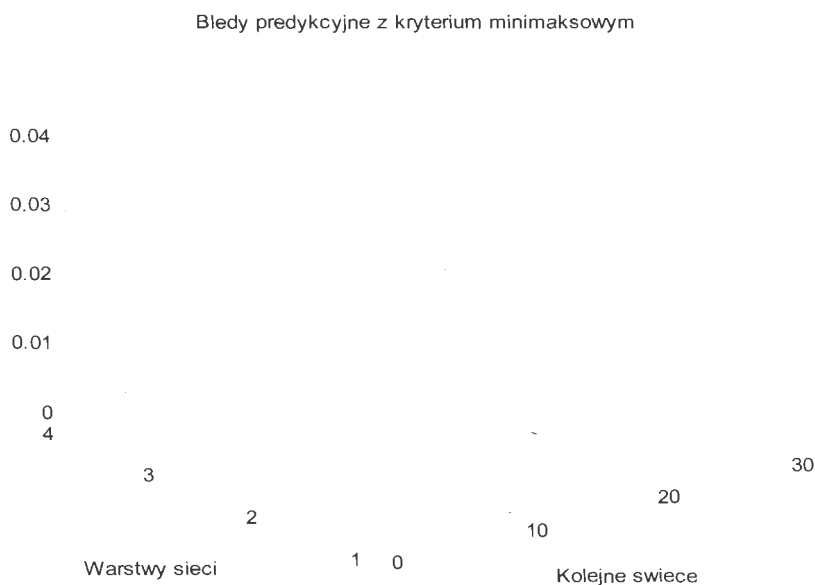
struktury modelu. Także względną stabilność struktury widać np. w szeregach czasowych od 2 do 4, gdzie modele (czwartej warstwy) praktycznie nie różnią się i korzystają zawsze z czterech tylko zmiennych wejściowych – o numerach 10, 19, 21 i 22. Z kolei w dwóch następnych krokach (szeregach czasowych) wykorzystywane są niemal te same zmienne, tylko zmienna o numerze 19 zastąpiona jest zmienną o numerze 14. Oczywiście dla badacza istotne jest co to za zmienne, gdyż dzięki temu pojawia się niejako w tle dodatkowa użyteczność metody. Staje się ona także narzędziem klasyfikacji zmiennych umożliwiając eliminację mniej przydatnych (rzadziej pojawiających się w strukturach modeli). Zmienne te będą opisane w kolejnym rozdziale pracy.



Rys. 40. Struktura najlepszych modeli z kryterium minimaxowym dla 10 kolejnych przypadków.

Przy tych założeniach (kryterium minimaxowe) błędy końcowe obliczone jako różnice pomiędzy odchyleniem po jednym odcinku wprzód (stopą zwrotu) także przedstawiają się zdecydowanie korzystniej (rys. 41) niż dla poprzedniego kryterium rozbieżności. Błędy te są wyraźnie mniejsze i przede wszystkim mniejsza jest ich wariancja (z wyjątkiem dwóch ekstremalnych zdarzeń w czwartej warstwie). Ta zgodność predykcji – ogólnie w zakresie poniżej 1% (wg kryterium (17)) oznacza możliwość zbudowania strategii, w której niewielkie ryzyko

porażki (np. przy stosunkowo dużym poziomie wartości *Stop Loss*) oznaczałyby możliwość uzyskiwania relatywnie wysokich zysków nie poprzez stosowanie mechanizmu *Take Profit*, lecz poprzez zamykanie pozycji w ustalonym czasie wygaśnięcia (np. zgodnie z horyzontem predykcji).



Rys. 41. Błędy predykcji popełnione w różnych warstwach sieci obliczeniowej dla kryterium opartego na zasadzie minimalizacji maksymalnego odchylenia w odcinku uczącym.

Podobnie jak w poprzednich badaniach sprawdzono częstotliwość występowania zgodności pierwszych pochodnych – zmiennej rzeczywistej  $y$  i modelowanej  $\hat{y}$ .

Wyniki przedstawione na rys. 42 są zaskakująco słabe, gorsze niż obserwowane w algorytmie CL2 B. Na 30 badanych szeregów czasowych uzyskano tu w poszczególnych warstwach zgodność znaków kolejno [18 18 17 12] razy.

Są to najśłabsze wyniki spośród wszystkich rozpatrywanych trzech algorytmów (kryteriów rozbieżności).

Szczególnie rzadko występuje zgodność znaków pierwszych pochodnych w czwartej warstwie. Dodatkowo o małej przydatności modeli tej warstwy (dla rozpatrywanego algorytmu CL2 C z kryterium minimaxowym) świadczą duże błędy (rys. 41). Oznacza to prawdopodobnie

większe „przeuczenie” modelu na odcinku uczącym nie przenoszące cech statystycznych na odcinek predykcyjny. Dla niższych warstw dokładność jest już tu porównywalna z poprzednimi algorytmami CL2 A i CL2 B.

Nie przekreśla to sensu stosowania tego algorytmu w ogóle, zwłaszcza przy ograniczeniu złożoności. Cechą szczególnie pozytywną są tu bowiem niewielkie błędy bezwzględne, które mogą predestynować algorytm do zastosowania przy świecach o większych odstępach czasowych niż tu rozpatrywane np. 30 minut, 1h, 4 h itd.



Rys. 42. Diagram zgodności pierwszej pochodnej przebiegu rzeczywistego i modelowanego z kryterium minimaxowym

Podsumowując rozdział poświęcony badaniom różnych algorytmów predykcyjnych z zastosowaniem różnych kryteriów rozbieżności przeprowadzono porównanie trzech rozważanych algorytmów pod kątem ich zdolności do generowania zysku skumulowanego w ewentualnym prostym systemie tradingowym.

System ten sprowadzał się do:

1. bezpośredniego otwierania pozycji długiej lub krótkiej, w zależności od znaku pierwszej pochodnej prognozy i zamykania tej pozycji po upływie okresu (jednego) badanej świecy;

2. otwierania pozycji jak w p. 1 z równoczesnym ustawieniem zabezpieczenia przed stratą w formie mechanizmu *Stop Loss* i zamykania pozycji także po upływie jednej świecy.

Każdy z algorytmów w rozważanym zakresie badań dawał określony wynik skumulowany, z którego można było obliczyć wartość średnią zwrotu (z zainwestowanego kapitału) na jedną świecę:

$$m_s = (1/k) \sum (y_{i+1} - y_i) \quad (33)$$

gdzie  $k$  – to liczba badanych świec.

Podobnie obliczając odchylenie standardowe dla badanego zbioru zwrotów:

$$s_s = \text{std} (y_{i+1} - y_i) \quad (34)$$

wprowadzono umowne kryterium oceny przydatności algorytmu dla systemu tradingowego – znormalizowaną stopę zwrotu (wskaźnik Sharpe'a nie uwzględniający zysku pozbawionego ryzyka [106]):

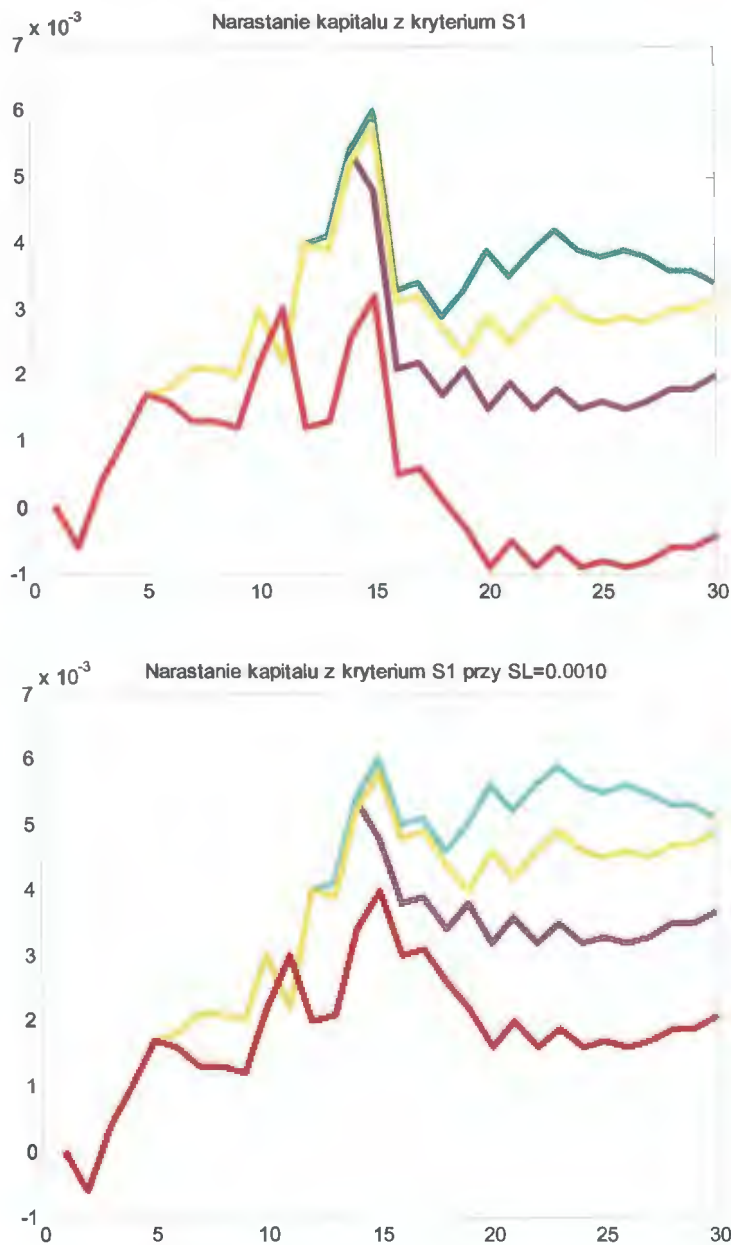
$$r_s = m_s / s_s \quad (35)$$

Wskaźnik (35) będzie jeszcze wielokrotnie używany do porównania skuteczności rozważanych strategii tradingowych.

Każdą z rozpatrywanych strategii rozróżnianych wg stosowanego kryterium rozbieżności (a więc wg algorytmu CL2 A, CL2 B i CL2 C) zbadano w symulacjach realizowanych w kolejnych, położonych obok siebie szeregach czasowych o liczności równej 30 szeregów z parametrami jak w tab. 1.

Na rys. 43 przedstawiono wykres narastania kapitału dla strategii inwestycyjnej, w której jako kryterium rozbieżności pomiędzy zmienną rzeczywistą i jej modelem ustalonym w czterech warstwach sieci GMDH zastosowano wskaźnik  $\Delta S_1$ .





Rys. 43. Wykresy narastania kapitału w strategii tradingowej (dla czterech warstw sieci obliczeniowej) zbudowanej z zastosowaniem kryterium  $\Delta S_1$  dla wariantu bez mechanizmu Stop Loss (lewy wykres) i z mechanizmem Stop Loss równym 0.0010 (10 pipsów).

Przedstawione na rys. 43 wyniki symulacji dotyczą czterech warstw sieci obliczeniowej. W każdej warstwie sieci wybierany był, w wyniku selekcji, najlepszy model wg kryterium  $\Delta S_1$  i zgodnie z rekomendacjami tej warstwy otwierana była długa lub krótka pozycja, zgodnie ze znakiem pierwszej pochodnej modelu. Zarówno na wykresie z lewej strony rys. 43 jak i z prawej strony tego rysunku krzywe skumulowanej wartości kapitału układają się dla poszczególnych warstw w tej samej kolejności – najlepsze (najwyżej położone krzywe) są krzywe zbudowane na podstawie prognoz uzyskiwanych w drugiej warstwie. Nieco gorsze są wyniki uzyskiwane na podstawie prognoz w trzeciej warstwie, następnie w pierwszej warstwie i najgorsze są wyniki uzyskiwane wg prognoz z czwartej, najbardziej złodzonej warstwy modeli.

W prawej części wykresu kolejność jest taka sama jednak wyniki końcowe skumulowanej wartości kapitału po 30 krokach są zdecydowanie lepsze. Średnie wartości znormalizowanej stopy zwrotu  $r_s$  wg (35) wynoszą odpowiednio dla poszczególnych warstw sieci dla obu wariantów systemu (z użyciem SL i bez) jak w tab. 2.

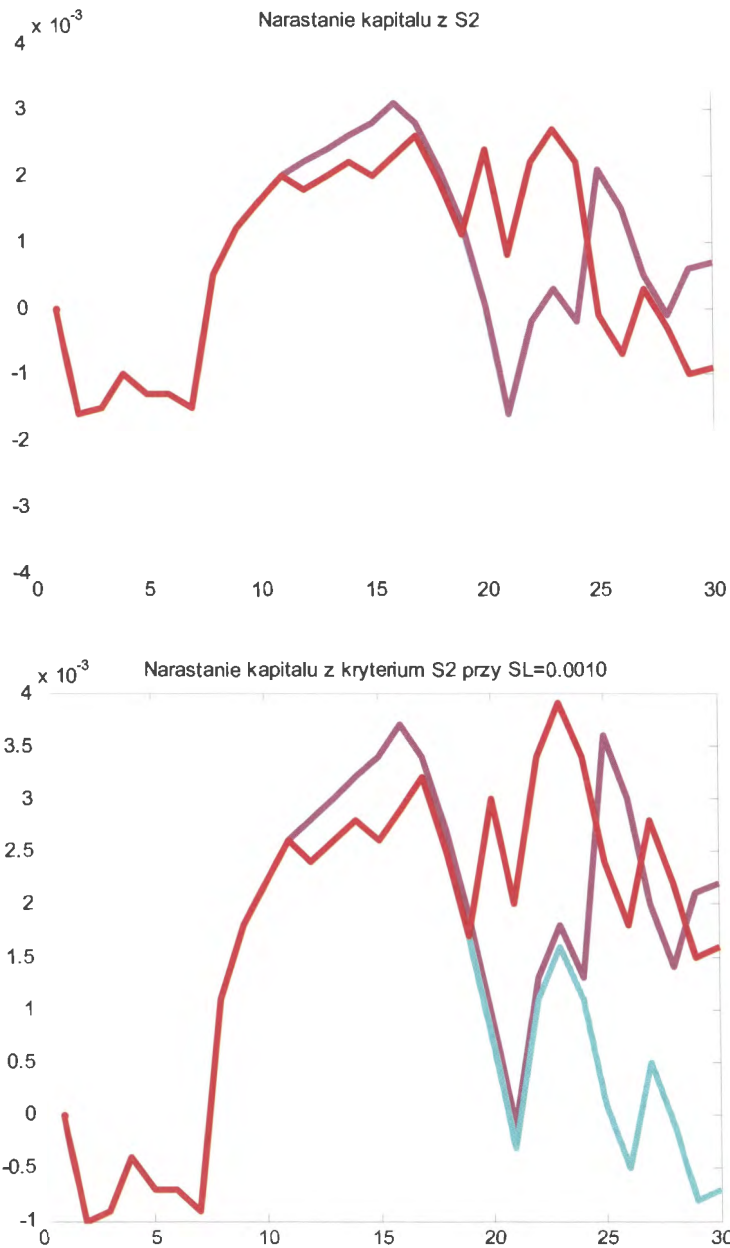
Tab. 2. Wyniki symulacji algorytmu predykcyjnego CL2 A – średnia normalizowana stopa zwrotu.

Wersja algorytmu	Warstwa 1	Warstwa 2	Warstwa 3	Warstwa 4
CL2 A bez SL	0.0865	0.1481	0.1392	-0.0172
CL2 A z SL	0.2019	0.2838	0.2718	0.1259

Przedstawione w tab. 2 wyniki dla wariantu algorytmu z użyciem Stop Loss<sup>23</sup> (drugi wiersz tabeli) otrzymane na wyjściu drugiej i trzeciej warstwy sieci mogą być już interesujące w praktyce. Patrząc na wykres (rys. 41) wyniki te, w prawej części wykresu, nie są może jednak zbyt przekonujące i konieczne byłoby zapewne rozszerzenie badań o zdecydowanie większą liczbę szeregów czasowych.

W drugim z kolei cyklu rozważań dotyczących algorytmu CL2 B otrzymano zdecydowanie gorsze wyniki. Przedstawiono je na rys. 44. Tu zarówno z lewej (bez Stop Loss) jak i z prawej otrzymano dużo gorsze wyniki niż poprzednio.

<sup>23</sup> Stop Loss – typowy mechanizm półautomatycznego zamknięcia pozycji przynoszącej straty (półautomatyczny – oznacza najczęściej „manualnie” ustawiana wartość dopuszczalnej straty i automatyczne jej wykonanie przez platformę brokerską.



Rys. 44. Wykresy narastania kapitału w strategii tradingowej (dla czterech warstw sieci obliczeniowej) zbudowanej z zastosowaniem kryterium  $\Delta S_2$  dla wariantu bez mechanizmu Stop Loss (lewy wykres) i z mechanizmem Stop Loss równym 0.0010 (10 pipsów).

## Rozdział 2. Badania algorytmów kombinatorycznych GMDH

Wyniki te przedstawiono w tab. 3. Widać tu wiele wartości ujemnych, a więc symulacji przynoszących stratę. Wskaźnik  $r_s$  w porównaniu z wynikami z tab. 3 jest tu zdecydowanie gorszy. Algorytm realizuje predykcje oparta na bardzo krótkim wektorze danych wejściowych ( $L_w = 6$ ) i na kryterium rozbieżności związanym z horyzontem predykcyjnym. Należy uznać próby tej predykcji za raczej mało obiecujące.

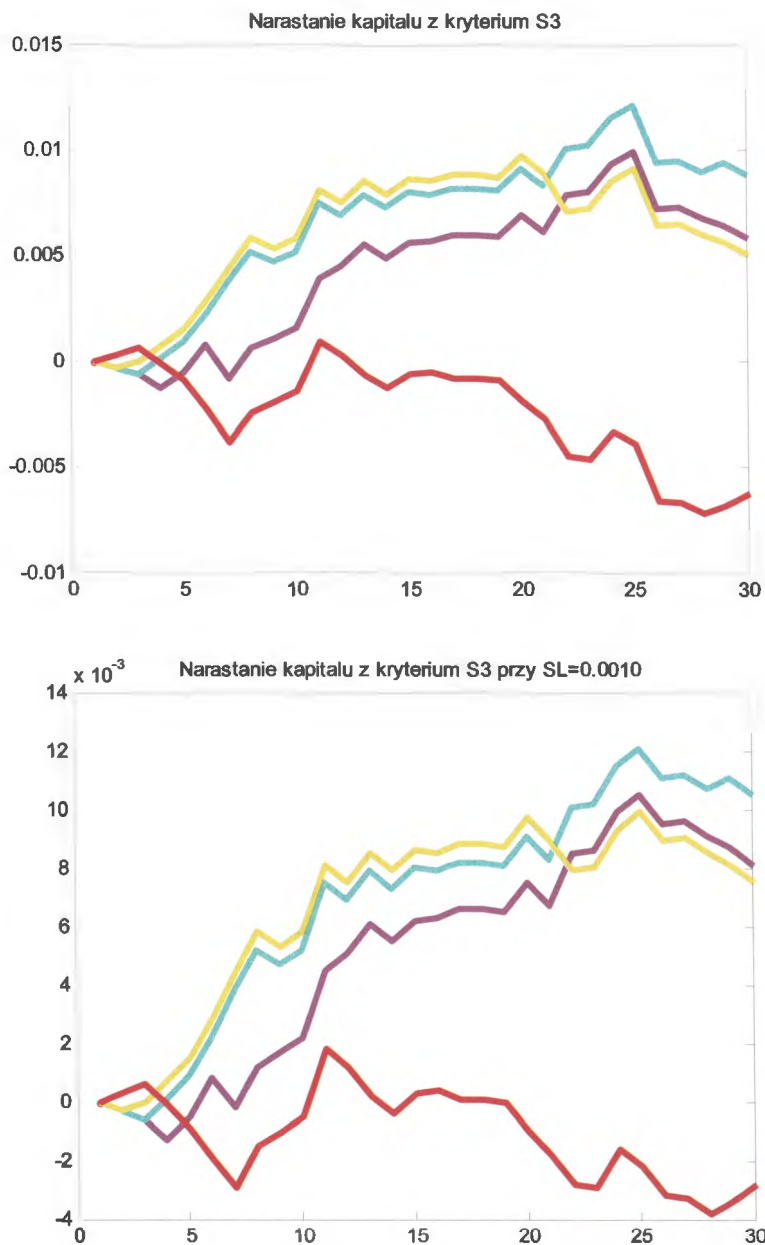
Tab. 3. Wyniki symulacji algorytmu predykcyjnego CL2 B – średnia znormalizowana stopa zwrotu

Wersja algorytmu	Warstwa 1	Warstwa 2	Warstwa 3	Warstwa 4
CL2 B bez SL	0.0264	-0.1331	-0.0339	-0.0339
CL2 B z SL	0.0910	-0.0326	0.0731	-0.0731

Trzecim badanym algorytmem był algorytm z kryterium minimum wg (17). Te badania, przedstawione dla każdej z czterech warstw na rys. 44 przedstawiają zdecydowanie korzystniejszy obraz symulacji. Na wykresie rys. 45 nawet w lewej części (bez użycia mechanizmu *Stop Loss*) widać wyraźnie trzy dodatnie wyniki symulacji na 30 świecach. Najlepszy wynik uzyskuje się tu w drugiej warstwie sieci neuronowej (tab. 4). Stosując ten sam wskaźnik co poprzednio  $r_s$ , tu uzyskuje się całkiem zadowalający rezultat w drugiej warstwie (0.3036). Wynik ten warto zapamiętać w kontekście przyszłych badań i porównań, tym bardziej, że jest on uzyskany bez żadnych dodatkowych wspomagań związanych z technologiami systemu tradingowego takimi jak mechanizmy zarządzania ryzykiem inwestycyjnym typu SL, TP czy czas zamknięcia pozycji. Wyniki te są niejako rezultatem „czystego” statystycznego przewidywania.

Tab. 4. Wyniki symulacji algorytmu predykcyjnego CL2 C – średnia znormalizowana stopa zwrotu.

Wersja algorytmu	Warstwa 1	Warstwa 2	Warstwa 3	Warstwa 4
CL2 C bez SL	0.1949	0.3036	0.1672	-0.2089
CL2 C z SL	0.3262	0.4227	0.3047	-0.1154



Rys. 45. Wykresy narastania kapitału w strategii tradingowej (dla czterech warstw sieci obliczeniowej) zbudowanej z zastosowaniem kryterium  $\Delta S_3$  dla wariantu bez mechanizmu Stop Loss (lewy wykres) i z mechanizmem Stop Loss równym 0.0010 (10 pipsów).

Jeszcze lepsze wyniki uzyskano po włączeniu jednego z tych mechanizmów, czyli SL. Najlepszy wynik, także dla drugiej warstwy, był tu rzędu 0.4227, a więc nieosiągalny dla poprzednio rozważanych algorytmów.

Dla sprawdzenia stabilności tego wyniku przeprowadzono dodatkowe badania algorytmu CL2 C tylko dla drugiej warstwy.



## Podsumowanie

Metoda GMDH sprzężona z mechanizmami półautomatycznego obrotu na platformach brokerskich tworzy bardzo interesujący i obiecujący system transakcyjny. Przykłady użycia w różnych obszarach rynków finansowych wskazują na jej uniwersalność w aspekcie miejsca użycia, ale też na ograniczone możliwości wykorzystania. Podstawowym ograniczeniem metody jest niska moc obliczeniowa współczesnych maszyn. W metodzie konieczne jest bowiem transformowanie przestrzeni predyktorów do nowych wielowymiarowych przestrzeni pochodnych. To z kolei wymusza czasochłonne przeszukiwanie przestrzeni zmiennych objaśniających w celu określenia najbardziej obiecujących modeli predykcyjnych. Dodatkowo, eksperymentalnie sprawdzono, że budowane modele tylko przez chwilę są zdolne do opisanía przyszłości i wymagają ustawicznej adaptacji, zarówno poprzez zmianę parametrów modelu (tu wielomianowego) jak też poprzez zmianę struktury tego modelu. Stąd, trudne wyzwanie obliczeniowe – sedno tzw. inteligencji obliczeniowej, które sprowadzić można do poszukiwania kompromisu pomiędzy dokładnością i pracochłonnością algorytmu. Można tu także skonstatować, że przydatność metody daje się stwierdzić wyłącznie poprzez obliczenia, poprzez eksperymenty na fizycznej rzeczywistości reprezentowanej przez macierze predyktorów. Przyszłość tej metodologii – chyba w wykorzystaniu mocy rozproszonych systemów obliczeniowych zgodnie z ideą *cloud computing*, ale także w poszukiwaniu nowych specyficznych dla metody algorytmów heurystycznego przeszukiwania przestrzeni zmiennych objaśniających.

Zapotrzebowanie na moc obliczeniową i lepsze algorytmu przeszukiwania wynika z przekonania, że warto budować modele o dużej złożoności. W rozdziale 1 widać było związek pomiędzy dokładnością predykcji a złożonością modelu.

Zaimplementowana przez autora do sześciu warstw sieć obliczeniowa z trudem daje się wyjaśnić do poziomu czwartej warstwy tak, by można było odtworzyć pierwotne współczynniki równań regresji (te, związane z pierwotnymi, nieprzetworzonymi danymi) dla najlepszych modeli. Z uwagi na złożoność tego opisu przeniesiony on został do dodatku, gdzie zajął ok. 5 stron. Opis warstwy piątej byłby już 4-5 krotnie obszerniejszy, a warstwy szóstej zająłby już kilkadziesiąt stron. W praktycznej implementacji tych algorytmów autor nie naśladował Aleksieja Ivakhnenki, który zazwyczaj przedstawiał modele wielomianowe w formie



uwzględniającej pierwotne predyktory z wyższymi stopniami złożoności (np. jako człony kwadratowe, sześciennie itd.). Ponieważ eksperymentalnie w niniejszej pracy dowiedziono, że model jest poprawny tylko przez przysłowiową chwilę, to dociekanie, które z danych wejściowych mają wpływ na ciągle zmieniającą się strukturę modelu nie było w istocie ważne. Ważna była zdolność systemu do samoorganizacji, do ciągłej zmiany struktury modelu i parametrów. Adaptacyjność tę wielokrotnie uzyskiwano i potwierdzono empirycznie jej zasadność i skuteczność.

Intencją było umożliwienie implementacji tej metody w dowolnym środowisku programowym. Środowiskiem wybranym przez autora jest Matlab. Zainteresowany czytelnik może wybrać dowolne inne o podobnych lub lepszych funkcjonalnościach.

Warto zauważyć, że obliczenia te dają się zrównoleglić z uwagi na wielokrotnie zagnieżdżaną pętlę. To istotny aspekt rozpatrywanego ograniczenia. Współczesne metody zrównoleglenia obliczeń uwzględniają możliwość tworzenia farm komputerów, czy zastosowania szybkich procesorów graficznych w celach naukowych. To jeden z kierunków rozwoju rozpatrywanej metody GMDH. Drugi, to spokojne oczekiwanie na spełnienie się prawa Moora „gwarantującego” wykładniczy rozwój mocy obliczeniowej.

Zupełnie inne podejście, to użycie intelektu zamiast maszyn. To zapewne, na razie, dopóty górujemy nad maszynami, najlepsze przyspieszenie wzrostu skuteczności metody. To poszukiwanie usprawnień metody GMDH w obszarze poszukiwania lepszych baz dla reprezentacji modeli i odpowiedniejszych kryteriów, poszukiwanie lepszych metod przeszukiwania przestrzeni predyktorów.

W pracy stosowano kilka kryteriów jakości prognoz charakterystycznych dla praktycznego handlu. Był więc stosowany w analizie portfelowej wskaźnik Sharpe’a czy Sortino, był także wskaźnik uwzględniający awersję do maksymalnego obsunięcia kapitału. W fazach testowania stosowane były w przypadku poszukiwania modeli regresyjnych nie tylko tradycyjne kryteria metody najmniejszych kwadratów, ale także minimaksowe. Zwłaszcza kryteria związane z awersją do maksymalnego obsunięcia są, zdaniem autora, bardziej naturalne i związane z osobowością inwestora niż kryteria stosowane w modelach Markowitza czy Sharpe’a. Oczywiście zawsze lepiej jest znaleźć model, który przy danym poziomie ryzyka pozwoli na osiągnięciu wyższego niż niższego zysku. Zawsze lepiej jest osiągnąć założony zysk przy niższym niż wyższym poziomie ryzyka. Jednakże jak zmierzyć tę szczególną cechę inwestora, który wybiera poziom ryzyka „nierozsądnie”? Dlaczego istnieją na rynku tyśiące alternatyw wobec bezpiecznych obligacji?

Praca jest w jakiejś mierze wyrazem fascynacji autora dorobkiem zmarłego niedawno (w 2008 roku) ukraińskiego wybitnego uczonego prof. Aleksieja Ivakhnenki. Ivakhnenko nigdy nie koncentrował się na zadaniach predykcji na rynkach finansowych. Z przeglądu licznych źródeł tego Autora wynika, że do kilku publikacji dotyczących tej problematyki namówili Go raczej niemieccy partnerzy – Johann-Adolf Mueller i Frank Lemke. Kontakt ten nie przyniósł spektakularnych praktycznych rezultatów, ale nie zniechęcił też (np. Franka Lemkego, z którym autor niniejszej pracy ma bezpośredni kontakt) do kontynuacji usprawnień metody. Nigdy jednak nie był, w tej grupie uczonych, rozpatrywany aspekt handlu automatycznego, częstotliwego, o zupełnie nieoczekiwanych cechach. Wprowadzenie infobotów do zarządzania kontami klientów to, w przekonaniu autora, niedaleka przyszłość. To nieodłączny atrybut członka przyszłego społeczeństwa informacyjnego, który będzie poszukiwał specyficznego, outsourcingowego wsparcia osobistych finansów. To, z dużą dozą prawdopodobieństwa, będzie typowe zachowanie drobnego inwestora zniecierpliwionego natarczywością i nieodpowiedzialnością banków, zwłaszcza po fali bankructw tych ostatnich w okresie trwającego kryzysu finansowego.

Handel tradycyjny i w pełni automatyczny to dwa różne światy – w aspekcie teorii podejmowania decyzji, strategii matematycznych, psychologii inwestora, szybkości działania, wpływu technologii komunikacyjnych na teorie racjonalnego inwestowania (i zarządzania ryzykiem). Zastosowanie tradycyjnych koncepcji dla handlu automatycznego może być zupełnie bezzasadne. To tak, jakby rozpatrywać dwa środki komunikacji – wóz konny i samolot. Przyczepiwszy skrzydła z boków wozu konnego, konstatować ze zdziwieniem, że nie unosi się.

Rozwiązaniem dla wozu jest zmiana koni. Na bardzo szybkie.



# Bibliografia

1. Banaszak Z., Józefczyk J. *Towards dedicated decision support tools*. Applied Computer Science vo. 1 no 1 2005, ATH, Bielsko-Biała 2005, pp. 161-181.
2. Bargiela A., Pedrycz W. *Recursive Information Granulation: aggregation and interpretation issues*. IEEE Transactions on Systems, Man and Cybernetics Part B – Cybernetics 33 (1), 96-112.
3. Borawski M. *Analysis of Properties of Mean Value - Pseudo -Variance Representation*. Polish Journal of Environmental Studies vol.17 No 3b, 2008, p. 39-43.
4. Breiman L., Friedman J.H., Olshen R.A., Stone C.J., *Classification and Regression Trees*, Wadsworth International Group, Monterey, 1984.
5. Brock W., Lakonishok j., LeBaron B. *Simple technical trading rules and stochastic properties of stock returns*. Journal of Finance 47(1992), 1731-1764.
6. Brown G., Wyatt J., Harris R., Yao X. *Diversity creation methods: A Survey and Categorisation*. University of Birmingham, 2004.
7. Burda Z. *Fizyka i zarządzanie ryzykiem finansowym*. Postępy fizyki, tom 57, zes.3, 2006
8. Cai B.M., Cai C.X., Keasey K. *Market Efficiency and Returns to Simple Technical Trading Rules: Further Evidence form US, UK, Asian and Chinese Stock Markets*. Asia-Pacific Financial Markets, Springer 2005.
9. Chan Kim W., Mauborgne R. *Strategia błękitnego oceanu*. MT Biznes, Warszawa 2005
10. Chua S. *Sammy Chua's Day Trade*. Wiley Trading, 2007.
11. Cizek P., Härdle W, Weron R. *Statistical Tools for Finance and Insurance*, Springer-Verlag, Berlin 2005.
12. Czapiewski P., Samborska-Owczarek A. *Applicability Verification of Pattern – Based Decision Making Method in Real-life Conditions*. Polish Journal of Environmental Studies vol.16 No 4a, 2007, p. 57-62.
13. DeBontd W., Thaler R. *Does the stock market overreact?* Finance of Journal 40(3) 793-805
14. Dehuri S., Cho S.-B. *Multicriterion Pareto based particle swarm optimized polynomial neural network for classification: A review and state-of-the-art*. Computer Science Review 3(2009) 19-40.

15. Duch W. *Computational Intelligence and Where Is It Going*. In Challenges for Computational Intelligence (Ed. J. Kacprzyk). Springer. vol. 63/2007 s. 1-13.
16. Duda R.O., Hart P.E., Stork D.G., *Pattern Classification*, John Wiley, New York (2000).
17. Elder A. *Come Into My Trading Room*. Wiley Trading, 2002.
18. Elder J.F. *A review of Machine Learning, Neural and Statistical Classification* (eds. Michie, Spiegelhalter & Taylor; Ellis Horwood, 1994), *Journal of the American Statistical Association* 91, no. 433: 436-437.
19. Elton E.J, Gruber M.J. *Nowoczesna teoria portfelowa i analiza papierów wartościowych*. WIG-PRESS, Warszawa 1998.
20. Fama E. Market efficiency, long-term returns, and behavioral finance. *Journal of Financial Economics* 49 (1998) 283-306.
21. Fiszeder P. *Dynamiczna alokacja aktywów – model Markowitza, Rynki finansowe – prognozy a decyzje*, Acta Universitatis Lodziensis, Folia Oeconomica, 177, Wydawnictwo Uniwersytetu Łódzkiego, Łódź 2004.
22. Fiszeder P. *Forecasting the Volatility of Polish Stock Index – WIG20*, w: Forecasting Financial Markets. Theory and Applications, Wydawnictwo Uniwersytetu Łódzkiego, Łódź 2005.
23. Fiszeder P. *Modelling Financial Processes with Long Memory in Mean and Variance*, Dynamic Econometric Models 7, UMK, Toruń, 2006.
24. Friesen G.C.F, Weller P, Dunham L.M. *Price trends and patterns in technical analysis: A theoretical and empirical examination*. *Journal of Banking & Finance* 33(2009) 1089-1100.
25. Fujimoto K., Nakabayashi S. *Applying GMDH Algorithm to Extract Rules from Examples*. *Systems Analysis Modelling Simulation*, Volume 43, Number 10, 2003/10 pp. 1311-1319.
26. Gatnar E. *What is Data Mining*. *Statistics in Transition* vol.5 no. 5 p. 837-842
27. Gencar R. *Linear, non-linear and essential foreign exchange rate prediction with simple technical trading rules*. *Journal of International Economics* 47(1999) 91-107.
28. Grabczewski K. *Zastosowanie kryterium separowalności do generowania reguł klasyfikacji na podstawie baz danych*. pr. dokt. IBS, Warszawa, 2003.
29. Greshilov A.A., Stakun V.A., Stakun A.A. *Matematicheskiye metody postroeniya prognozov*. Radio i sviaz. Moskva 1997.

## Bibliografia

30. Hashem S. , Schmeiser B. *Improving Model Accuracy using Optimal Linear Combinations of Trained Neural Networks*. IEEE Transactions on Neural Networks, 1995.
31. Hashem S. , Schmeiser B. *Improving Model Accuracy using Optimal Linear Combinations of Trained Neural Networks*. IEEE Transactions on Neural Networks, 1995.
32. Hia Jong Teoh, Tai Liang Chen, Ching Hsue Cheng, Hsing Hui Chu. *A hybrid multi-order fuzzy time series for forecasting stock markets*. Expert Systems with Applications 36 (2009) 7888-7897.
33. Hia Jong Teoh, Tai-Liang Chen, Ching-Hsue Cheng. *Frequency-Weighted Fuzzy Time-Series Based on Fibonacci Sequence for TAIEX Forecasting*. Emerging Technologies in Knowledge Discovery and Data Mining. Springer 4819/2007.
34. Ivakhnenko A., Ivakhnenko G. *Problems of Further Development of the Group Method of Data Handling Algorithms*. Part I. Pattern Recognition and Image Analysis vol.10 No.2, 2000, pp. 187-194.
35. Ivakhnenko A., Ivakhnenko G. *Problems of Further Development of the Group Method of Data Handling Algorithms, Part I*. Pattern Recognition and Image Analysis vol.10 No.2, 2000, pp. 187-194.
36. Ivakhnenko A., Ivakhnenko G., Mueller J. *Self- organization of Neural Network with Active Neurons*. Pattern Recognition and Image Analysis 1999 v.4 no. 2 pp. 185-196.
37. Ivakhnenko A., Stepashko V. *Pomekhoustoicivost modelirovanija*, Icjier 1986 , Naukovaja Dumka.
38. Ivakhnenko A.G., *An Inductive Sorting Method for the Forecasting of Multi-dimensional Random Processes and Events with the Help of Analogs Forecast Complexing*, Pattern Recognition and Image Analysis, 1991, vol. 1, no.1, pp.99-108.
39. Ivakhnenko A.G., *Dolgosrochnoe Prognozirovanie i Upravlennie Slozhnymi Sistemami* (Long-Term Forecasting and Control of Complex Systems), Kiev: Tekhnika, 1975.
40. Ivakhnenko A.G., Ivakhnenko G.A., Müller J.A., *Self-Organization of Neural Nets with Active Neurons*, Pattern Recognition and Image Analysis, 1994, vol. 4, no. 2. pp. 185-196.
41. Ivakhnenko A.G., Ivakhnenko G.A., Müller J.A., *Self-Organization of Optimum Physical Clustering of the Data Sample for Weakened Description and Forecasting of Fuzzy Objects*, Pattern Recognition and Image Analysis, 1993, vol. 3, no. 4, pp. 415-422.

42. Ivakhnenko A.G., Ivakhnenko G.A., *Obzor zadac po algoritmam Metoda Gruppovogo Uceta Argumentov*. Pattern Recognition and Image Analysis, 1995 vol. 5 no. 4 pp. 527-535.
43. Ivakhnenko A.G., Ivakhnenko G.A., *Perceptron Synthesis according to Clustering-Balance Criterion*. Pattern Recognition and Image Analysis, 1995, vol. 5, no. 3, pp. 337-341.
44. Ivakhnenko A.G., Ivakhnenko G.A., *Simplified Linear Programming Algorithm as Basic Tool for Open-Loop Control*, Systems Analysis, Modelling, and Simulation, (SAMS), 1995, vol.18-19, pp.315-319.
45. Ivakhnenko A.G., Müller J.A., *Parametric and Non-parametric Selection Procedures in Experimental Systems Analysis*, Systems Analysis, Modeling, and Simulation, 1992, vol. 9, pp. 157-175.
46. Ivakhnenko A.G., Müller J.A., *Problems of Computer Clustering of the Data Sampling of Objects under Study*, Sov. J. Automation Inf. Sci., 1991, vol. 24, no. 1, pp. 58-67.
47. Ivakhnenko A.G., *Samoobujascijesja systemy raspoznavanija i avtomatycznego upravljenija*. Tekhnika, Kijev 1969.
48. Ivakhnenko A.G., Stepashko V.S., *Pomekhoustoichivost' Modelirovaniya* (Noise Immunity of Modeling), Kiev: Naukova Dumka, 1985.
49. Ivakhnenko A.G., Yurachkovskij Yu.P., *Modelirovanie Slozhnykh Sistem po Eksperimental'nym Dannym* (Modeling of Complex Systems from Experimental Data), Moscow: Radio i Svyaz', 1987.
50. Ivakhnenko A.G., Zaichenko J.P., Dimitrov V.D. *Priniatije resenij na osnove samoorganizacii*. Sovetskoe Radio, Moskwa 1976.
51. Ivakhnenko G.A., *Self-Organization of Neuronet with Active Neurons for Effects of Nuclear Test Explosions Forecastings*. System Analysis Modeling Simulation (SAMS), 1995, vol.20, pp.107-116.
52. Jajuga K. *Metody statystyczne w finansach*. Statsoft Polska, 2003.
53. Jajuga K., Jajuga T. *Jak inwestować w papiery wartościowe*. PWN, Warszawa 1994.
54. Jajuga K., Kuziak K., Markowski P. *Inwestycje finansowe*, Wydawnictwo Akademii Ekonomicznej im. Oskara Langego we Wrocławiu, Wrocław 1998.
55. Jing Dang, Brabazon A., Edelman D., O'Neill M. *An Introduction to Natural Computing in Finance*. LNCS, PublisherSpringer Berlin / Heidelberg Volume 5484/2009.

56. Jinhong You, Xian Zhou, Li-Xing Zhu. *Inference on a regression model with noised variables and serially correlated errors*. Journal of Multivariate Analysis 100(2009) 1182-1197.
57. Kacprzyk J., Wilbik A., Zadrozny S. *A linguistic quantifier based aggregation for a human consistent summarization of time series*. In: J. Lawry, E. Miranda, A. Bugarin, S. Li, M. A. Gil, P. Grzegorzewski, O. Hryniewicz (Eds.): *Soft Methods for Integrated Uncertainty Modeling*. Springer-Verlag, Berlin Heidelberg 2006, ss. 183-190, Seria: *Advances in Soft Computing*.
58. Kacprzyk J., Wilbik A., Zadrozny S. *Using a genetic algorithm to derive a linguistic summary of trends in numerical time series*. In: *Proceedings of 2006 International Symposium on Evolving Fuzzy Systems, Ambleside, UK, September 7-9, 2006*, IEEE Press, ss. 137-142.
59. Kacprzyk J., Zadrozny S. *Fuzzy linguistic data summaries as a human consistent, user adaptable solution to data mining*. In: Gabrys B., Leiviskä K., Strackeljhan J. (Eds.): *Do smart adaptive systems exist?* Springer Verlag, Heidelberg New York 2005, ss. 321-339, Seria: *Studies in fuzziness and soft computing*.
60. Kohavi R. *A Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection*. Computer Science Department, Stanford University, 1995.
61. Kompa K., Matuszewska Janica A. *Examination of Warsaw Stock Exchange Indexes Behaviour: Applications of Rolling Windows Variance Ratio Test*. Polish Journal of Environmental Studies vol.17 No 3b, 2008, p. 150-154.
62. Korbicz J., Marcin Mrugalski M. *Confidence estimation of GMDH neural networks and its application in fault detection systems*. International Journal of Systems Science, 2008, Vol. 39, no 8, pp. 783-800.
63. Krawczak M. *Theoretical foundations of neural networks prediction*. W: Studziński J., Drelichowski L., Hryniewicz O., Kacprzyk J. (red.) *Technologie informatyczne w zarządzaniu, systemy wspomaganie decyzji*. IBS PAN, Warszawa 2000, pp. 245-254.
64. Krawczak M., Jakubowski A., Konieczny P., Kulikowski R., Miklewski A., Szkatuła G. *Aktywne zarządzanie inwestycjami finansowymi – wybrane zagadnienia*, Akademicka Oficyna Wydawnicza EXIT, Warszawa 2003.
65. Krogh A., Vedelsby J., *Neural Network Ensembles, Cross Validation and Active Learning*. NIPS, 7 of 1995.
66. Krutsinger J. *Systemy transakcyjne. Sekrety mistrzów*. Wig-Press, Warszawa 1999.
67. Kulikowski R. *Wspomaganie inwestycji obarczonych ryzykiem*. W: Krawczak M., Jakubowski A., Konieczny P., Kulikowski R., Miklewski A., Szkatuła



- G.: *Aktywne zarządzanie inwestycjami finansowymi – wybrane zagadnienia*, Akademicka Oficyna Wydawnicza EXIT, Warszawa 2003, ss. 253-268.
68. Kuś J., Korbicz J. *Statyczne i dynamiczne sieci GMDH w Sieci neuronowe* (eds. Duch, Korbicz, Rutkowski, Tadeusiewicz) w serii pod red. M. Nałęczca *Biocybernetyka i inżynieria medyczna*, PAN, Warszawa 2000.
  69. Kutner R. *Niegaussowskie procesy stochastyczne i niedebye'owska relaksacja oraz zdarzenia rzadkie w świecie realnym*. Uniwersytet Warszawski, 2006.
  70. Lemke F., Mueller J.A. *Self-Organizing Data Mining A Portfolio Trading System*. *Journal of Computational Intelligence in Finance*, 1997/05 pp 12-26.
  71. LeSage J., Sudipto Banerjee, Fischer M.M., Congdon P. *Spatial Statistics: Methods, models & computation*. Computational Statistics and Data Analysis 53(2009) 2781-2785.
  72. Łuniewska M. *The Construction and Evaluation of Sectorial Portfolios*. Polish Journal of Environmental Studies vol.16 No 4a, 2007, p. 161-165.
  73. Madala H.R., Ivakhnenko A. G. *Inductive Learning Algorithms for Complex Systems Modelling*. CRC Press 1994.
  74. Makridakis S., Wheelwright S., Hyndman R. *Forecasting: Methods and Applications*. John Wiley and Sons Inc. 1998.
  75. Marjak H. *The Architecture Selection for Neural Network – a Financial Criterion or a Prediction Criterion*. Polish Journal of Environmental Studies vol.17 No 3b, 2008, p. 255-261.
  76. McLachlan G.J. *Discriminant Analysis and Statistical Pattern Recognition*, Wiley Interscience, New York, 1992.
  77. Merz Ch. J. *Classification and Regression by Combining Models*. Ph.D. thesis, University of California, Irvine, 1998.
  78. Mohsen Mehrara, Ali Moeini, Mehdi Ahrari, Ali Erfanifard. *Investing the efficiency in oil futures market based on GMDH approach*. *Expert Systems with Applications* 36(2009) 7479-7483.
  79. Mrugalski, M., Arinton, E., Korbicz, J. *Sieci neuronowe typu GMDH w diagnostyce systemów przemysłowych*. *Pomiary, Automatyka, Kontrola*. 2002, R. 48, nr 2, s. 10-14.
  80. Mueller J., Lemke F. *Self-organizing Data Mining. An Intelligent Approach to Extract Knowledge from Data*. Berlin, 1999.
  81. Muriel A., *Short-term predictions in forex trading*, *Physica A* 344 (2004) 190-193

## Bibliografia

82. Niederliński A. *Systemy komputerowe automatyki przemysłowej – zastosowania*. WNT, Warszawa 1985.
83. Nowak R. *Statystyka dla fizyków*. WN PWN, Warszawa 2002.
84. Onwubolu G. *Hybrid Self-Organizing Modeling Systems*. Springer 2009.
85. Owens S., Lizotte O. *When to trade*. FX Engines, 2004.
86. Palit A., Popovic D. *Computational Intelligence in Time Series Forecasting*. Springer 2005.
87. Park B.J., Oh S.K., Ahn T.C., Pedrycz W. *A study on multilayer fuzzy polynomial interference system based on an extended GMDH algorithm*. Proc. 8<sup>th</sup> IEEE Inf. Conference on Fuzzy Systems, Seoul 1999.
88. Pawlak Z. *Rough sets and intelligent data analysis*. Information Sciences, 147, 2002, 1-12.
89. Pawlak Z., Polkowski L., Skowron A. *Rough Set Theory*. In: B. Wan, (ed.), *Encyclopedia of Computer Science and Engineering*, Wiley, New York 2004.
90. Pedrycz W. *Computational Intelligence: An Introduction*. CRC Press 1997.
91. Penrose R. *Droga do rzeczywistości*. Prószyński i Ska. Warszawa 2006.
92. Piotrowski E. *Dwoistości wartości kapitału*. Politechnika Białostocka 2002.
93. Piotrowski E., Śładkowski J. *Quantum Diffusion of Prices and Profits*. Physica A, 345(2005) s. 185-195.
94. Piper J. *The Way to Trade*. Harriman Modern Classic, 2006.
95. Raghuraj R.K., Lakshminarayanan S. *Variable Predictive Models – A new multivariate classification approach for pattern recognition application*. Elsevier, Pattern Recognition vol. 42/1, 2009 p.7-17 [sciencedirect.com].
96. Raghuraj R.K., Lakshminarayanan S. *VPMCD: Variable Interaction Modeling Approach for Class Discrimination in Biological Systems*. FEBS Letters, 581, 826-830, 2007.
97. Rejer I. *Fuzzy Expert Model of Index WIG20*. Polish Journal of Environmental Studies vol.17 No 3b, 2008, p. 365-370.
98. Rua A., Nunes C. *International comovement of stock market: A wavelet analysis*. Journal of Empirical Finance, 2009.
99. Rutkowski L. *Metody i techniki sztucznej inteligencji*, WNT, Warszawa 2006.
100. Satchwell Ch. *Pattern Recognition and Trading Decisions*. McGraw Hill, New York 2005.

101. Sawaragi, Y., Soeda, T., Tamura, H. et al., *Statistical Prediction of Air Pollution Levels Using Non-Physical Models*, Automatica (IFAC), 1979, vol. 15, no. 4.
102. Sawaragi, Y., Soeda, T., Tamura, H. et al., *Statistical Prediction of Air Pollution Levels Using Non-Physical Models*, Automatica (IFAC), 1979, vol. 15, no. 4, pp. 441-452.
103. Schwager J.D. *Analiza techniczna rynków terminowych*. WIG-Press, Warszawa 2002.
104. *Self-Organizing Methods in Modeling*, Statistics: Textbooks and Monographs, Farlow, S.J., Ed., New York: Marcel Dekker Inc., 1984, vol. 54.
105. Shahwan T. , Lemke F. *Forecasting Commodity Prices for Predictive Decision Support Systems*. EFITA/WCCA 2005, Via Real, Portugal.
106. Sharpe W.F. *The Sharpe Ratio*. The Journal of Portfolio Management Fall, 1994 s.49-58
107. Sheng-Hsun Hsu, Po-An Hsieh, Ting-Chih Chih, Kuei-Chu Hsu. *A two-stage architecture for stock price forecasting by integrating self-organizing map and support vector regression*. Expert Systems with Applications 36(2009) 7947-7951.
108. Shleifer A. *An inefficient market: An introduction to behavioral finance*. Clarendon Lectures. Oxford University Press 2000.
109. Sortino F., Van der Meer R. *Downside Risk*. The Journal of Portfolio Management Fall., 1991 s.27-31.
110. Sung-Kwun Oh, Pedrycz W. *The design of self-organizing Polynomial Neural Networks*. Information Sciences, 141 (2002) pp.237-258.
111. Szymański S. *Integracja metod klasyfikacji odkrywania wiedzy w systemie ekspertowym*. pr. dokt. Wydział Informatyki Politechniki Szczecińskiej, 2006.
112. Tadeusiewicz R. *Sieci neuronowe*. Akademicka Oficyna Wydawnicza RM, Warszawa 1993.
113. Tak-chung Fu, Fu-lai Chung, Luk R., Chak-man Ng. *Representing Financial Time Series based on data point importance*. In: Engineering Applications of Artificial Inteligence 21(2008) p.277-300.
114. Tarczyński W, Łuniewska M. *Dywersyfikacja ryzyka na polskim rynku kapitałowym*. Placet, Warszawa 2004.
115. Tarczyński W, Łuniewska M. *Metody wielowymiarowej analizy porównawczej na rynku kapitałowym*. PWN, Warszawa 2006.

## Bibliografia

116. Triseev, Y. P., *Approaches to the Solution of Mathematical Programming Problems on the Basis of Heuristic Self-Organization*, Sov. J. Automation Inf. Sci., 1987, vol. 20, no. 2, pp. 30-37.
117. Tsoumakas G., Katakis J., Vlahavas I. *Effective Voting of Heterogeneous Classifiers*. Department of Informatics, Aristotele University of Thessaloniki, 2004.
118. Tsymbal A., Pechenizkiy M., Cunningham P. *Diversity in search strategies for Ensemble Feature Selection*. Department of Computer Science, Trinity College Dublin, 2003.
119. Vapnik V. *Statistical Learning Theory*, Wiley-Interscience, New York, 1998.
120. Wagner N., Michalewicz Z., Khouja M., Mc Gregor R.R. *Forecasting with a Dynamic Window of Time: The DyFor Genetic Program Model*. Springer Berlin/Heidelberg vol.3490/2005, 205-215.
121. Wagner N., Michalewicz Z., Khouja M., Mc Gregor R.R. *Time Series Forecasting for Dynamic Environment: The DyFor Genetic Program Model*. IEEE Transaction on Evolutionary Computation, vol. 11 n. 4 /2007, s.433-453.
122. Weiss S., Kulikowski C., *Computer Systems that Learn: Classification and Prediction Methods from Statistics, Neural Networks, Machine Learning and Expert Systems*, San Francisco, 1991.
123. Weron A., Weron R. *Inżynieria finansowa: Wycena instrumentów pochodnych, Symulacje komputerowe, Statystyka rynku*. WNT, Warszawa 2005.
124. Wierzbicki A. *Badania operacyjne i systemowe w epoce cywilizacji wiedzy*. Instytut Badań Systemowych PAN seria Badania systemowe tom 64, Warszawa 2008, s. 13-35.
125. Wiliński A. *Analiza szeregów czasowych w ujęciu bayesowskim dla funkcji dyskryminacyjnej opartej na rozkładzie gaussowskim*. XI Sesja Informatyki Wydział Informatyki Politechniki Szczecińskiej, Wydawnictwo Informa, Szczecin 2006.
126. Wilinski A. *GMDH Based Computational Intelligence Method for Prediction of Financial Markets*. "Intellectual Systems of Decision Making and The Problems of Computational Intelligence (ISDMCI'2008)", Eupatoria, Tom 1 cz.1, 2008.
127. Wiliński A. *Risk Estimation of Consecutive Losses in Automated Currency Trading*. Polish Journal of Environmental Studies vol.16 No 4a, 2007, p. 380-385.
128. Wiliński A., Czapiewski P. *GMDH-based Multiple Prediction of Financial Markets*. Polish Journal of Environmental Studies vol.17 No 3b, 2008 p.73-78.

129. Wilinski A., Czapiewski P., Sienkiewicz S. *Model interfejsu graficznego w systemie wspomagania decyzji w obrocie instrumentami finansowymi*. X Sesja Informatyki Wydział Informatyki Politechniki Szczecińskiej, Wydawnictwo Informa, Szczecin 2005.
130. Witkowska D. *Artificial Neural Networks Application to Forecast the Participation Units of the Open-End Investment Fund*. Polish Journal of Environmental Studies vol.17 No 3b, 2008, p. 523-529.
131. Witkowska D., Matuszewska A., Kompa K. *Wprowadzenie do ekonometrii dynamicznej i finansowej*. Wydawnictwo SGGW, Warszawa 2008.
132. Yungho Leu, Chien-Pang Lee, Yie-Zu Jou. *A distance-based fuzzy time series for exchange rates forecasting*. Expert Systems with Applications 36(2009) 8107-8114.
133. Zhang Yudong, Wu Lenan. *Stock market prediction of S&P 500 via combination of improved BCO approach and BP neural network*. Expert Systems with Applications 36(2009) 8849-8854.
134. Zhohnarskij, A. A., *Agglomerative Cluster Analysis Procedures for Multidimensional Objects: A Test for Convergence*. Pattern Recognition and Image Analysis, 1992, vol. 2, no. 4, pp. 388-390.
135. Weisstein, Eric W. *Moore-Penrose Matrix Inverse*. Wolfram MathWorld [online] <http://mathworld.wolfram.com/Moore-PenroseMatrixInverse.html> [dostęp: 02/2008]
136. Stokalski B. Cloud computing – informatyka bez kłopotów [dostęp 05/2008] <http://www.infovidematrix.pl/inspiracje/?p=61#more-61>
137. [www.e-forex.ro](http://www.e-forex.ro) [dostęp 08/2008]
138. [www.oanda.com](http://www.oanda.com) [dostęp 05/2007 – 09/2008]
139. [www.investopedia.com](http://www.investopedia.com) [dostęp 02/2005-10/2008]
140. [www.wanadoo.com](http://www.wanadoo.com) [dostęp 02/2006-11/2008]
141. [www.gmdh.com](http://www.gmdh.com) [dostęp 06/2007-11/2008]
142. [www.fxengines.com](http://www.fxengines.com) [dostęp 12/2007 – 02/2009]





Zasadniczym celem niniejszej monografii jest potwierdzenie tezy autora, że fuzja matematyki i technologii egzemplifikowana powiązaniem metod predykcji ze strategią inwestycyjną opartą na mechanizmach technologicznych platform brokerskich daje nową synergetyczną wartość.

Praca koncentruje się na pragmatycznym powiązaniu predykcji z możliwościami jej internetowej (automatycznej) realizacji. Wymagało to uwzględnienia szeregu niespodziewanych ograniczeń natury technologicznej. W pracy przedstawiono wyniki wielu eksperymentów w przestrzeniach historycznych różnych instrumentów finansowych.

**ISSN 0208-8029**

**ISBN 9788389475237**

---

---

**Instytut Badań Systemowych PAN**

**Tel. Centrala 022-38 10 100 / fax 022-38 10 105 e-mail: [ibs@ibspan.waw.pl](mailto:ibs@ibspan.waw.pl)**