



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 poznany*  
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 Mercurego<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 Krutsinger czy Thomas DeMark, opisani w pracy Joe Krutsingera [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 selekcyonerów 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.

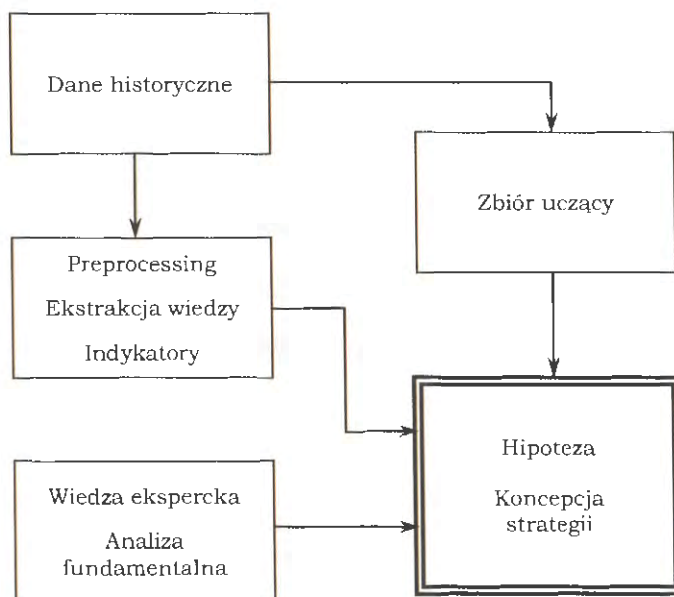
<sup>9</sup> Group Method of Data Handling – angielskie tłumaczenie metody opracowanej przez Ivakhnenkę w oryginale zwane *Metod Gruppovogo Uceta Argumetov* (ros.). W języku polskim nie ma dotąd jednoznacznego przyjętego powszechnie tłumaczenia. Właściwe wydaje się być – Metoda Grupowania Atrybutów czy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



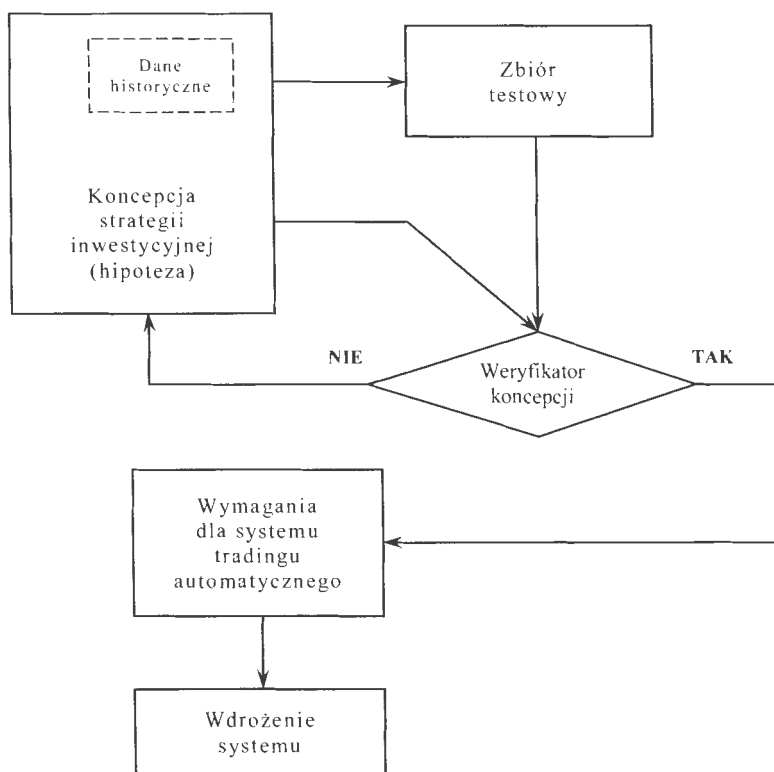
Rys. 1. Schemat blokowy procesu powstawania hipotez badawczych dotyczących koncepcji strategii inwestycyjnych.

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 wskaźników 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.

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.





# 6

## Predykcyjne algorytmy GMDH w analizie portfelowej na przykładzie GPW

*Trudno powiedzieć, co jest prawdą, ale czasami tak łatwo rozpoznać fałsz*

Albert Einstein (1879-1955)

Dotychczas rozpatrywane przykłady zastosowania wybranych algorytmów metody GMDH dotyczyły bardzo płynnych rynków finansowych, w tym najbardziej płynnego – kursu EUR/USD dominującego na rynku pozagiełdowym OTC. Rozpatrywana regresja wieloraka dotyczyła wprawdzie indeksu WIG20, ale koncepcja uzależnienia go od wybranych zmiennych wejściowych opierała się już na najważniejszych czynnikach światowej gospodarki rynkach – cenach ropy, złota, indeksach największych giełd, a więc także na zmiennych o wysokiej płynności, łatwości nabycia lub sprzedaży, a co za tym idzie – na niskich kosztach transakcyjnych. Dziś, w chwili pisania tej pracy, dwupipsowy spread<sup>25</sup> na najważniejszych parach walutowych zaczyna być normą, tak jak 5 lat temu normą był spread trzypipsowy. Ta płynność i rynkowa łaskawość wobec inwestora czyni współczesne platformy brokerskie i bankowe miejscem atrakcyjnym dla tradingu automatycznego. Daleko od tej normy jest ciągle giełda warszawska z wysokimi kosztami transakcji, praktycznie

---

<sup>25</sup> dwupipsowy spread – używane w żargonie środowiskowym określenie prowizji brokera wyrażonej w pipsach (dziesięciotysięcznych wartościach większości kursów walutowych).

wykluczającymi częstotliwie otwieranie pozycji w handlu dziennym. W związku z tym interesujące może być zmierzenie się z tym trudnym środowiskiem, stosując rozważaną tu metodę GMDH w roli narzędzia predykcyjnego. Zapewne to pierwsza taka próba na Giełdzie Papierów Wartościowych w Warszawie (GPW).

## 6.1. Ogólna charakterystyka systemu transakcyjnego

Tym razem w rozważaniach predykcyjnych rozpatrywane będą wybrane akcje GPW pod kątem ewentualnej syntezy kilku rozpatrywanych wcześniej podejść. Wśród nich znaczenie strategiczne będzie miała idea regresji wielorakiej.

W systemie zakłada się więc możliwość wystąpienia statystycznego związku pomiędzy pewną zmienną obserwowaną, zwaną dalej wybraną akcją a pewną przestrzenią danych wejściowych (tu – innymi wybranymi akcjami z GPW). Podobnie więc jak poprzednio dokonuje się wyboru kilkudziesięciu akcji GPW jako pewnej przestrzeni danych wejściowych (przestrzeni zmieniającej się w czasie, w wyniku kolejnego sprawdzania, które z nich byłyby przydatnymi zmiennymi obserwowanymi). Ponieważ w ogólnym przypadku, przy założeniu niestacjonarności procesów stochastycznych w obserwowanych szeregach czasowych nie wiadomo, czy w ogóle statystyczny związek pomiędzy zmiennymi występuje, a jeżeli występuje – to w zależności, od których zmiennych i w jak długim horyzoncie czasowym, to konieczna była wstępna selekcja akcji pretendujących do roli zmiennych obserwowanych. Wybrana zmienna obserwowana, o ile spełniałaby oczekiwania inwestora i zachowywała się zgodnie z modelem predykcyjnym, mogłaby być akcją, która byłaby przedmiotem obrotu giełdowego (wcześniej używano w podobnej sytuacji sformułowania – przedmiotem tradingu). Ta wstępna selekcja jest kolejnym przykładem konieczności udziału człowieka – eksperta w procesach decyzyjnych o dużej złożoności. Wymaga bowiem uwzględnienia takich, trudnych do formalnego opisanie czynników jak pozycja spółki na giełdzie, pogłoski o możliwych przyszłych zachowaniach, przewidywana płynność obrotu. Nie należy więc rozpatrywać w modelu spółek małych, o niewielkich obrotach giełdowych. Mogą być one przedmiotem ciekawych badań o fundamentalnym jednak charakterze, typowym dla np. stylu Warrena Buffeta.

Ponieważ dodatkowym ograniczeniem i jednocześnie wyzwaniem dla rozpatrywanego systemu transakcyjnego jest zamiar jego pełnej automatyzacji, to wynikają z tego także pewne specyficzne cechy

sposobu otwierania i zamykania pozycji. Zakłada się mianowicie, że po upływie kroku obserwacji (tu – w rozpatrywanej macierzy danych – co najmniej co dobę) nastąpi zamknięcie otwartej pozycji i otwarcie nowej. Zakłada się także, że dla każdej otwartej pozycji możliwe jest ustawienie mechanizmu *Stop Loss* skutkującego zamknięciem pozycji długiej w przypadku spadku ceny akcji w ciągu doby (przed jej upływem) poniżej wybranej granicy (np. 2% ceny otwarcia danej akcji). W badaniach nie stosowano mechanizmu *Take Profit* (zamykania pozycji przed upływem doby w przypadku osiągnięcia pewnego satysfakcjonującego inwestora poziomu zysku z obawy przed pełną lub częściową jego utratą).

Symulacja przeprowadzona na danych historycznych sprawdziła się do wielokrotnego zastosowania modelu predykcyjnego zbudowanego zgodnie z zasadami GMDH w taki sposób, że w danej rozpatrywanej chwili czasu poszukiwano modelu regresyjnego (liniowego, o 5-7 wyrazach) najlepiej opisującego przeszłość (tu – przebieg zmian ceny wybranej akcji). Podejście to zbliżone jest w tej części do pierwszych kroków metody samoorganizującego się portfela opisanej przez Muellera i Lemkego [70]. Autorzy ci skoncentrowali się jednak jedynie na optymalizacji złożoności modelu regresyjnego modelu, w którym dla jednej wybranej zmiennej obserwowanej stosowano różne zestawy (struktury) zmiennych objaśniających. Tu – nie przesądza się, która z dostępnych w historycznej przeszłości zmiennych będzie tą właśnie obserwowaną. Stawia się tezę, że zapewne najskuteczniejsza (wg obranego kryterium) będzie zmienna struktura portfela, a więc co pewien okres inne zmienne przejmować będą role zmiennych obserwowanych.

Wzrost liczby wyrazów w modelu regresyjnym, ponad wspomniane 5-7, wiąże się z natychmiastowym katastrofalnym dla komputera wzrostem złożoności i czasu niezbędnego do wykonania obliczeń. Na podstawie najlepszego w danej chwili spośród ponad 100 tys. modeli (dla modeli 5-wyrazowych) dokonywano predykcji ceny na dwa kroki wprzód. Następnie, w kolejnym kroku ponownie dla nowej chwili czasowej poszukiwano najlepszego spośród ok. 100 tys. modeli i jeżeli obie predykcje wykazywały zgodność, co do kierunku zmian ceny to podejmowana była decyzja inwestycyjna.

Decyzja ta polegała na symulacji zakupu akcji, jeżeli obie predykcje wskazywały na wzrost ceny. W przeciwnym przypadku decyzja sprowadzała się do zaniechania działania. Jeżeli jednak byłaby możliwa krótka sprzedaż to rozpatrywana była także predykcja spadku ceny i podejmowana decyzja o otwarciu pozycji krótkiej (wykluczonej obecnie na GPW bezpośrednio na akcjach, a w ograniczonym zakresie moż-

liwa do wykonania za pomocą kontraktów terminowych, lecz wyłącznie na kilku wybranych walorach).

Następnie po upływie symulowanego kroku porównywane były rzeczywiste (z danych historycznych) i modelowane przebiegi cen wybranego waloru. Jeżeli prognoza była udana (tzn. kierunek prognozowanej zmiany ceny pokrywał się z jej rzeczywistą realizacją), to odnotowywano dodatni zwrot na inwestycji (krótkoterminowej, na odcinku jednego kroku czasowego – tu dobowego), jeżeli nie, to sprawdzano, czy porażka nie przekracza dopuszczalnej granicy *Stop Loss* (SL). Jeżeli ją przekraczała, to przyjmowano porażkę w wysokości SL.

Badania przeprowadzone były na danych pozyskanych z GPW w 2008 roku. Macierz danych o wymiarach  $N \times M = [400 \times 28]$  zawiera 400 wierszy pobranych co dobę cen 28 wybranych akcji. Każda kolumna jest więc szeregiem czasowym zmian ceny wybranej akcji. Ostatnią 28. kolumną są zmiany indeksu WIG20. Fragment tej macierzy danych zaprezentowano w listingu 9. Widać tu duże zróżnicowanie cen akcji, gdyż takie w rzeczywistości występuje. Stanowi to pewna niedogodność w zarządzaniu portfelem o zmiennej strukturze w związku z tym będzie zaproponowana specyficzna normalizacja tej macierzy.

```
C = [34.50, 50.40, 2.82, 320.10, 74.95, 220.00, 41.00,  
129.60, 12.57, 42.00, 79.40, 196.00, 102.00, 37.50,  
46.53, 7.80, 3.06, 266.00, 214.00, 3.23, 51.70,  
42.30, 7.59, 6.08, 89.90, 22.79, 24.40, 3232.61,  
  
36.00, 51.40, 2.79, 333.00, 75.00, 221.50, 41.40,  
130.00, 12.82, 42.60, 79.00, 195.00, 102.80, 36.40,  
47.15, 7.80, 3.09, 265.50, 223.00, 3.24, 52.30,  
45.70, 7.62, 6.20, 89.10, 23.60, 24.28, 3324.23,  
  
36.00, 51.40, 2.77, 340.00, 74.50, 223.80, 41.75,  
132.20, 12.82, 42.43, 79.00, 200.00, 106.10, 36.30,  
48.15, 7.80, 3.07, 263.60, 233.00, 3.39, 52.25,  
47.29, 7.21, 6.38, 87.90, 23.81, 24.40, 3395.46 ...
```

*Listing 9. Przykład macierzy danych (trzy pierwsze wiersze) zawierającej informacje o cenach akcji z GPW. Ostatnią zmienną w 28. składnikowym wierszu jest wartość WIG20.*

### **6.1.1. Założenia do algorytmu opartego na zmiennej strukturze portfela akcji**

Zakłada się, że badania będą przeprowadzone na portfelach akcji o liczności 5 (następnie 7) akcji o różnych strukturach i składach. Przez strukturę portfela będzie się rozumiało kombinacje indeksów (numery kolumn akcji, a przez skład – udział wartościowy poszczególnych akcji w kwocie zainwestowanego kapitału. Dla ułatwienia zarządzania środkami w postaci pakietów akcji wszystkie kolumny zostały znormalizowane. Normalizację przeprowadzono w taki sposób, że wartość początkowa akcji (cena zamknięcia w pierwszy wierszu macierzy danych) została pomnożona przez najwyższą możliwą liczbę całkowitą (liczbę akcji), taką, że iloczyn nie przekraczał zadanej umownej kwoty (w badaniach przyjęto ten pułap na poziomie 10000 PLN). Po normalizacji macierz danych przedstawiona na listingu 9 zmieniła wartości do zbliżonych do 10000 PLN jak to przedstawiono w listingu 10. Pierwsza kolumna zawiera liczby akcji pozwalające na zbudowanie pakietu nieprzekraczającego 10000 PLN. Np. w pierwszym wierszu 289 akcji, które początkowo kosztują 34.50 PLN tworzą pakiet o wartości 9970 PLN. Kolejne kolumny to następne notowania tego samego pakietu (podobnie jak w listingu 9 rozpatrywane są w tym przykładzie tylko trzy pierwsze notowania).

Zaproponowana normalizacja pozwala na traktowanie każdego waloru tak, jak gdyby na początku symulacji wszystkie akcje miały jednakową w przybliżeniu cenę. W miarę upływu czasu niektóre z nich będą znacznie przekraczać próg odniesienia (tu 10000 PLN), inne będą spadać poniżej swej wartości początkowej. Zmian tych nie widać w listingu 10 gdyż przedstawia on w kolejnych kolumnach zaledwie trzy pierwsze zmiany (3 dni) pakietów akcji.

### **6.1.2. Fuzja analizy portfelowej i GMDH**

Do zbudowania systemu transakcyjnego łączącego typowe podejście portfelowe z rozpatrywaną tu metodą grupowania argumentów zastosowano złożony wieloetapowy algorytm. Jego pierwsza część miała na celu znalezienie struktury optymalnego (przez pewien okres w rozpatrywanej przestrzeni danych historycznych) portfela akcji.

Pojawiają się tu natychmiast pytania co to znaczy pewien okres, jakie jest kryterium optymalności portfela, jak selekcjonować akcje, jaki powinien być skład (udziały wagowe poszczególnych pakietów akcji). Otóż większość tych problemów rozwiązywano poprzez wielokrotnie

testowane modele regresyjne z zastosowaniem idei regresji wielorakiej (wielokrotnie, oznacza do granic mocy obliczeniowej komputera lub systemu przetwarzania złożonego z wielu maszyn).

Cn =	[0.0289	0.9970	1.0404	1.0404
	0.0198	0.9979	1.0177	1.0177
	0.3546	1.0000	0.9893	0.9822
	0.0031	0.9923	1.0323	1.0540
	0.0133	0.9968	0.9975	0.9909
	0.0045	0.9900	0.9968	1.0071
	0.0243	0.9963	1.0060	1.0145
	0.0077	0.9979	1.0010	1.0179
	0.0795	0.9993	1.0192	1.0192
	0.0238	0.9996	1.0139	1.0098
	0.0125	0.9925	0.9875	0.9875
	0.0051	0.9996	0.9945	1.0200
	0.0098	0.9996	1.0074	1.0398
	0.0266	0.9975	0.9682	0.9656
	0.0214	0.9957	1.0090	1.0304
	0.1282	1.0000	1.0000	1.0000
	0.3267	0.9997	1.0095	1.0030
	0.0037	0.9842	0.9823	0.9753
	0.0046	0.9844	1.0258	1.0718
	0.3095	0.9997	1.0028	1.0492
	0.0193	0.9978	1.0094	1.0084
	0.0236	0.9983	1.0785	1.1160
	0.1317	0.9996	1.0036	0.9496
	0.1644	0.9996	1.0193	1.0489
	0.0111	0.9979	0.9890	0.9757
	0.0438	0.9982	1.0337	1.0429
	0.0409	0.9980	0.9931	0.9980
	0.0003	0.9698	0.9973	1.0186]

Listing 10. Przykład macierzy danych po normalizacji. Wszystkie dane są wynikiem dzielenia wartości znormalizowanych przez 10000.

Kryterium optymalizacji w każdym z tysięcy lub setek tysięcy rozpatrywanych modeli regresyjnych ustalane było na podstawie zbieżności przebiegu rzeczywistego zmian ceny wybranej akcji i jej regresyjnego modelu liniowego o 5 wyrazach zbudowanych na wartościach innych akcji w rozpatrywanym okresie. Liczbę wyrazów w modelu (tu 5) wybrano arbitralnie, jednak głównie pod presją natychmiastowego

wzrostu czasu obliczeń. Wzrost liczby wyrazów np. z 5 do 6 oznaczał bowiem sześciokrotnie więcej modeli do rozpatrzenia, a wzrost liczby wyrazów z 5 do 7 oznaczał  $6 \times 7$  czyli 42-krotny wzrost liczby modeli i w przybliżeniu takie również wydłużenie czasu obliczeń. Wiele z prezentowanych tu modeli wymagało wielogodzinnych obliczeń na stosunkowo szybkiej maszynie. Kilkudziesięciokrotny wzrost czasu przetwarzania czynił metodę praktycznie nieprzydatną.

Stąd właśnie konieczność stosowania wielu heurystycznych zabiegów, które w przeszłości często pomijał wielokrotnie cytowany Ivaikhnenko i jego następcy.

Składniki modelu regresyjnego (wyrazy równania stanowiącego model zmian obserwowanej zmiennej) dobierane były wg jednego z wcześniej prezentowanych autorskich algorytmów kombinatorycznych GMDH. Składnikami modelu regresyjnego każdej rozpatrywanej kolejno zmiennej objaśnianej (akcji) mogą być wybrane pozostałe zmienne (akcje) odpowiednio wybrane, ale także pewne nowe zmienne uzyskane z przesunięcia w czasie zmiennych pierwotnych o krok, dwa lub więcej wprzód (przesunięcie kolumny wprzód w dziedzinie czasu powoduje, że w każdym wierszu macierzy danych pojawiają się dodatkowe dane z bezpośredniej „przeszłości”). W ten sposób macierz początkowa danych wejściowych o 28 kolumnach została rozszerzona poprzez konkatencję poziomą do macierzy o 109 kolumnach. W kolumnach od 29 do 109 umieszczone zostały wektory przesuniętych cen akcji. Pozostawiając wybraną kolumnę akcji obserwowanej a pozostałe 27 kolumn umieszczając obok w grupach kolejno bez przesunięcia, z przesunięciem o 1, 2, 3 wiersze wprzód uzyskano liczbę  $1+4 \times 27$  kolumn, czyli wspomniana macierz o wymiarach  $[400 \times 109]$ .

W wyniku przeszukiwania tysięcy modeli regresyjnych tworzonych z udziałem kolumn przesuniętych znajdowany był model najlepszy wg ustalonego kryterium zbieżności. Najlepszy model na danym kroku symulacji wykorzystywany był do prognozowania zmiany ceny akcji objaśnianej w horyzoncie predykcyjnym o długości dwóch kroków wprzód. Jeżeli prognoza wykonywana w danym kroku pokrywała się z prognozą kroku poprzedniego to wykonywane było symulowane otwarcie pozycji długiej na danym, opisywanym walorze. Po upływie kroku skuteczność predykcji była oceniana i na tej podstawie po wielu krokach można było estymować przydatność waloru. Jakość akcji (waloru) pretendującej do portfela oceniana była na podstawie wskaźnika Sharpe’a na wybranym odcinku kilkudziesięciu kroków (czyli tu – dni).

Ta zasada została początkowo sprawdzona nie na portfelu, lecz na jednej wybranej akcji uznanej za właściwą do pełnienia roli akcji ob-



serwowanej. Akcja ta została wybrana w wyniku wstępnego przeglądu efektywności wszystkich dostępnych akcji (kolumn) w wybranym odcinku czasu. W rozpatrywanych tu badaniach była to akcja nr 13. Mniej w tej chwili istotna jest identyfikacja spółki, która kryje się za tym indeksem, wszak jej atrakcyjność i tak zmienia się w dłuższym horyzoncie czasowym (co częściowo widać na wykresie rys. 94). Wartość początkowa tej akcji wynosiła 102 PLN. W strategii przyjęto, że obrotu dokonuje się wyłącznie jedną akcją, otwierając pozycje długie lub krótkie i zamykając je po upływie doby (jednej świecy).

Badania przeprowadzono przy następujących dodatkowych ograniczeniach:

ww – numer wybranej zmiennej obserwowanej  
(tu np. ww = 13);

K – liczba kroków wykonanych w badaniach symulacyjnych (tu – 350 dla dostępnych 400 wierszy);

hp – horyzont predykcyjny (tu hp=2 kroki wprzód);

zakr – liczba kroków wstecz rozpatrywanych w modelu regresyjnym (tu, po kilku badaniach wybrano zakr=30);

SL – *Stop Loss* – granica dopuszczalnej straty tu ustawiana na poziomie np. 1/50 wartości początkowej (otwarcia) wybranego waloru;

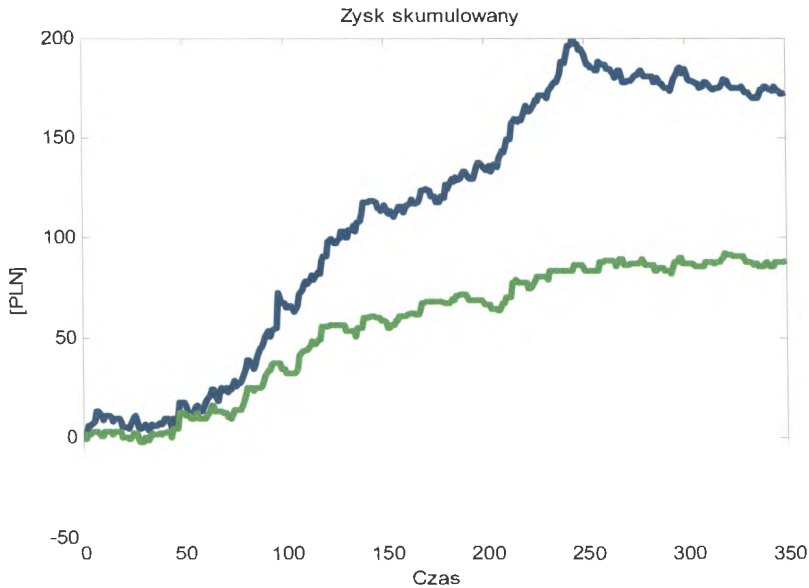
n – liczba zmiennych (kolumn) wykorzystywanych w tworzeniu kolejnych modeli regresyjnych dla tej zmiennej (tu np. n=27, czyli tyle, ile wprowadzono zmiennych wejściowych pomniejszych o jedną rozpatrywana jako zmienna obserwowana).

Najważniejsze wyniki badań symulacyjnych skryptu opisującego model ze zmienną nr 13 jako wyjściową przedstawione są na rys. 94. Na wykresie widać stosunkowo płynne narastanie kapitału przez pierwsze 250 świec (ok. roku) a następnie istotne pogorszenie jakości tradingu, za przyczyną gorszej skuteczności pozycji krótkich. Na zielonej krzywej przedstawiającej skuteczność otwierania tylko długich pozycji tego załamania nie widać.

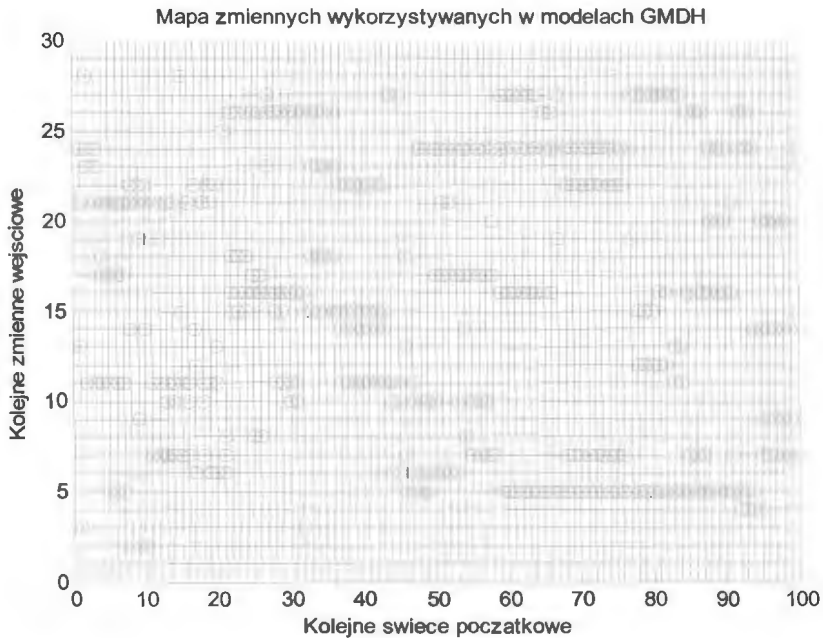
Wartości współczynników Sharpa są tu odpowiednio równe 0.26 dla krzywej niebieskiej i 0.18 dla zielonej. Wynik końcowy dla dopuszczalnych na giełdzie warszawskiej wyłącznie pozycji długich jest rzędu 90 PLN, co oznacza zysk rzędu 90% po roku stosowania tej strategii (z inwestowaniem jedną akcją o wartości początkowej 102 PLN). To interesujące wyniki zwłaszcza, że uzyskane dla jednej wybranej akcji. Jest to jednak akcja najbardziej obiecująca w rozpatrywanym okresie,

a w końcowej fazie symulacji widać już (na krzywej niebieskiej) zdecydowane pogorszenie wyniku. Nie należy więc opierać na tak niepewnym instrumencie koncepcji zasadniczej strategii inwestycyjnej. Zapewne występuje silna zmienność w czasie przydatności poszczególnych akcji. Pewną miarą tej niestacjonarności procesu może być mapa zmiennych stosowanych w algorytmie GMDH do regresyjnego opisu przeszłości i budowania bieżących modeli predykcyjnych. Sytuację tę przedstawia rys. 95, na którym na każdej pionowej zaznaczone są numery zmiennych wejściowych tworzących aktualnie najlepszy model predykcyjny. Pomimo zmienności, widoczne są dla niektórych zmiennych wejściowych stosunkowo długie okresy wykorzystywania ich w kolejnych modelach.

Interesujące mogą być także wnioski z obserwacji dokładności aproksymacji tej zmiennej za pomocą różnych modeli (wykorzystujących różny zestaw zmiennych objaśniających wskazywany na rys. 95).



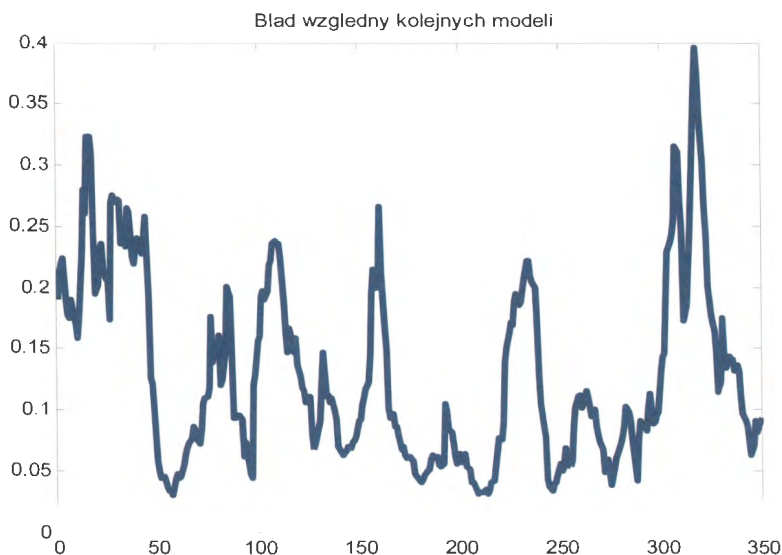
Rys. 94. Wykres narastania kapitału dla strategii z pojedynczą akcją (tu nr 13 o wartości początkowej 102,00 PLN). Krzywa niebieska to wynik dla strategii umożliwiającej otwieranie pozycji w obu kierunkach, zielona – tylko dla pozycji długich.



Rys. 95. Mapa zmienności modeli regresyjnych w symulacji systemu transakcyjnego z jedną zmienną objaśnianą i czterema zmiennymi objaśniającymi (w każdej chwili na wykresie wskazane są 4 numery zmiennych wejściowych).

Wyniki wg kryterium (2) przedstawiono na rys. 96. Widać tu wyraźnie dość dobre wyniki prawie w całym zakresie badań. Kryterium często spada poniżej 10%, a więc świadczy o bardzo dobrym dopasowaniu modelu do obserwowanej rzeczywistości (w przeszłości). W zakresie, gdzie te wyniki są nieco gorsze – w końcowym okresie symulacji – występuje także pogorszenie na wykresie narastania kapitału (rys. 94).

Drugim ważnym badaniem, w zasadzie poprzedzającym opisaną wyżej symulację pojedynczego waloru było badanie predykcji wielorakiej. Opracowany autorski skrypt umożliwił porównanie wskaźników Sharpe'a równocześnie wszystkich rozpatrywanych zmiennych (tu 28 akcji). Przydatność tych zmiennych była badana dla sytuacji, gdyby dopuszczalne było inwestowanie w obu kierunkach, oraz dla sytuacji dopuszczalności inwestowania tylko w kierunku zakupu walorów.



Rys. 96. Błąd względny dokładności aproksymacji w kolejnych modelach regresyjnych.

Badania przeprowadzone były na danych pozyskanych z GPW w 2008. Macierz danych o wymiarach [400x28] zawierała 400 wierszy pobranych co dobę cen 28 wybranych akcji. Każda kolumna jest więc szeregiem czasowym zmian ceny wybranej akcji. Ostatnią kolumną są zmiany indeksu WIG20. W ten sposób wybrana została zmienna nr 13, której wyniki testowania, jako najlepszej w całym rozpatrywanym horyzoncie czasowym, przedstawiono powyżej.

Wykaz rozpatrywanych akcji w macierzy danych:

AGORA	ASSECO	POL	BIOTON	BRE	BUDIMEX	BZWBK	CERSANIT
CEZ	GETIN	GTC	HANDLOWY	KETY	KGHM	KOPEX	
LOTOS	MILLENNIUM	MOSTALEXP	PBG	PEKAO	PGNIG	PKNORLEN	
PKOBP	POLICE	POLIMEXMS	POLNORD	TPSA	TVN	WIG20	

Listing 11. Wykaz akcji wybranych do badań symulacyjnych w kolejności ich występowania w macierzy danych.

Omówione cechy systemu opartego na transakcjach z jedną zmienną zadecydowały, że rozważono system portfelowy. Struktura portfela powstawała w wyniku wyboru 5 najlepszych akcji posiadających najwyższe w fazie uczenia wskaźniki Sharpe'a. Liczba akcji dobrana została arbitralnie, aczkolwiek z dozą rozsądku – większa liczba akcji to nieliniowo narastające wydłużenie czasu obliczeń. Struktura portfela (skład jakościowy) zmieniana była na każdym kroku symulacji w fazie uczenia systemu. W ten sposób po zakończeniu tej fazy powstawała macierz indeksów najlepszych akcji IX o wymiarach [LLx5] (w badaniach stosowano liczbę serii symulacji np. LL=10), której każdy wiersz określał portfel inwestycyjny poprzez numery akcji w nim wymienione. W każdym wierszu akcje wymienione były w kolejności od posiadającej najwyższy wskaźnik Sharpe'a do najniższego.

Do uczenia systemu stosowano odrębny plik. Po nauczeniu systemu – które portfele (o jakiej strukturze i składzie) w bezpośredniej historii poprzedzającej chwilę bieżącą były najlepsze, następowała faza weryfikacji tych portfeli na nowych danych (z zarejestrowanej „przyszłości”), nie używanych w fazie uczenia. Przez strukturę portfela będzie się rozumiało kombinacje indeksów (numerów) akcji, a przez skład – udział wartościowy poszczególnych akcji w kwocie zainwestowanego kapitału.

Testowanie portfeli odbywało się za pomocą innego pliku. Najważniejsze stosowane w rozpatrywanych badaniach definicje:

KK – liczba symulacji wykonywanych z portfelem o stałej strukturze; w fazie uczenia (na danych z „przeszłości”) to najczęściej 30 symulacji wykonywanych co świecę (jedną dobę), a w fazie testowania (weryfikacji na danych z „przyszłości”) przyjmowano KK = 5, 10 lub 15 świec (najczęściej 15);

LL – liczba serii symulacji o licznosci KK wykonywanych co zadana liczbę świec (dni, gdy świece są dobowe). W badaniach najczęściej rozpatrywano 10 serii symulacji rozpoczynanych co 10 dni. Oznaczało to szczelne wypełnienie całej przestrzeni danych. Po zakończeniu każdej serii następowało testowanie portfela.

zakr – zakres lokalnie rozpatrywanych świec od chwili bieżącej wstecz w celu obliczenia współczynników modelu regresyjnego; tu, najczęściej zakr=30;

IP – wektor numerów świec początków symulacji;

kr – inkrement w przestrzeni modeli regresyjnych; z uwagi na czas obliczeń nierealne jest przeszukanie całej przestrzeni

wszystkich możliwych modeli, stosowano tu przeszukiwanie polegające na obliczaniu co kr=6 modelu.

Dla ułatwienia zarządzania środkami w postaci pakietów akcji wszystkie kolumny zostały znormalizowane. Normalizację przeprowadzono w taki sposób, że wartość początkowa akcji została pomnożona przez najwyższą możliwą liczbę całkowitą (liczbę akcji), taką, że iloczyn nie przekraczał zadanej umownej kwoty (w badaniach przyjęto ten pułap na poziomie 10000 PLN).

Zastosowana będzie metoda GMDH dla ustalenia chwilowej struktury optymalnej portfela (wg Markowitza) na podstawie wielokrotnie powtarzanej regresji wielorakiej z kryterium optymalizacji ustalonym na podstawie zbieżności rzeczywistego i regresyjnego modelu liniowego o 5 wyrazach (arbitralnie). Składniki modelu regresyjnego dobierane są wg jednego z przedstawianych wcześniej autorskich algorytmów kombinatorycznych GMDH.

Składnikami modelu regresyjnego każdej rozpatrywanej kolejno zmiennej objaśnianej (akcji) mogą być wybrane pozostałe zmienne (akcje) odpowiednio wybrane, ale także pewne nowe zmienne uzyskane z przesunięcia w czasie zmiennych pierwotnych o krok, dwa lub więcej wstecz. W ten sposób macierz początkowa danych wejściowych o 28 kolumnach została rozszerzona poprzez konkatenację poziomą do macierzy o 109 kolumnach (w kolumnach od 29 do 109 umieszczone zostały wektory przesuniętych cen akcji).

W wyniku przeszukiwania tysięcy modeli regresyjnych tworzonych z udziałem kolumn przesuniętych znajdowany był model najlepszy wg ustalonego kryterium zbieżności. Najlepszy model na danym kroku symulacji wykorzystywany był do prognozowania zmiany ceny akcji objaśnianej w horyzoncie predykcyjnym o długości dwóch kroków wprzód. Jeżeli prognoza wykonywana w danym kroku pokrywa się z prognozą kroku poprzedniego to wykonywane jest symulowane otwarcie pozycji długiej na danym, opisywanym walorze. Po upływie kroku skuteczność predykcji jest oceniana i na tej podstawie po wielu krokach można estymować przydatność waloru. Jakość akcji (waloru) pretendującej do portfela oceniana jest na podstawie wskaźnika Sharpe'a na wybranym odcinku kilkudziesięciu kroków (czyli tu dni).

Formalnie algorytm ten przedstawić można jako sekwencję następujących kroków:

A. Cześć pierwsza – faza uczenia:

1. Przygotowanie podstawowej macierzy danych w postaci  $X_{ij}$ ,  $i=1,2, \dots, N$ ;  $j=1,2, \dots, M$ , gdzie  $i$  to indeks wiersza danych związany z chwilą czasu, w której pobrano dane, a  $j$  – to indeks zmiennej (akcji) – kolumny danych;
2. Wybór jednej z akcji (kolumn)  $j=1,2, \dots, M$  jako zmiennej objaśnianej, która będzie opisywana jako zmienna wyjściowa w modelu regresyjnym  $Y=f(X_j)$ ;
3. Zbudowanie rozszerzonej poprzez konkatencję poziomą macierzy danych wejściowych  $X_{rij}$ ,  $i=1,2,\dots, N$ ,  $j=1,2, \dots, 1+4*(M-1)$ . Dodatkowe kolumny tej macierzy uzyskano poprzez przesunięcie kolumn danych podstawowych o 1-3 wiersze wpród.
4. Określenie chwili początkowej symulacji  $t_0$  i horyzontu badań do chwili końcowej  $t_k$  z krokiem  $dt$  (w rozpatrywanych danych była to jedna doba).
5. Poszukiwanie najlepszego modelu regresyjnego zbudowanego na pięciu (arbitralnie dobrana liczba, wcześniej uzasadniona czasem obliczeń) wybranych zmiennych objaśniających  $j_1, j_2, \dots, j_5$  należących do zbioru indeksów kolumn  $\{1,2,\dots, 109\}$ . Liczba 109 to liczba kolumn macierzy rozszerzonej:  $1+4*(28-1) = 109$ . Poszukiwanie tego modelu odbywa się poprzez badanie tylko niektórych kolumn – oddalonych od siebie o pewien dobrany racjonalnie inkrement  $kr$  (z powodu nieliniowo narastającego czasu obliczeń przy próbach stosowania „gestego” przeszukiwania macierzy danych zastosowano inkrement  $kr = 6$ ). Poszukiwanie najlepszego modelu prowadzono wg kryterium zbieżności na odcinku predykcyjnym ( $hp=2$ ) a w innych badaniach, wg kryterium zbieżności na całym wektorze wyników modelowania porównanych z rzeczywistymi wartościami zmiennej objaśnianej. Kryteria te były wyjaśnione w rozdziale 1 jako (2) – (4).
6. Na podstawie znalezionej najlepszego modelu wykonywana była predykcja na dwa kroki (doby) wpród i podejmowana była symulowana decyzja o otwarciu pozycji (zakupie akcji) tylko wtedy, gdy bieżąca predykcja pokrywała się co do znaku zmiany z predykcją wykonaną w poprzednim kroku.

7. Ponieważ w chwili bieżącej znany jest wynik prognozy z poprzedniego kroku wynik ten był dla danej zmiennej zapamiętywany i wizualizowany.
  8. Po wykonaniu kroku w dziedzinie czasu następował powrót do początku pętli rozpoczynającej się w p.4.
  9. Gdy osiągnięty był koniec symulacji dla danej zmiennej objaśnianej obliczano jej wskaźnik Sharpe'a dla uzyskanych wyników.
  10. Rozpoczynając od p. 2 wybierano kolejną akcję jako zmienną objaśnianą i kolejno doprowadzano do ustalenia jej atrakcyjności inwestycyjnej na podstawie kryterium Sharpe'a z p. 9.
  11. Wszystkie oceniane akcje były sortowane wg malejącego wskaźnika Sharpe'a. Najlepsze z nich (tu – 5 najlepszych) stanowiły chwilową strukturę portfela inwestycyjnego o składzie zrównoważonym kapitałowo (w przybliżeniu po 20%).
- B. Część druga – faza testowania:
1. Portfel o ustalonej w p. 11 strukturze był inwestowany przez 1 krok.
  2. W innych badaniach struktury portfela nie zmieniano przez 10 kroków.

Wyniki tych badań przedstawia się poniżej.

## **6.2. Badania portfela akcji GPW utworzonego wg zasad GMDH**

Zgodnie z opisanym wyżej algorytmem badania przeprowadzono w dwóch fazach – pierwsza poświęcona była przygotowaniu portfela (faza uczenia systemu), druga testowaniu portfela określonego w fazie pierwszej.

Badania prowadzono w trzech wybranych odcinkach czasowych macierzy danych z uwagi na długi, wielogodzinny czas obliczeń. Prowadzono je za pomocą dwóch odrębnych skryptów matlabowych – oddzielnie dla uczenia i testowania.



### 6.2.1. Faza uczenia systemu – generowanie propozycji portfeli inwestycyjnych

Badania przeprowadzono w trzech kolejno po sobie następujących fragmentach przestrzeni danych. Każdy fragment obejmował 10 symulacji rozpoczynanych co 10 świec. W ten sposób przestrzeń danych była szczelnie wypełniona predykcjami. W serii symulacji rozpoczynanych co 10 świec nie zmieniano struktury portfela. (gielDa3deBlinlagS).

Czas wykonania każdego z trzech skryptów dla kolejnych trzech fragmentów przestrzeni danych (współczesny laptop) – to ok. 3 godz. Parametry w tabeli 15 dotyczą pierwszego z trzech okresów symulacji, wyjaśnione są one w poprzednim podrozdziale.

Tab. 15. Parametry symulacji dla pierwszego okresu uczenia systemu.

Parametry symulacji	Wartości parametrów	Nazwa zmiennej w pliku
Nazwa pliku	gielDa3deBlinlagS	
Liczba akcji w portfelu	5	
Nr świecy początkowej	40	IP
Odległość między świecami pocz.	10	
Liczba świec w symulacji	30	KK
Liczba symulacji w badaniu	10	LL
Zakres lokalnego zbioru uczącego	30	zakr
Inkrement w przestrzeni modeli	6	kr

W wyniku przeprowadzenia symulacji zgodnie z opisanym algorytmem uzyskano macierz B wszystkich 28 współczynników Sharpe'a uporządkowaną malejąco. Kolumny dopowiadają kolejnym zmiennym (akcjom), wiersze kolejnym symulacjom rozpoczynanym co 10 świec. Wyniki siedmiu najlepszych wskaźników przedstawiono w listingu 12 a listingu 13 ujawnione są numery akcji, które uzyskały te najbardziej obiecujące wyniki. Przedstawiono tu jedynie po siedem najlepszych wyników i odpowiadającym im numerów akcji.

Najlepsze wskaźniki Sharpe'a						
0.2119	0.1960	0.1788	0.1701	0.1695	0.1679	0.1652
0.4198	0.2512	0.2473	0.2333	0.2173	0.2125	0.2070
0.3188	0.3050	0.3037	0.2907	0.2486	0.2375	0.2287
0.3994	0.3334	0.2757	0.2661	0.2637	0.2499	0.2352
0.2795	0.2497	0.2189	0.2093	0.2011	0.1959	0.1803
0.2860	0.2842	0.2094	0.2063	0.2002	0.1972	0.1831
0.3534	0.2859	0.2721	0.2535	0.2116	0.1812	0.1795
0.3866	0.3863	0.3224	0.3141	0.2958	0.2214	0.1729
0.3959	0.3507	0.2573	0.2297	0.2202	0.2181	0.1852
0.4589	0.2187	0.2007	0.1951	0.1648	0.1613	0.1490

Listing 12. Wskaźniki Sharpe'a dla najlepszych akcji pretendujących do włączenia w skład portfela inwestycyjnego.

Portfele 7 najlepszych akcji:						
27	8	4	17	11	28	18
28	18	27	23	9	8	4
28	1	24	8	2	6	15
8	24	28	25	23	2	19
12	25	2	21	20	19	23
12	20	25	6	9	18	19
20	6	13	12	19	2	26
13	6	21	26	28	23	12
13	20	21	6	28	22	23
13	28	21	23	27	8	19

Listing 13. Wskaźniki Sharpe'a dla najlepszych akcji pretendujących do włączenia w skład portfela inwestycyjnego.

Interesujący jest zwłaszcza obraz migracji najlepszych akcji. Z listingu 13 widać, że są akcje - efemerydy, które na krótko goszczą w zbiorze pretendentek do portfela np. akcja nr 27 pojawiła się dwa razy, w tym raz na pierwszym miejscu, akcja nr 1 tylko raz, są także akcje wielokrotnie występujące np. opisana wcześniej akcja nr 13, także 28, 4 czy 23. Można także zauważyć całkowitą absencję niektórych akcji, np. o numerach 3, 7, 10, 16.

W następnej kolejności przeprowadzono symulacje dla danych rozpoczynających się tuż po pierwszym fragmencie przestrzeni. Ponieważ zakresem badań w pierwszym okresie było 10 serii rozpoczynanych co 10 świec, to druga seria badań rozpoczynała się 100 świec po początkowej świecy pierwszej serii, czyli od świecy numer 140.

Tab. 16. Parametry symulacji dla drugiego okresu uczenia systemu.

Parametry symulacji	Wartości parametrów	Nazwa zmiennej w pliku
Nazwa pliku	gielda3deBlinlagS	
Liczba akcji w portfelu	5	
Nr świecy początkowej	140	IP
Odległość między świecami pocz.	10	
Liczba świec w symulacji	30	KK
Liczba symulacji w badaniu	15	LL
Zakres lokalnego zbioru uczącego	30	Zakr
Inkrement w przestrzeni modeli	6	Kr

Macierz B współczynników Sharpe'a uzyskana dla tego fragmentu danych przedstawiona jest w listingu 14. To macierz współczynników uporządkowana jest malejąco. Kolumny dopowiadają kolejnym akcjom, wiersze kolejnym symulacjom rozpoczynanym co 10 świec.

Najlepsze wskaźniki Sharpe'a						
0.2616	0.1797	0.1410	0.1226	0.1035	0.0965	0.0621
0.1785	0.1384	0.1200	0.0980	0.0885	0.0743	0.0632
0.3189	0.2423	0.2058	0.2025	0.1561	0.1102	0.1084
0.2069	0.1988	0.1797	0.1689	0.1649	0.1437	0.1420
0.2908	0.2808	0.2484	0.2479	0.1917	0.1749	0.1630
0.4423	0.3719	0.3485	0.2439	0.2311	0.2083	0.2017
0.5728	0.3688	0.2997	0.2624	0.2324	0.1922	0.1915
0.4061	0.2831	0.2205	0.2047	0.1952	0.1664	0.1441
0.2318	0.2010	0.1697	0.1156	0.1103	0.1005	0.0713
0.2181	0.1827	0.1716	0.1695	0.1639	0.1209	0.1134

Listing 14. Wskaźniki Sharpe'a dla najlepszych akcji pretendujących do włączenia w skład portfela inwestycyjnego w drugim odcinku symulacji.

Portfele 7 najlepszych akcji:						
13	27	20	8	28	18	5
11	21	14	22	18	1	12
18	11	14	16	1	7	9
7	16	13	1	18	24	11
20	10	7	25	18	8	6
20	6	10	13	8	7	23
20	22	10	6	28	8	27
20	22	28	10	24	2	23
5	22	8	23	25	18	2
15	8	18	1	26	22	7

Listing 15. Wskaźniki Sharpe'a dla najlepszych akcji pretendujących do włączenia w skład portfela inwestycyjnego w drugim odcinku symulacji.

Widać tu zbliżone do pierwszego okresu wartości w macierzy B i niestety zupełnie inne sekwencje indeksów akcji. Podobnie niepokojące wyniki uzyskano dla kolejnego trzeciego okresu symulacji, rozpoczynającego się od 240 świecy.

Tab. 17. Parametry symulacji dla trzeciego okresu uczenia systemu.

Parametry symulacji	Wartości parametrów	Nazwa zmiennej w pliku
Nazwa pliku	gielda3deBlinlagS	
Liczba akcji w portfelu	5	
Nr świecy początkowej	240	IP
Odległość między świecami pocz.	10	
Liczba świec w symulacji	30	KK
Liczba symulacji w badaniu	10	LL
Zakres lokalnego zbioru uczącego	30	zacr
Inkrement w przestrzeni modeli	6	kr

Podobnie jak w dwóch pierwszych okresach wyniki przedstawione są w postaci macierzy B oraz macierzy numerów akcji.

Najlepsze wskaźniki Sharpe'a:						
0.2468	0.2465	0.2018	0.1996	0.1967	0.1929	0.1766
0.2390	0.2110	0.2088	0.1863	0.1357	0.0777	0.0543
0.2105	0.1973	0.1626	0.1625	0.1501	0.1345	0.1255
0.3140	0.2777	0.2674	0.2097	0.2050	0.1875	0.1430
0.4127	0.3646	0.2212	0.2041	0.2039	0.1998	0.1844
0.3019	0.2654	0.2329	0.1788	0.1774	0.1601	0.1311
0.2667	0.2635	0.1873	0.1790	0.1736	0.1637	0.1392
0.4118	0.2870	0.1932	0.1449	0.1444	0.1378	0.1286
0.2958	0.2498	0.2494	0.1589	0.1212	0.0864	0.0840
0.2539	0.1591	0.1338	0.1071	0.1006	0.0840	0.0713

Listing 16. Wskaźniki Sharpe'a dla najlepszych akcji pretendujących do włączenia w skład portfela inwestycyjnego w trzecim odcinku symulacji.

Portfele 7 najlepszych akcji:						
27	26	18	1	3	22	16
13	9	27	26	3	1	17
13	1	9	18	27	7	26
2	1	13	24	7	21	26
2	13	1	21	5	18	22
13	21	16	1	2	15	22
20	16	21	6	23	8	22
23	20	16	12	6	21	11
23	20	12	11	14	28	21
23	20	26	21	24	14	4

Listing 17. Wskaźniki Sharpe'a dla najlepszych akcji pretendujących do włączenia w skład portfela inwestycyjnego w trzecim odcinku symulacji.

Wyniki jakościowe, czyli macierz B dla trzeciego okresu są średnio nieco gorsze. Także nie widać wyraźnej dominacji niektórych akcji nad innymi, jeżeli chodzi o częstość ich pojawiania się. Z wcześniejszych badań wynika, że akcja najlepsza w całym badanym okresie, czyli akcja nr 13, także nie gwarantuje stabilności zachowań w czasie, gdyż są takie okresy (wiersze macierzy), gdy w ogóle nie pojawia się ta najlepsza akcja w sugerowanym składzie portfela.

Celowość utrzymywania portfeli w niezmiennym składzie przez zbadany odcinek 10 świecowy będzie można ocenić dopiero po przetestowaniu tych składów w „przyszłej” przestrzeni danych, takich, które nie były wykorzystywane w fazie uczenia.

### 6.2.2. Faza weryfikacji systemu – sprawdzanie skuteczności predykcji

Przeprowadzone badania symulacyjne w trzech różnych fragmentach szeregu czasowego portfela akcji umożliwiają weryfikację poprawności tezy, że charakterystyczne cechy zachowania się portfela w całości dadzą podstawę do udanej predykcji na niewielkim, chociaż odcinku wprzód. Przez ten niewielki odcinek predykcyjny rozumie się tu 2-3 kroki próbkowania szeregu czasowego, czyli 2-3 świece rozpatrywanego szeregu. Precyzyjniej, jest to jednak predykcja na jeden okres wprzód zweryfikowana dodatkowo predykcją z poprzedniego kroku. Te dwie predykcje, o ile są zgodne, pozwalają na podjęcie decyzji inwestycyjnej o otwarciu pozycji – zakupie portfela akcji o ustalonej strukturze. Struktura ustalana jest na początku serii 10 symulacji.

Przedstawione dalej wyniki dotyczą właśnie weryfikacji poprawności tak postawionej tezy. Przeprowadzone badania poprzedniego podrozdziału pozwalają na określenie struktur portfeli, a poniżej weryfikuje się ich poprawność w nowej przestrzeni danych, z której ani jeden wiersz nie był użyty do zbudowania modelu portfela.

Podobnie jak w powyższym podrozdziale testy podzielone są na trzy obszary w szeregu czasowym określonym przez macierz danych. Otóż, po zakończeniu fazy uczenia na wierszach rozpoczynających się od wiersza, kolejno nr 40, 50, 60... aż do 130 i prowadzonych do wiersza odpowiednio nr 70, 80, 90 itd. następuje testowanie rozpoczynające się od wiersza następującego tuż po zakończeniu uczenia, a więc odpowiednio od 70, 80, 90 itd. Te trzy odrębne obszary wydzielono z uwagi na, wspomniany już, wielogodzinny czas obliczeń niezbędny dla przeprowadzenia obliczeń w jednym obszarze.

W tab. 17 przedstawiono parametry symulacji pierwszego okresu badań zawartego pomiędzy świecami nr 70 i 170. Są to kontynuacje badań następujące tuż po wykonaniu skryptów uczących, z których każdy rozpoczynał się kolejno 30 świec wcześniej. W każdym pliku testowym wykorzystywana była część przestrzeni danych historycznych, taka, która wcześniej nie była wykorzystana w fazie uczenia. (Badania pliku `gielda3deBlin5normLs` od 70 świecy (kontynuacja badań z plikiem uczącym `gielda3deBlinlagS` na 30 świecach rozpoczynającym symulacje od 40 świecy)).

Wyniki symulacji o parametrach jak w tab. 18 przedstawione są w tab. 19. W tabeli tej kolejno podana jest informacja o liczbie wykonanych symulacji rozpoczynających się co 10 świec. Było to 10 takich symulacji, w czasie których na odcinku 15 świec nie zmieniano struktury



portfela (oznaczało to, że kolejną serię rozpoczynano przed zakończeniem poprzedniej). Zakładano, że do dyspozycji inwestora jest kwota ok. 10000 PLN i na portfel składano w przybliżeniu jednakowe wartościowo pakiety akcji (po ok. 2000 PLN) o numerach przedstawionych w listingu 13. Użyte przybliżenie „około” oznacza, że inwestowana była kwota wynikająca z iloczynu całkowitej liczby akcji i jej bieżącej ceny.

Podsumowując całe 10 serii symulacji uzyskano wynik rzędu 3000 PLN (z zainwestowanych 10000 PLN) po 100 dniach badań (bez uwzględnienia prowizji maklera).

Biorąc pod uwagę zmienność wyniku w poszczególnych seriach uzyskano średnia wartość wyniku w jednej serii równy ok. 300 PLN a wskaźnik Sharpe’a dla całego okresu badań (serii 10 symulacji) nieco poniżej 1.0. Nałożone na siebie wykresy krzywych narastania kapitału na wspomnianym odcinku 15 kroków wprzód dla 10 serii po 5 akcji w każdym portfelu przedstawiono na rys. 97.

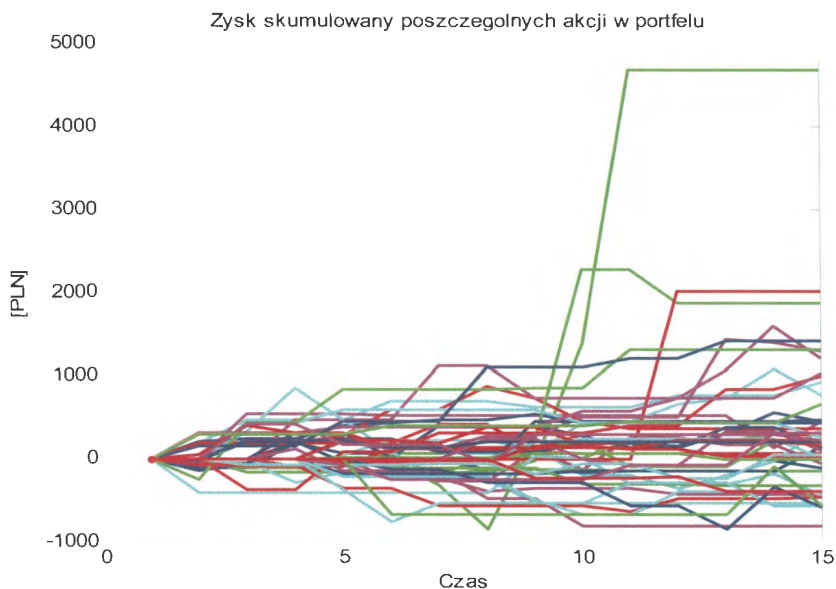
Tab. 18. Parametry symulacji weryfikującej poprawność modeli pierwszego okresu przestrzeni danych.

Parametry symulacji	Wartości parametrów	Nazwa zmiennej w pliku
Nazwa pliku	gielda3deBlin5normLS	
Liczba akcji w portfelu	5	
Nr świecy początkowej	70, 80, 90, ...	IP
Odległość między świecami pocz.	10	
Liczba świec w symulacji	15	KK
Liczba symulacji w badaniu	5	LL
Zakres lokalnego zbioru uczącego	30	Zakr
Inkrement w przestrzeni modeli	6	Kr

Tab. 19. Wyniki symulacji testującej poprawność pierwszego okresu przestrzeni danych.

Obserwowany wynik	Wartości
Liczba symulacji	10
Wartość średnia	303.60
Odchylenie standardowe	317.36
Wskaźnik Sharpe’a	0.9566
Wynik końcowy [PLN]	3036.10

Każda akcja obserwowana jest przez 15 okresów – świec. Wykresów zakończonych wynikiem dodatnim jest zdecydowanie więcej niż zakończonych stratą.



Rys. 97. Wykresy narastania kapitału dla poszczególnych akcji w kolejnych 10 portfelach inwestycyjnych. W serii 10 kolejnych badań symulacyjnych otwieranych co 10 świec i obserwowanych przez 15 świec. Ponieważ w każdym portfelu jest 5 akcji więc na wykresie obserwuje się ogólny obraz 50 przebiegów.

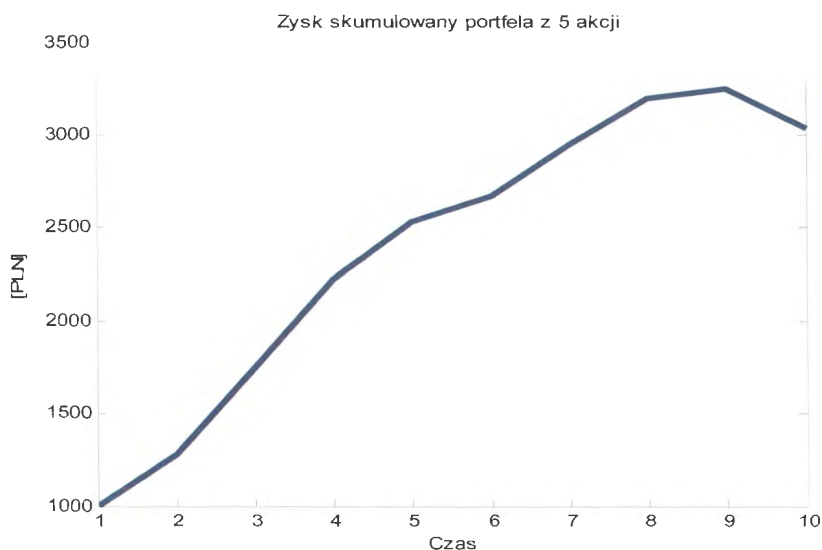
Wyniki symulacji dla portfela 5 akcji dla kolejnych 10 realizacji co 10 świec, z okresem symulacji wynoszącym 15 świec od 70 świecy poczynając. Na wykresie występuje 50 krzywych narastania kapitału w 15 krokach. Kolor krzywej oznacza jej pozycję w portfelu – np. kolor niebieski to pozycja pierwsza (o najwyższym wskaźniku Sharpe'a, kolor cjan – druga itd.). W każdym ustalonym kolorze na wykresie jest 10 krzywych – tyle, ile symulacji wykonywanych co 10 świec. Każda symulacja wykonywana jest w horyzoncie 15 kroków.

Wektor wyników cząstkowych [x 1000 PLN], uzyskiwanych po każdej zakończonej symulacji złożonej z 15 kroków wykonanych co 10 świec dobowych, o średniej równej 303.60 PLN i odchyleniu standardowym równym 317.36 PLN przedstawiał się następująco:



1.0025 0.2818 0.4677 0.4733 0.3055 0.1401 0.2776 0.2490  
0.0533 -0.2148

W wektorze występuje tylko jeden przypadek straty. Wykres narastania kapitału zbudowany na tym wektorze przedstawiony jest na rys. 98.



Rys. 98. Krzywa narastania kapitału – 5 akcji w portfelu, od 70 poczynając, 10 symulacji każda na 15 świecach.

Wykres ten podsumowuje pierwszy okres symulacji testującej. Uzyskany wynik jest bardzo dobry, daje doskonały wskaźnik Sharpe'a i może mieć duże znaczenie praktyczne, o ile nie naruszy tego obrazu ewidentne zakłócenie o naturze silnej niestacjonarności w kolejnych dwóch badanych okresach.

W drugim okresie testowania rozpoczynającym się od 170 świecy (tj. następującym bezpośrednio po fazie uczenia rozpoczynającej się od 140 świecy i kończącej się właśnie na 170) badania prowadzone były jak poprzednio – w 10 seriach symulacji rozpoczynających się co 10 świec. Każda seria była testowana na odcinku 15 świec, więc częściowo nachodziła na odcinek testowania kolejnej serii. Parametry tego testu przedstawione są w tab. 20.

Tab. 20. Parametry symulacji weryfikującej poprawność modeli drugiego okresu przestrzeni danych.

Parametry symulacji	Wartości parametrów	Nazwa zmiennej w pliku
Nazwa pliku	gielda3deBlin5normLs	
Liczba akcji w portfelu	5	
Nr świecy początkowej	170	ipocz
Odległość między świecami pocz.	10	
Liczba świec w symulacji	15	KK
Liczba symulacji	10	LL

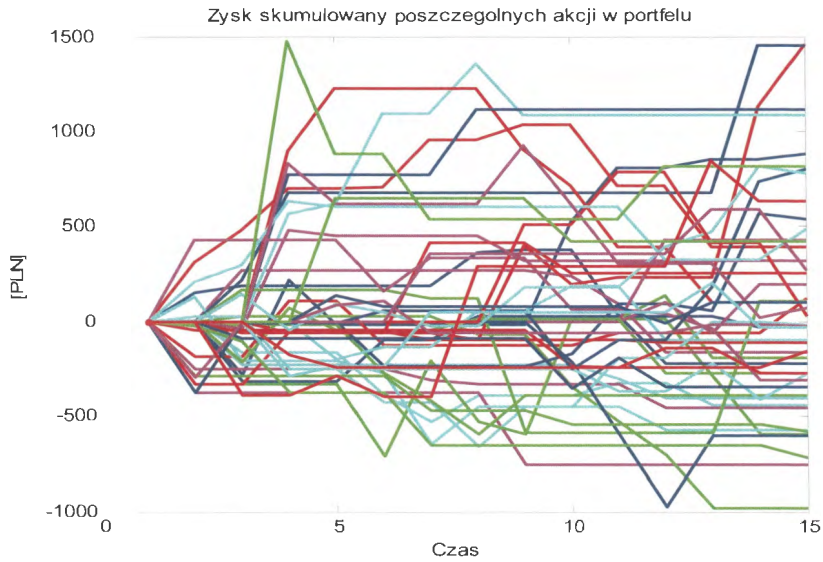
Wyniki tej symulacji przedstawione są w tab. 21. Wynik jest zdecydowanie gorszy niż w pierwszym okresie. Warto także zauważyć (na rys. 100), że występują tu trzy kolejne porażki dające w sumie dość istotne obsunięcie kapitału. W sumie jednak wskaźnik Sharpe'a na poziomie 0.27 jest ciągle dość wysoki, podobnie jak pewnie dodatni wynik końcowy. Rys. 99 daje inny obraz symulacji narastanie kapitału poszczególnych akcji w porównaniu z rys. 97. Na poprzednim obserwowano wyraźną asymetrię na korzyść przebiegów dodatnich, na rys. 99 tej asymetrii praktycznie nie widać.

Tab. 21. Wyniki symulacji testującej poprawność drugiego okresu przestrzeni danych.

Obserwowany wynik	Wartości
Liczba symulacji	5
Wartość średnia	97.61
Odchylenie standardowe	356.90
Wskaźnik Sharpe'a	0.2735
Wynik końcowy [PLN]	976.14

Wektor wyników cząstkowych, uzyskiwanych po każdej zakończonej symulacji złożonej z 15 kroków wykonanych co 10 świec dobowych, o średniej równej 97.61 PLN i odchyleniu standardowym równym 356.90 PLN przedstawiał się następująco:

-111.6233 -275.8115 -288.8245 330.3665 730.3525 336.9555 -  
390.4330 285.1755 266.9260 93.0535



Rys. 99. Wyniki symulacji dla portfela 5 akcji dla kolejnych 10 realizacji co 10 świec, z okresem symulacji wynoszącym 15 świec od 170 poczynając.



Rys. 100. Krzywa narastania kapitału dla drugiego okresu testowania – 5 akcji w portfelu, 10 realizacji po 15 świec w każdej, od 170 poczynając. Wynik jest zdecydowanie najgorszy.

Wyniki symulacji dynamicznie zmieniającego się portfela 5 akcji w końcowym odcinku badań, od 270 świecy, podobnie jak poprzednio – 10 serii co 10 świec, każda realizowana na odcinku 15 kroków przedstawione są w kolejnych tabelach. W tab. 21 przedstawiono parametry symulacji trzeciego okresu a w tab. 22 jej wyniki.

Tab. 22. Parametry symulacji weryfikującej poprawność modeli trzeciego okresu przestrzeni danych.

Parametry symulacji	Wartości parametrów	Nazwa zmiennej w pliku
Nazwa pliku	gielda3deBlin5normLs	
Liczba akcji w portfelu	5	
Nr świecy początkowej	270	Pocz.
Odległość między świecami pocz.	10	
Liczba świec w symulacji	15	KK
Liczba symulacji	10	LL

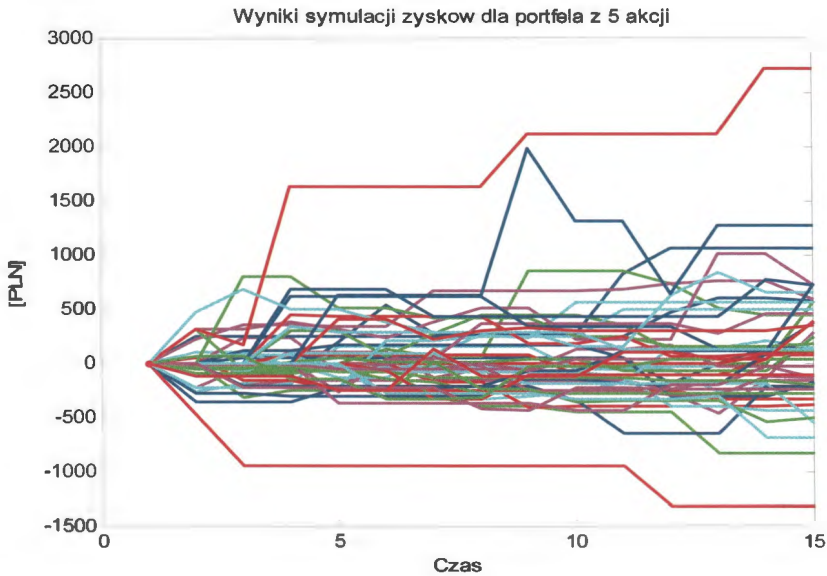
Wyniki te są lepsze od poprzednich i nieco gorsze od wyników pierwszego okresu.

Tab. 23. Wyniki symulacji testującej poprawność pierwszego okresu przestrzeni danych.

Obserwowany wynik	Wartości
Liczba symulacji	10
Wartość średnia	160.82
Odchylenie standardowe	178.01
Wskaźnik Sharpe'a	0.9035
Wynik końcowy [PLN]	1608.30

Wektor wyników cząstkowych uzyskiwanych po każdej zakończonej symulacji złożonej z 15 kroków wykonanych co 10 świec dobowych, o średniej równej 160.82 PLN i odchyleniu standardowym równym 178.01 PLN przedstawiał się następująco:

179.3245    253.5387    238.0940    -77.4825    -202.0402    347.1975  
 373.9645    137.6280    194.3495    163.7168



Rys. 101. Wyniki symulacji dla portfela 5 akcji dla kolejnych 10 realizacji co 10 świec, z okresem symulacji wynoszącym 15 świec od 270 poczynając.



Rys. 102. Krzywa narastania kapitału dla danych – 5 akcji w portfelu, 10 realizacji po 15 świec w każdej, od 270 poczynając.

Tym razem występują dwa ujemne wyniki cząstkowe, ponownie obok siebie. Na rys. 101 podobnie jak w pierwszym okresie daje się zauważyć pozytywna asymetria przebiegów, a na rys. 102 przedstawione jest narastanie kapitału dla trzeciego okresu symulacji.

### 6.2.3. Konkluzje

Wszystkie trzy rozpatrzone odcinki szeregu czasowego (od 70 świecy do 170, potem do 270 i 370) dają zróżnicowane rezultaty przedstawione powyżej. Po połączeniu wyników w jeden szereg czasowy przedstawiają się one jak na wykresie rys.103.



Rys. 103. Wykres narastania kapitału w całym okresie symulacji, od świecy nr 70 do 370, z inkreментом przeszukiwania modeli  $kr=6$ , z każdorazowym rozpatrywaniem 15 kroków wprzód, uruchamianych co 10 kroków. Każda jednostka na osi zmiennej niezależnej oznacza 10 świec dobowych, czyli 10 dni.

Uzyskano wyniki:

Średni przyrost kapitału na jedna symulacje (10 dni) = 187.34 PLN przy zaangażowaniu kapitału wynoszącym 10000 PLN.

Odchylenie standardowe dla serii 30 symulacji o średniej jw. = 297.15 PLN

Wskaźnik Sharpe'a dla wszystkich symulacji  $SR = 0.6305$ . Należy przypomnieć, że są to wyniki uzyskane przy charakterystycznym dla GPW ograniczeniu otwierania wyłącznie długich pozycji.

Przy tym ograniczeniu – uprawnieniu wyłącznie jednostronnego kierunku inwestowania, wynik należy uznać za doskonały. Gdyby dopuszczalne było na GPW powszechne otwieranie pozycji w obu kierunkach – długim i krótkim, to rezultaty transakcji zawieranych z wykorzystaniem predykcji opartych na GMDH byłyby zdecydowanie lepsze. Należy zakładać, że tak się stanie. Należy zakładać także, że zdecydowanie zmniejszy się przeciętna prowizja maklerska w Polsce, związana z płynnością akcji czy ich derywatów. Wówczas zaprezentowana w tym rozdziale strategia portfelowa może stać się bazą dla stabilnych systemów tradingowych.

Tak więc GPW jest jeszcze mało cywilizowaną platformą transakcyjną z bardzo dużymi prowizjami. Nadzieja dla GMDH to predykcja na co najmniej kilkudobowych świecach i zarządzaniu portfelami walorów. Podejmowanie prób predykcji na podstawie wyłącznie analizy technicznej dla pojedynczej akcji jest raczej skazane na niepowodzenie. Także w obrocie derywatami a nie instrumentem bazowym (pomimo tego, że są niższe koszty). Zagrozeniem będzie niższa płynność.





## 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 opisanie 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 tysiące 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*, Icijer 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 Upravlenie 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 upravlenija*. 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.



## Bibliografia

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. Kruttsinger 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. Nałęcza *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)**