



Instytut Badań Systemowych

Polskiej Akademii Nauk

---

*Seria:*

**BADANIA SYSTEMOWE**

**TOM 65**

Redaktor Naukowy

Prof. dr hab. inż. Jakub Gutenbaum

## **Rada Redakcyjna:**

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



Instytut Badań Systemowych  
Polskiej Akademii Nauk

**Antoni Wiliński**

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

**Warszawa - Szczecin**

Copyright © by Instytut Badań Systemowych PAN  
Warszawa 2009

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

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

### **Recenzenci:**

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

### **Redaktor techniczny:**

Dr inż. Anna Samborska-Owczarek

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

ISBN 9788389475237  
**ISSN 0208-8029**



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

## Wprowadzenie

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

---

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

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

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

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

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

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

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

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

---

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

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

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

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

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



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

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

Większość uznanych traderów<sup>6</sup> takich jak Larry Williams, Jack Schwager, Joe 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 selekcyjną np. roślin czy zwierząt, którzy pozostawiają zawsze kilka-kilkanaście najlepszych par czy osobników. W ten sposób uzyskuje się niejako szerszą bazę odwzorowania rzeczywistości. Nie wszystkie najlepsze cechy mogą być przekazane przez pojedynczy egzemplarz (przez najlepszy model rzeczywistości). Ta chwilowa dominacja może być przypadkowa i w wielokrotnie powtarzanej predykcji może się nie sprawdzić.

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

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

---

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

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

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

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

---

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

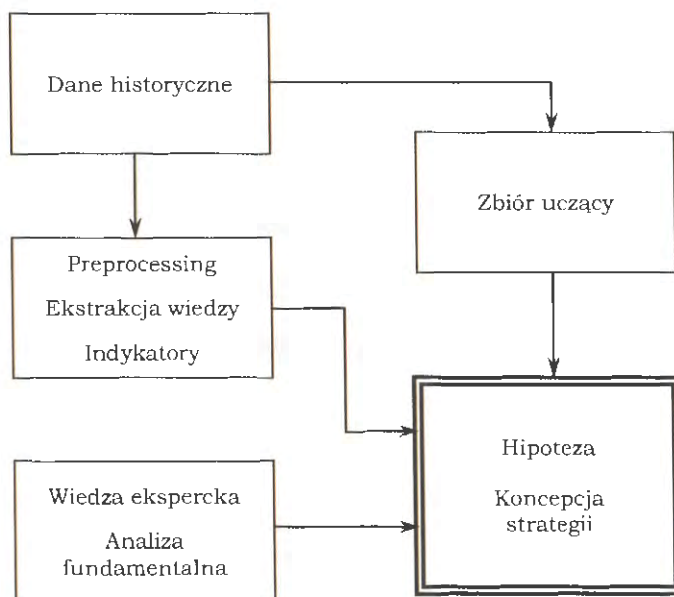
<sup>9</sup> Group Method of Data Handling – angielskie tłumaczenie metody opracowanej przez Ivakhnenkę w oryginale zwane *Metod Gruppovogo Uceta Argumetov* (ros.). W języku polskim nie ma dotąd jednoznacznego przyjętego powszechnie tłumaczenia. Właściwe wydaje się być – Metoda Grupowania Atrybutów czy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

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

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

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

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



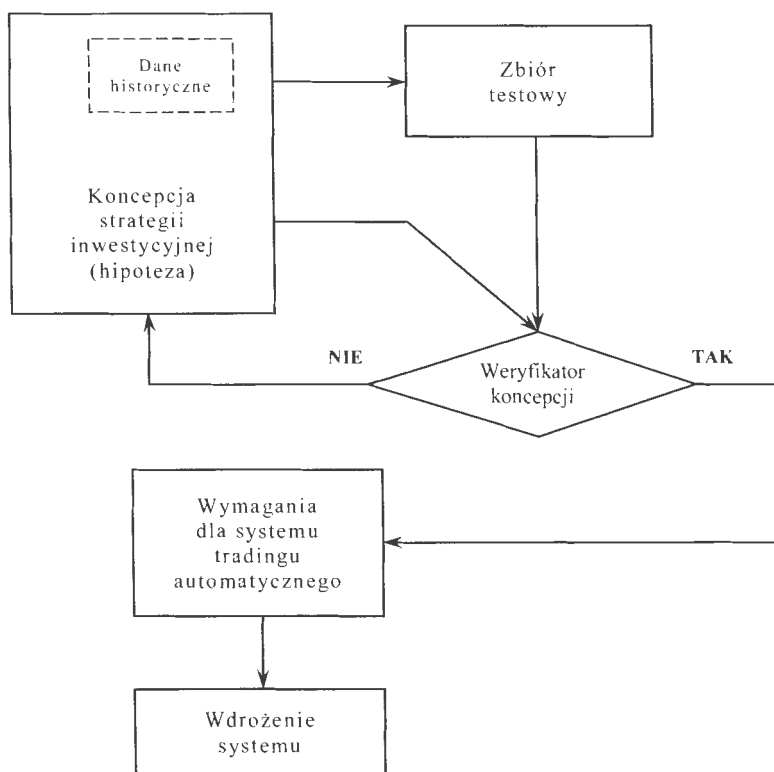
Rys. 1. Schemat blokowy procesu powstawania hipotez badawczych dotyczących koncepcji strategii inwestycyjnych.

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

## Wprowadzenie

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

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



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

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

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

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

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

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

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

---

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

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.





# 5

## Predykcyjne algorytmy GMDH oparte na przestrzeni czynników behawioralnych

*Prawdą jest to, co wytrzyma próbę doświadczenia*  
Albert Einstein (1879-1955)

Wśród zagorzałych zwolenników analizy technicznej powszechne jest przekonanie, że każdy obserwowany rynek uwzględnia przyszłe zdarzenia (ekonomiczne, polityczne) poprzez poprzedzanie każdej, dużej i małej, decyzji inwestycyjnej rozważaniami uwzględniającymi całą dostępną inwestorowi informację o przeszłości i przyszłości, skażonymi także różnymi indywidualnymi emocjami i uprzedzeniami. To co cały świat obserwuje jest sumą tych rozważań, emocji i podjętych decyzji, gdyż w istocie rzeczy, każda z tych decyzji jest rezultatem przetwarzania dostępnej informacji i sposobu jej interpretacji. Interesujące może być rozpatrywanie bieżących zachowań uczestników rynku, nie poprzez obserwację kursu waluty, lecz poprzez badanie przyczyn zmian tego kursu wynikających ze złożonych zleceń i otwartych pozycji na rynku spot [20].

## 5.1. Charakterystyka źródła danych

W światowym Internecie odszukać można wiele źródeł informacji dającej się zakwalifikować do analizy fundamentalnej. Niektóre z nich mogą mieć duże znaczenia dla rozwoju systemów handlu automatycznego. Np. bardzo charakterystyczne jako dodatkowe dynamiczne czynniki oddziałujące na ocenę instrumentu finansowego są tzw. *economic events* związane z ogłaszaniem rozmaitych komunikatów o przewidywanym stanie wskaźników ekonomicznych danego kraju (im ważniejszy kraj, tym zapewne większe przewidywane oddziaływanie na globalne instrumenty finansowe). Analizę fundamentalną od technicznej odróżnia podejście do rzeczywistości – analiza fundamentalna skupia się na badaniu przyczyn, a nie skutków, jak czyni się to stosując analizę techniczną. Przykładem takiego właśnie podejścia jest doszukiwanie się przyczyn zmian ceny waloru rynkowego w aktywnych postawach inwestorów. W najprostszym przypadku pojedynczej akcji, im większy będzie na nią popyt, tym wyższa będzie jej cena. Gdyby więc można było rozpoznać zamiary inwestorów, co do zakupu lub sprzedaży waloru, to można byłoby objaśnić przyczynę przyszłej zmiany ceny. Ważne byłoby przy tym, by rozpoznać zamiary możliwie dużej grupy uczestników rynku.

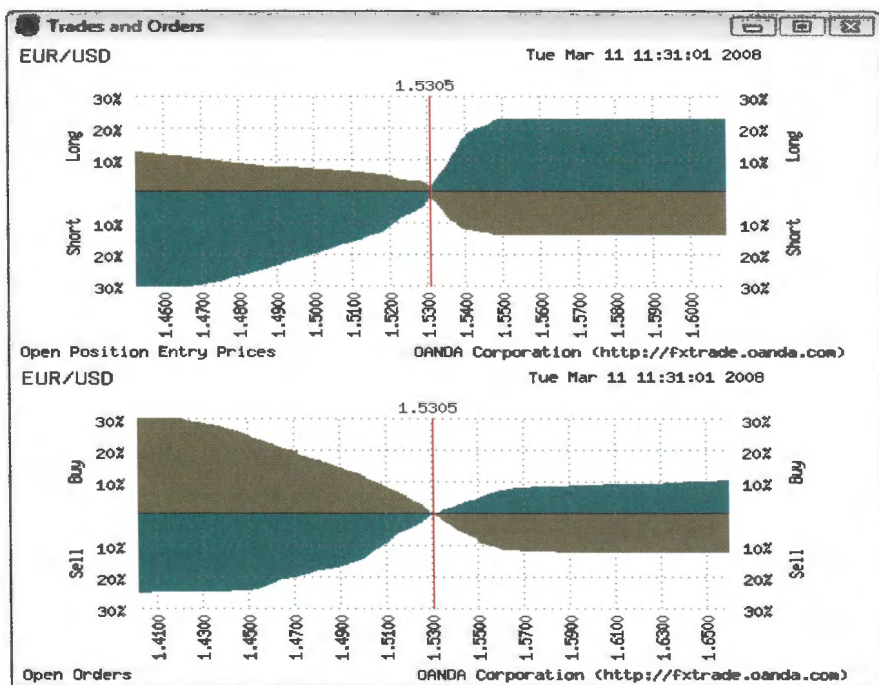
Przykładem platformy prezentującej behawioryzm jej użytkowników jest aplikacja *Small Speculant Sentiment Analyzer* [140] przedstawiająca intencje wybranej części użytkowników znanej platformy brokerskiej oanda.com [138]. Aplikacja ta to kolejny przykład możliwości zautomatyzowania procesu podejmowania decyzji na rynkach dostępnych przez Internet. Wspomniana aplikacja wizualizuje zamiary kilkudziesięciu tysięcy użytkowników w formie dwóch diagramów, z których górny przedstawia rozkład otwartych pozycji rynku spot, a dolny rozkład zleceń (rys. 87). Każdy kwadrant diagramu może być zinterpretowany jako krzywa całkowita liczby decyzji (otwartych pozycji lub zleceń) w funkcji odległości od bieżącej wartości rozpatrywanego instrumentu finansowego - na przedstawionym wykresie na przykładzie eurodolara.

Kolejno, np. diagram w górnej prawej ćwiartce (kwadrancie) oznacza narastającą liczbę (względną, wyrażoną w procentach) pozycji długich otwartych powyżej bieżącej wartości kursu. Wykres przedstawia sytuację na rynku EUR/USD przy wartości bieżącej kursu  $C = 1,5305$ . Każda wartość (w %) w prawo od środka układu oznacza narastającą względną liczbę pozycji długich otwartych pomiędzy środkiem układu (wartością bieżącą kursu) a daną rozpatrywaną wartością kursu. Są to oczywiście pozycje w danej chwili przegrane, gdyż położone są powyżej

wartości bieżącej. Na wykresie widać, że wszystkich pozycji długich położonych powyżej wartości bieżącej jest ok. 24%.

Podobnie rozpoznać można, że wygrywających pozycji krótkich jest w rozpatrywanej chwili ok. 13% (dolny prawy kwadrant), pozycji wygrywających długich (górny lewy kwadrant) jest ok. 14%, a pozycji przegrywających krótkich (dolny lewy) ok. 35%. Wskutek niedoskonałości wykresu wielkości te nie sumują się do 100%. Wykres jest obciążony i w takim stanie przedstawiany jest w Internecie [138]. Z punktu widzenia prezentowanej tu koncepcji istotna jest jednak wzajemna relacja pomiędzy czterema rozpatrywanymi kwadrantami.

Dla sformalizowania tych obserwacji niech  $x_b$  oznacza wartość bieżącą kursu rozpatrywanej pary walutowej, a  $G_1(x_b, x_b + \delta)$  oznacza udział pozycji długich otwartych w przedziale  $< x_b, x_b + \delta >$  w całkowitej liczbie wszystkich otwartych pozycji (długich i krótkich). Niech  $\delta$  oznacza pewną, racjonalnie wybraną, stałą wielkość rozpatrywanego przedziału (np. 100, 200 lub więcej pipsów). Zmienna ta winna umożliwić w miarę dokładne odczytanie udziału otwartych pozycji z wykresu (rys. 87).



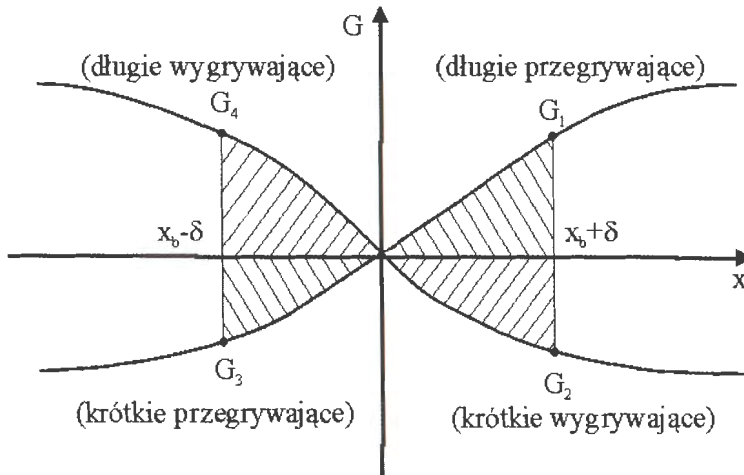
Rys. 87. Diagram liczby otwartych pozycji (górny) i złożonych zleceń (dolny).

Przez analogię, odpowiednio (rys. 88):

$G_2(x_b, x_b + \delta)$  – oznaczać będzie udział pozycji krótkich otwartych powyżej wartości bieżącej kursu w przedziale  $\langle x_b, x_b + \delta \rangle$ ;

$G_3(x_b - \delta, x_b)$  – oznaczać będzie udział pozycji krótkich otwartych poniżej wartości bieżącej kursu w przedziale  $\langle x_b - \delta, x_b \rangle$ ;

$G_4(x_b - \delta, x_b)$  – oznaczać będzie udział pozycji długich otwartych w przedziale  $\langle x_b - \delta, x_b \rangle$ , czyli poniżej wartości bieżącej kursu.



Rys. 88. Definicja udziałów  $G_i$  w odległości  $\delta$  od wartości bieżącej kursu  $x_b$ .

Wprowadzenie parametru  $\delta$  powoduje, że suma czterech rozpartzonych całek  $G_1$  do  $G_4$  nie będzie równa 100%, lecz pozwoli na powtarzalność obliczeń pól w tych samych warunkach.

Podobnie dla nierozpatrywanego do tej pory diagramu dolnego na rys. 87 niech odpowiednio:

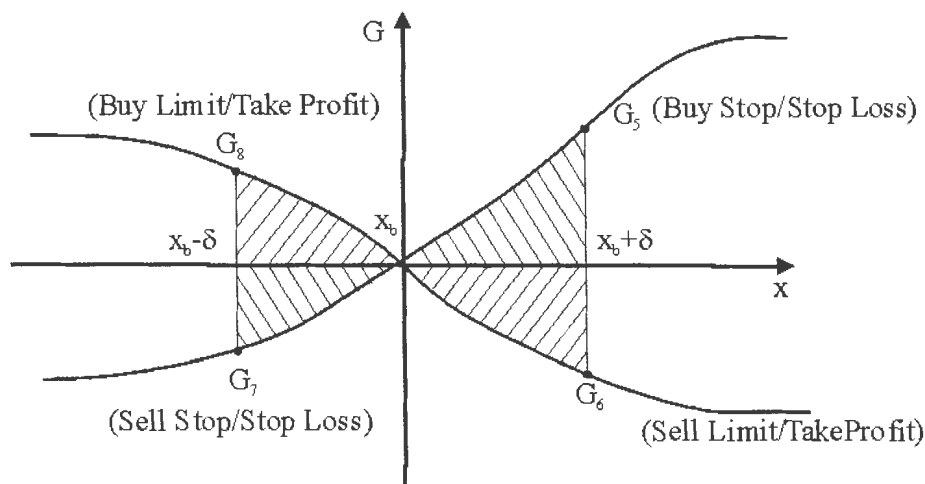
$G_5(x_b, x_b + \delta)$  oznacza udział zleceń kupna złożonych z limitem położonym w przedziale  $\langle x_b, x_b + \delta \rangle$ . Zlecenia te mogą wynikać z zamian otwarcia pozycji długiej powyżej bieżącej wartości kursu (są to tzw. zlecenia *Buy Stop*), lub z zamiaru zamknięcia przegrywającej pozycji krótkiej (zlecenie *Stop Loss*). Oba te przypadki są rezultatem takiej samej postawy inwestora – jego przekonania, że kurs waluty wzrośnie. W pierwszym przypadku otwarcie pozycji długiej następuje w nadziei dalszego wzrostu, w drugim przypadku zamknięcie pozycji krótkiej (technicznie tożsame z otwarciem pozycji długiej) następuje w obawie także przed dalszym wzrostem kursu i powiększaniem straty.

$G_6(x_b, x_b + \delta)$  oznacza udział zleceń sprzedaży złożonych z limitem położonym powyżej bieżącej wartości kursu. Są to zlecenia wynikające z zamiaru zamknięcia pozycji długiej wygrywającej (*Take Profit*) lub otwarcia pozycji krótkiej. Oba wynikają z takiego samego przekonania inwestorów o spadku kursu (po osiągnięciu poziomu wyznaczonego w zleceniu). Inwestor w przypadku posiadania pozycji długiej składa zlecenie typu *Take Profit* lokując je powyżej bieżącej wartości kursu w przekonaniu, że kurs dalej nie będzie rósł. Inny inwestor składa zlecenie otwarcia pozycji krótkiej *Sell Limit* także w przekonaniu, że kurs po osiągnięciu poziomu określonego w zleceniu zacznie spadać.

$G_7(x_b - \delta, x_b)$  to udział zleceń sprzedaży złożonych poniżej wartości bieżącej kursu. Zlecenia te wynikają z zamiaru inwestora otwarcia pozycji krótkiej (*Sell Stop*) lub zamknięcia przerywającej pozycji długiej (*Stop Loss*). Oba więc wynikają z przekonania, że kurs będzie spadał.

$G_8(x_b - \delta, x_b)$  to udział zleceń kupna złożonych poniżej wartości bieżącej kursu. Zlecenia takie są składane przez inwestora w celu otwarcia w celu otwarcia pozycji długiej (*Buy Limit*) lub dla zamknięcia krótkiej pozycji wygrywającej (*Take Profit*). Oba typy zleceń wynikają z tego samego przekonania inwestorów, że kurs po osiągnięciu poziomu wyznaczonego przez zlecenie, będzie następnie rósł.

Całki  $G_5$  do  $G_8$  dotyczące zleceń przedstawione są na rys. 89. Przedział  $\delta$  (odległość od wartości bieżącej kursu) może to być oczywiście inny niż na rys. 88.



Rys. 89. Definicja udziałów zleceń w odległości  $\delta$  od wartości bieżącej kursu  $x_b$ .

Jeżeli w powyższych definicjach  $G_i$  oraz  $\delta$  wprowadzić pojęcie skończenie małego przyrostu kursu  $dx$  nie mniejszego niż 1 pips oraz pojęcie rozkładu gęstości  $g_i(x)$  złożonych zleceń lub otwartych pozycji, to ogólna definicja całki  $G_i$  będzie:

$$G_i = \int_{x_b}^{x_{b+\delta}} g_i(x) dx \quad (64)$$

dla  $i = 1, 2 \dots 8$ ,  $\delta = 100, 200, 300 \dots$  pipsów.

Każdy indeks  $i$  oznacza zupełnie inną funkcję, w aspekcie systemu transakcyjnego, inny świat – zwycięstw lub porażek, optymizmu lub pesymizmu.

Dla pewnego ustalonego arbitralnie, aczkolwiek rozsądnie  $\delta$ , można przeprowadzić normalizację całek  $G_i$  w dwóch grupach – dla  $i \in \{1, 2, 3, 4\}$  oraz  $i \in \{5, 6, 7, 8\}$ , taką że:

$$G_i^n = \frac{G_i}{\sum_1^4 G_i} \quad (65)$$

dla  $i = 1, 2 \dots 4$ , oraz:

$$G_i^n = \frac{G_i}{\sum_4^8 G_i} \quad (66)$$

dla  $i = 5, 6 \dots 8$ .

Zabiegi powyższe mają na celu wyodrębnienie takiego fragmentu obrazu (rys. 87), który w sposób powtarzalny, dla wielu tysięcy prób z serwera, zapewni zawsze tę samą procedurę znajdowania proporcji pomiędzy otwartymi pozycjami w czterech kwadrantach górnej części rys. 87 oraz złożonych zleceń w dolnej części tego rysunku a dodatkowo zapewni stałą wartość sumy całek.

Po przeprowadzeniu normalizacji wg (65) i (66) przygotowano do dalszego przetwarzania macierzy danych obserwowanych  $X_{M \times N}$  o następującej strukturze:

M – liczba wierszy (pomiarów, obserwacji – tu świec obserwowanych w cyklu jednogodzinnym);

N – liczba kolumn, tu  $N = 10$ .

W kolumnach macierzy  $X$  zapisano odpowiednio:

$X_1$  – kurs badanej co godzinę pary walutowej EUR/USD;

$X_2 \div X_5$  - kolejno obserwowane na podstawie górnego diagramu (rys. 87) i przetworzone zgodnie z (2) wartości udziałów  $G_1^4 \div G_4^4$ , oznaczające względną liczbę pozycji otwartych w rozpatrywanych kwadrantach układu rys. 2;

$X_6 \div X_9$  - obserwowane na podstawie dolnego diagramu (rys. 87) i przetworzone wg (66) wartości udziałów złożonych zleceń  $G_5^4 \div G_8^4$ , oznaczające względną liczbę w kwadrantach rys. 3.

$X_{10} \equiv Y$  - obserwowana zmienna wyjściowa - różnica pomiędzy bieżącą a poprzednią wartością kursu, stanowiąca podstawę do rozważań predykcyjnych. Formalnie,  $X_{i,10} = X_{i+1,1} - X_{i,1}$ .

Fragment tak zbudowanej macierzy danych przedstawiono w tab. 13 i tab. 14. Tabele te (macierz danych X) są bazą dla dalszych rozważań i zbudowania modelu predykcyjnego.

Tab. 13. Przykładowe dane obserwowane w aplikacji Small Speculant i przetworzone wg (65) dla całek związanych z otwartymi pozycjami.

Kurs EUR/USD	$G_1$	$G_2$	$G_3$	$G_4$	Zwrot
1.5305	14.7323	31.6047	21.1190	32.5439	-0.0004
1.5345	14.2370	31.4428	21.2531	33.0672	0.0040
1.5334	13.9713	32.1354	20.8157	33.0776	-0.0011
1.5326	14.1918	32.3827	20.8884	32.5371	-0.0008
1.5324	14.1873	32.3527	21.2208	32.2391	-0.0002
1.5326	15.7094	31.3924	20.9195	31.9787	0.0002

Tab. 14. Przykładowe dane obserwowane w aplikacji Small Speculant i przetworzone wg (66) dla całek związanych ze zleceniami.

Kurs EUR/USD	$G_5$	$G_6$	$G_7$	$G_8$	Zwrot
1.5305	35.6927	11.2699	22.0578	30.9796	-0.0004
1.5345	25.9228	22.8024	10.5920	40.6829	0.0040
1.5334	33.5726	13.5982	15.9851	36.8442	-0.0011
1.5326	33.7731	12.2998	19.1508	34.7763	-0.0008
1.5324	32.3408	14.2092	17.6952	35.7549	-0.0002
1.5326	31.1768	15.6575	16.3646	36.8011	0.0002



## 5.2. Sformułowanie problemu badawczego

Przedstawiona aplikacja internetowa jest zgodnie z jej sugerowaną nazwą (*Small Speculant*) źródłem wielu pomysłowych rozwiązań i miejscem ścierania się różnych poglądów spontanicznie powstającej wokół niej społeczności. Wiele z nich koncentruje się, jak to wynika z dyskusji na forum, na próbie realizacji tezy, że skoro ok. 95% wszystkich drobnych inwestorów przegrywa, to należy poszukiwać wzorca rozkładu udziałów  $G_1 \div G_8$  takiego, że prawdopodobna i możliwa będzie porażka tak znacznej części populacji inwestorów reprezentowanych poprzez bieżący rozkład otwartych pozycji i zleceń.

Rozpatrywana tu autorska koncepcja jest całkowicie odmienna od wyżej przedstawionej. Zbieżne są one jedynie w aspekcie założenia, że istnieją pewne wzorce wyrażane poprzez konfiguracje danych, których rozpoznanie umożliwi lub uprawdopodobni skuteczność prognozy. W przypadku opisanym jako popularny sposób zachowania się większości inwestorów wzorcem takim mogłaby być sytuacja na diagramach rys. 87 – jako informacja wejściowa dla znalezienia odpowiedzi na pytanie jaka możliwa najbliższa zmiana rynku spowodowałaby porażkę większości graczy. Na diagramie tym dla większości inwestorów gorszy byłby wzrost kursu – spowodowałby on pogłębienie się strat ok. 35% populacji z tytułu otwartych pozycji krótkich i poprawiałby sytuację ok. 24% przegrywających długich pozycji. W diagramie złożonych zleceń ruch byłby o wiele mniejszy niż w przypadku spadku kursu.

Ponieważ historycznie dzieje się tak, że większość przegrywa, to logicznym (w sensie myślenia indukcyjnego, nie zaś myślenia logicznego) ruchem byłoby wykonanie operacji (np. otwarcie pozycji), wygrywającej przy równoczesnej porażce większości. Na zasadzie przekory należałoby więc otworzyć więc pozycję długą. Z psychologicznego punktu widzenia, to niezwykle trudna decyzja, o wiele łatwiejsza dla odpowiednio zaprogramowanego infobota.

W rozpatrywanym tu modelu zakłada się także występowanie wzorców uzasadniających postawienie następującej tezy. Jeżeli wielokrotnie weryfikowana konfiguracje zmiennych wejściowych  $X_2 \div X_9$  pozwala na zbudowanie modelu regresyjnego (niekoniecznie liniowego) takiego, że wg ustalonego kryterium akceptowalne mogą być błędy wynikające z zastosowania tego modelu na odcinku predykcyjnym w fazie uczenia, to oznaczać to może, że znaleziono wzorzec (model regresyjny), na podstawie którego można zbudować system transakcyjny (oparty na koncepcji prognozy).

Jak zwykle, ważne jest rozróżnienie modelu predykcyjnego od modelu systemu transakcyjnego (inwestycyjnego). Pierwszy jest wyrazem zastosowania wyrafinowanych metod analizy technicznej (badania skutków zmian rynku) w celu prognostycznym, a drugi jest zastosowaniem (wykorzystaniem) wyników pierwszego modelu w warunkach proponowanych przez konkretną platformę inwestycyjną.

Podobnie jak w pracach [36, 50, 73, 79] do rozwiązania zadań prognostycznych użyta będzie rozpatrywana tu metoda GMDH. Zastosowany będzie tu algorytm oparty na nieliniowej regresji (rozdział 2), w której zmienną obserwowaną będzie zwrot po zakończeniu kolejnej świecy a zmiennymi wejściowymi kurs waluty i osiem opisanych całek  $G_i$ ,  $i = 1, 2, \dots, 8$ . Przez zwrot rozumie się tu różnice (w pipsach) pomiędzy wartościami kursu w chwili bieżącej i poprzedniej, oddalonymi o okres jednej świecy.

Zanim przystąpiono do zbudowania modelu regresyjnego rozszerzona została macierz danych wejściowych poprzez konkatencję horyzontalną nowych kolumn o długości  $N$ , z których każda jest iloczynem wartości w kolumnach z danymi pierwotnymi. Tych kolumn z danymi pierwotnymi jest 8, gdyż uwzględniono tylko kolumny reprezentujące całki  $G_1$  do  $G_8$ . W związku z tym kolejne kolumny o indeksie większym o liczbę  $k$  od dotychczasowej liczby kolumn  $N$  kolejno zawierały:

$$X_{N+k} = X_i * X_j \quad (67)$$

dla  $k = 1, 2, \dots, i = 2, 3, \dots, 9, j = i+1, i+2, \dots, 9$ , gdzie operator „\*” oznacza mnożenie tablicowe wektorów wyraz po wyrazie, tzn. wynikiem mnożenia jest kolejny wektor o takim samym wymiarze, którego wyrazy są iloczynami odpowiednich wyrazów dwóch kolumn wejściowych.

Liczba dodatkowych kolumn wg (67) z uwzględnieniem kwadratów wartości kolumn pierwotnych wyrazów  $N_d = N(N-1)/2 + N$ . Razem po konkatencji macierz danych ma wymiar  $M \times (N + N_d)$ .

Dla zbudowania modeli prognostycznych na podstawie tak przetworzonych danych pierwsza kolumnę ( $X_1$ ) traktuje się jako zmienną obserwowaną  $Y$  (odnosi się ona do zwrotu po kolejnym kroku), a wszystkie pozostałe kolumny jako zmienne wejściowe (objaśniające). W ten sposób ostateczny model opisujący np.  $K$  wierszy zmiennej wyjściowej miał postać formalnie liniową:

$$\hat{y}_{i+1} = X_i * A_{i+1} \quad (68)$$

$$[1 \times 1] \quad [1 \times (L+1)] \quad [L+1] \times 1$$

dla  $L = 3, 4, 5, \dots, i = 1, 2, \dots$

W modelu tym zmienna  $\hat{y}_{i+1}$  oznacza predykcyjną modelowaną wartość zmiennej obserwowanej, która w chwili bieżącej wynosi  $y_i$ .  $X_i$  jest wektorem  $(L+1)$  zmiennych wejściowych wybranym wg opisanego dalej algorytmu spośród  $N_d$  zmiennych rozszerzonej o wyrazy kwadratowe macierzy  $X$  zaobserwowanym w chwili bieżącej  $i$ . Wektor  $A_{i+1}$  jest zestawem  $(L+1)$  współczynników równania liniowego.

$$A_{i+1} = [a^0_{i+1}, a^1_{i+1}, a^2_{i+1}, \dots, A^L_{i+1}] \quad (69)$$

obliczonym w przeszłości na podstawie dostępnych danych historycznych dla wierszy danych rozpoczynających się w chwili  $i-K-1$ , a kończących się w chwili  $i-1$ .

Ponieważ zakłada się, że jest dana informacja do chwili  $i$ -tej to wykorzystując dane wejściowe do chwili  $i-1$  a zmienną obserwowaną do chwili  $i$ -tej uzyskuje się możliwość wykonania prognozy o jeden krok wprzód. Liczba  $K$  obserwowanych wierszy jest (na razie) przyjmowana arbitralnie i może być także przedmiotem dalszej optymalizacji. Wektor  $A_{i+1}$  obliczany jest z równania regresji wykonanej na  $K$  wierszach i  $L$  kolumnach:

$$A'_{i+1} = X_i \setminus Y_i \quad (70)$$

$[1 \times (L+1)] \quad [(L+1) \times K] \quad [(K \times 1)]$

które można także zapisać [104, 110] jako:

$$A'_{i+1} = \text{pinv}(X_{i-1}) * Y'_i \quad (71)$$

$[(L+1) \times 1] \quad [(L+1) \times K] \quad [K \times 1]$

gdzie  $Y'$  oznacza wektor transponowany a  $\text{pinv}$  jest operatorem pseudo-inwersji Moore'a-Penrose'a (inwersji uogólnionej) [135]. Dla tak obliczonego wektora współczynników  $A_{i+1}$  można wyznaczyć  $K$  wartości zmiennej wyjściowej:

$$\hat{Y}'_i = X_{i-1} * A_{i+1} \quad (72)$$

$[K \times 1] \quad [K \times (L+1)] \quad [(L+1) \times 1]$

gdzie  $X_{i-1}$  jest macierzą danych obserwowanych od wiersza leżącego  $K$  wierszy wstecz od chwili bieżącej do wiersza (chwili) o indeksie  $i-1$ . Indeks  $(i+1)$  przy wektorze  $A_{i+1}$  oznacza, że współczynniki są obliczone wektora  $Y$  przesuniętego o jeden krok wprzód w stosunku do danej  $X_{i-1}$ , po to by umożliwić predykcje (68).

Niewiadomą w tych rozważaniach jest liczba  $L$  współczynników równania jak i struktura modelu (czy do równania wchodzi tylko wyrazy liniowe, kwadratowe czy model jest kombinacją tych wyrazów).

Problem ten rozwiązuje algorytm oparty na idei Group Method of Data Handling dokonujący dla każdego i-tego kroku wielokrotnego wyboru L wyrazów dla  $L=3,4,5,\dots,N_d$ . Tworząc na K wierszach danych macierz danych do (i-1) wiersza:

$$X_{i-1} \left[ \mathbf{1} \ X_{i-1}^{j_1} \ X_{i-1}^{j_2} \ \dots \dots \ X_{i-1}^{j_L} \right] \quad (73)$$

gdzie:

$j_1, j_2, \dots, j_L \in \{1, 2, \dots, N_d\}$

$\mathbf{1}$  – oznacza wektor jedynek o długości K;

$X_{i-1}^{j_1}$  – oznacza wektor danych  $j_1$ -tej kolumny, takiej, że:

$j_1 \in \{1, 2, \dots, N_d\}$  kończący się na (i-1) wierszu, mający długość K.

Dla każdego obliczanego wg (72) z użyciem (73) modelu rozpatrywane jest kryterium dokładności odtworzenia przebiegu zmiennej obserwowanej jako:

$$K_{1L}^J = \frac{1}{K} \sum \|\hat{Y}_i - Y_i\|^2 \rightarrow \min \quad (74)$$

Indeksy w kryterium  $K_{1L}^J$  oznaczają:

L – jest to kryterium obliczone dla modelu zawierającego (L+1) wyrazów (L oraz wektor jedynek przeznaczony w modelu do obliczenia wyrazu wolnego równania regresji);

J =  $\{j_1, j_2, \dots, j_L\}$  – określa, które spośród kolumn macierzy danych  $X_{i-1}$  zostały wykorzystane do obliczenia modelu  $\hat{Y}_i$ .

W ten sposób obliczając  $L_{1L}$  dla wszystkich kombinacji  $C_{L_{Md}}$  dla każdego  $L = 3, 4, 5, \dots$  znajdowane jest  $K_{1L, \min}$  z odpowiadającym mu wektorem L indeksów kolumn  $\{j_1, j_2, \dots, j_L\}$ . Opisany tu algorytm poszukiwania najlepszego modelu można przedstawić następująco:

for  $i=1: N_d-3$

for  $j = i+1 : N_d-2$

for  $k = j+1 : N_d-1$

- 1)  $X_{i-1}$  wg (10) dla  $L=3$
- 2) Obliczenie  $A_{i+1}$  wg (7),  $\hat{Y}_i$  wg (72) i kryterium wg (74)
- 3) Zapamiętywanie kryterium, jeżeli jest najlepsze
- 4) Zapamiętywanie modelu, jeżeli kryterium jest najlepsze

```

                                for l = k+1 : Nd
                                    Xi-1 wg (10) dla L=4
                                    Podobnie jak wyżej zapamiętywanie naj-
                                    lepszego modelu
                                end / l
                            end / k
                        end / j
                    end / i

```

Tak znaleziony model użyty jest do wykonania predykcji (68). Jest to predykcja wykonana jeden krok wprzód przy założeniu, że uznany za najlepszy wg kryterium (74) model  $\hat{Y}_i$ , będzie skuteczny jeszcze jeden krok i da satysfakcjonujące przybliżenie  $\hat{y}_{i+1}$  wg (68). Gdy upłynie czas trwania jednego kroku, będzie można sprawdzić, czy wartość  $y_{i+1}$  (której w danej chwili jeszcze nie ma) będzie bliska modelowi  $\hat{y}_{i+1}$ . Wielokrotne powtarzanie tej procedury i ocena jakości modeli  $\hat{y}_{i+1}$  będzie przedmiotem badań. Należy przy tym zauważyć, że w każdym kroku model  $\hat{Y}_i$  budowany jest z użyciem innych zmiennych i z zastosowaniem różnej struktury. Posiada także, co oczywiste, zupełnie inny zestaw współczynników  $A_{i+1}$ .

W identyczny sposób można zbudować model (modele) umożliwiające predykcję na dwa kroki wprzód (także więcej). Wówczas:

$$\hat{y}_{i+2} = A_{i+2} \cdot X_i \quad (75)$$

gdzie:

$$A_{i+2} = X_{i-2} \setminus Y_i \quad (76)$$

$$\hat{Y}_i = A_{i+2} \cdot X_i' \quad (77)$$

gdzie  $X_{i-2} = [1 \ X_{i-2}^{j1} \ X_{i-2}^{j2} \ \dots \ X_{i-2}^{jL}]$ .

W modelu (75-76) dane macierzy rozszerzonej przez konkatenację poziomą są wykorzystywane jedynie do wiersza (i-2) od chwili bieżącej, a wektor zmiennej obserwowanej, tak jak poprzednio, do chwili bieżącej. Stąd inne współczynniki  $A_{i+2}$  i inne struktury i wielkości modeli (76). Model umożliwiający predykcję dwa kroki wprzód ma jeszcze dodatkowa interesująca właściwość. Otóż stosując go konsekwentnie krok po kroku można doprowadzić do kolejnego potwierdzenia prognozy. Jeżeli bowiem w jakiejś chwili  $t_i$  wykonana jest predykcja na dwa kroki wprzód, czyli dla chwili  $t_{i+1}$  i  $t_{i+2}$  to również w tej samej chwili znana jest predykcja z poprzedniego kroku dotycząca chwili  $t_i$  oraz  $t_{i+1}$ . Tak więc w chwili  $t_i$  istnieją dwie predykcje dla chwili  $t_{i+1}$ . Niech będą one

oznaczone odpowiednio  $p_{i+1}^i$  i  $p_{i+1}^{i-1}$ , gdzie indeks dolny oznacza, dla której chwili predykcja jest wykonana a indeks górny, w której chwili została wykonana. Z różnicy indeksów widać, że pierwsza predykcja  $p_{i+1}^i$  to predykcja jeden krok wprzód a predykcja  $p_{i+1}^{i-1}$  wykonana jest dwa kroki wprzód.

Przez predykcję niech będzie rozumiane odchylenie zmiennej obserwowanej od wartości w kroku poprzednim np.:

$$\begin{aligned} p_{i+1}^i &= \hat{y}_{i+1} - y_i \\ p_{i+1}^{i-1} &= \hat{y}_{i+1} - \hat{y}_i \end{aligned} \quad (78)$$

lecz także:

$$p_{i+1}^{i-1} = \hat{y}_i - y_{i-1} \quad (79)$$

Ponieważ w systemach transakcyjnych najważniejsza jest trafność kierunku zmian zmiennej obserwowanej, a nie dokładne określenie jej wartości, to w dalszych rozważaniach istotne będzie czy znak:

$$\text{sign}(p_{i+1}^i) = \text{sign}(p_{i+1}^{i-1}) \quad (80)$$

Jeżeli warunek (80) zachodzi to interesująca staje się wartość znaku. Jeżeli  $\text{sign}(p_{i+1}^i) = 1$ , oznacza to, że prognozowany jest wzrost rozpatrywanego kursu (tu ceny pary walutowej EUR/USD) i należy otworzyć pozycję długą, a gdy  $\text{sign}(p_{i+1}^i) = -1$ , należy otworzyć pozycję krótką. Jeżeli warunek (80) nie był spełniony nie wykonano żadnej operacji.

Formalnie uwarunkowania te można przedstawić wyprowadzając pewną zmienną decyzyjną  $d_i$  określającą, co powinien system transakcyjny wykonać.

$$d_i = \begin{cases} 1 & \text{gdy spełniony jest warunek (16)} \\ & \text{oraz gdy } \text{sign}(p_{i+1}^i) = 1 \\ 0 & \text{gdy nie spełniony warunek (16)} \\ -1 & \text{gdy spełniony jest warunek (16)} \\ & \text{oraz gdy } \text{sign}(p_{i+1}^i) = -1 \end{cases} \quad (81)$$

dla  $i = 2, 3, \dots$

W definicji (81)  $d_i$  oznacza decyzję otwarcia pozycji, gdy  $d_i \neq 0$  lub bezczynność, gdy  $d_i = 0$ . Każda otwarta pozycja była zamykana po jednym kroku. Dla tak zdefiniowanej strategii z użyciem modelu (75-76)

i danych jak w tab. 13 przeprowadzono badania. Badania przeprowadzono z wykorzystaniem danych zgromadzonych w wyniku dwumiesięcznego automatycznego monitoringu platformy oanda.com [138].

Dla 800 kolejnych kroków uzyskano modelowaną zgodność predykcji w pierwszym i drugim kroku w 688 przypadkach spośród których 350 (52,4%) otwarte pozycje dały sukces a 338 porażkę.

Stosowana L zmienna od 3 do 4 oraz długość wektorów danych w równaniach regresji równa  $K=45$ . Dla zabezpieczenia strat zastosowano mechanizm Stop Loss (SL) ustawiony na poziomie 50 pipsów, co dla świec jednogodzinnych jest poziomem trudnym do osiągnięcia, ale osiąganym (w rozpatrywanym horyzoncie 800 kroków SL było uruchamiane 16 razy). Po każdym  $k$  kroku mierzono zwrot (w pipsach):

$$z_k = (y_k - y_{k-1}) \cdot \text{sign}(p_k^{k-1}) \quad (82)$$

Zapis (82) oznacza, że wartość dodatnią (sukces) osiągnano wówczas, gdy przy otwartej pozycji długiej ( $\text{sign}(p_k^{k-1}=1)$ ) nastąpił wzrost kursu lub przy otwartej pozycji krótkiej ( $\text{sign}(p_k^{k-1} = -1)$ ) nastąpił spadek. Wynik skumulowany przedstawiono na rys. 90.

Przedstawiony na rys 90. wykres skumulowanej wartości  $z_k$ :

$$z_s(s) = \sum_{i=1}^k z_i \quad (83)$$

osiąga wartość wskaźnika Sharpe'a równą 0.0882, a wskaźnika Sortino równą 0.5485. Wskaźnik Sortino to wskaźnik Sharpe'a odniesiony do wariancji jedynie tych  $z_k$ , które przyniosły straty.

Wskaźnik Sharpe'a obliczany jest jako:

$$SR = z(k) / \text{std}(z) / k \quad (84)$$

a wskaźnik Sortino jako:

$$SoR = z(k) / \text{std}(z_u) / k \quad (85)$$

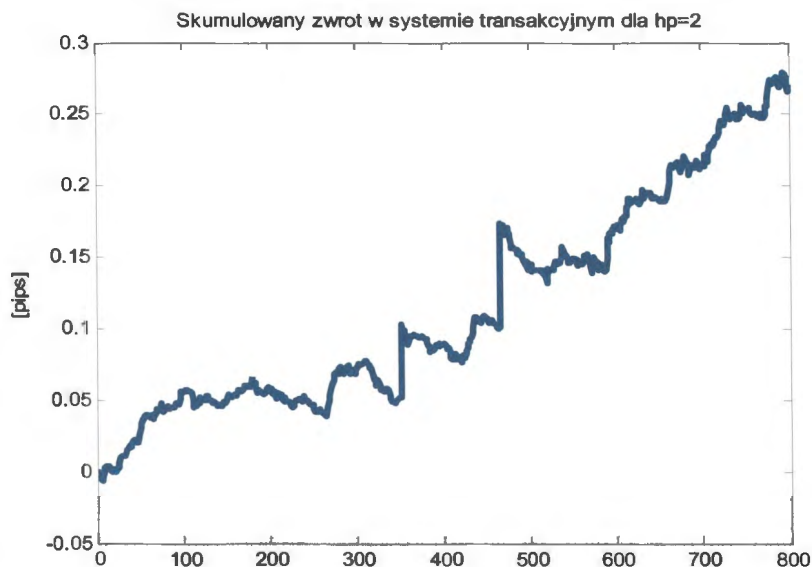
gdzie:

$z(k)$  – skumulowany zwrot po  $k$  kroku;

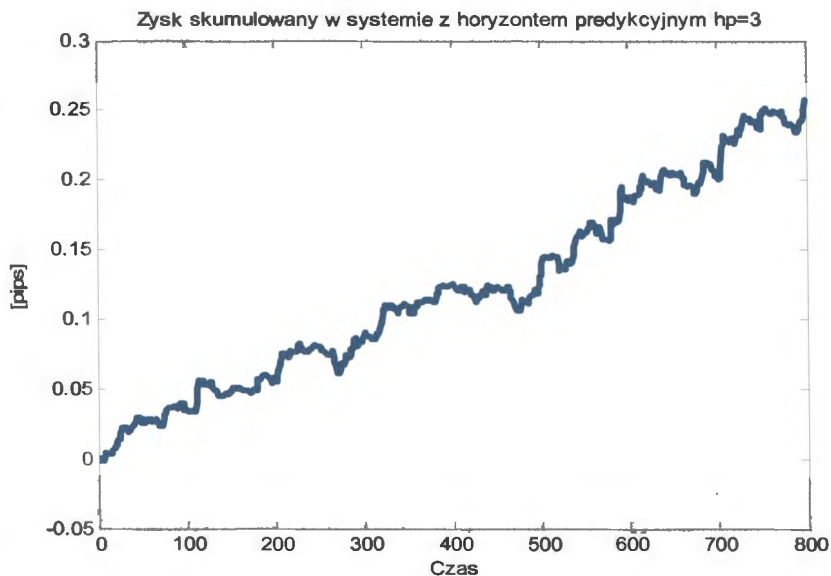
$\text{std}(z)$  – odchylenie standardowe w szeregu czasowym z zwrotów;

$\text{std}(z_u)$  – odchylenie standardowe w szeregu czasowym tylko ujemnych zwrotów.

To uproszczone formuły, nieuwzględniające stopy zwrotu wolnej od ryzyka, ale tak zdefiniowane używane są tylko w obrębie pracy.



Rys. 90. Skumulowany zwrot po 800 krokach systemu transakcyjnego z horyzontem predycyjnym  $hp=2$  kroki wprzód. Wskaźnik Sharpe'a równy 0.0882.

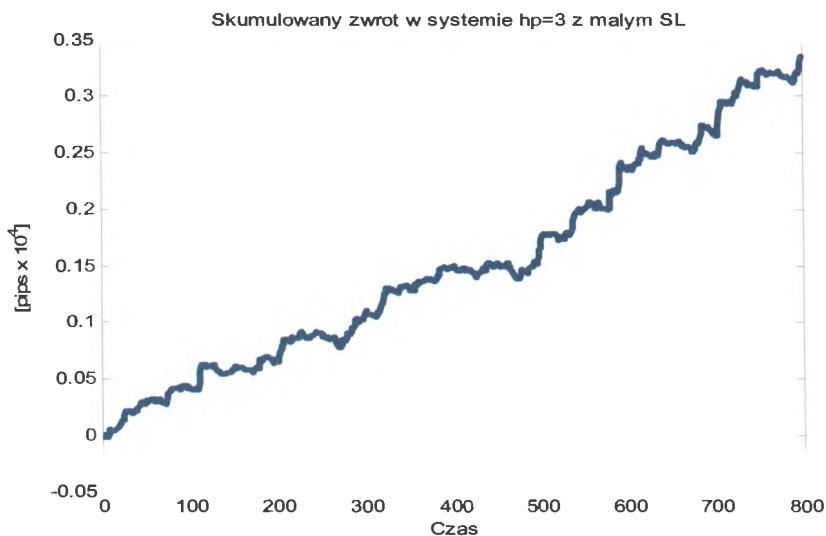


Rys. 91. Skumulowany zwrot po 800 krokach systemu transakcyjnego z horyzontem predycyjnym  $hp=3$  kroki wprzód. Wskaźnik Sharpe'a równy 0.1291.



Na rys. 91 przedstawiono wyniki narastania kapitału dla horyzontu predycyjnego wynoszącego 3 kroki wprzód. Podobnie jak poprzednio decyzja inwestycyjna była każdorazowo podejmowana tylko po potwierdzeniu predykcji. Na 800 symulowanych przypadków uzyskano potwierdzenie 575 razy z tego 310 razy (53.9%) udane. Wskaźnik Sharpe'a dla tej symulacji wynosił 0.1291, a więc był znacznie lepszy niż w poprzednio opisanym przypadku.

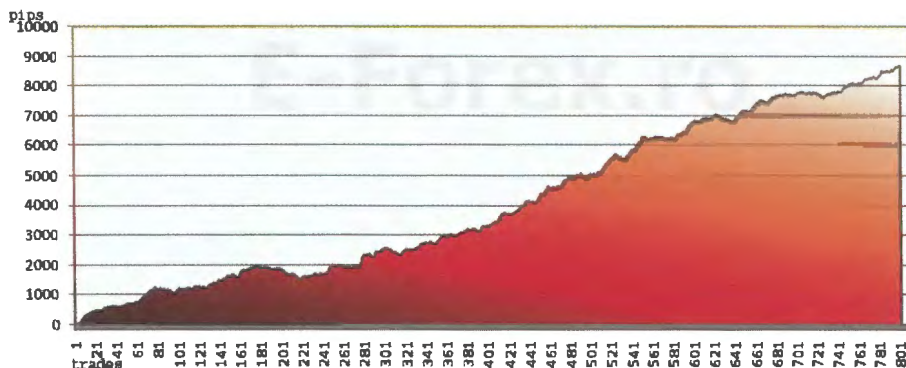
Obie symulacje przeprowadzono przy celowo dużej wartości SL (*Stop Loss*), zwiększającej ryzyko strategii. Nie rozpatrywano problemu optymalizacji wartości tego zabezpieczenia, ale np. po dwukrotnym zmniejszeniu SL do wartości 25 pipsów (w świecach godzinnych) uzyskano wyniki przedstawione na rys. 92. Wizualnie krzywa narastania kapitału przedstawia się tu jeszcze korzystniej, a wskaźnik Sharpe'a osiąga już poziom 0.2056.



Rys. 92. Skumulowany zwrot po 800 krokach systemu transakcyjnego z horyzontem predycyjnym  $hp=3$  kroki wprzód i zmniejszonym SL. Wskaźnik Sharpe'a równy 0.2056.

Należy zwrócić uwagę na dość dobre wyniki wszystkich symulacji pomimo niewielkiej przewagi prognoz poprawnych nad błędnymi w aspekcie kierunku zmiany wartości kursu. Doszukiwać się w tym sukcesie należy pozytywnego wpływu istoty metody – prognozuje ona nie tyle kierunek zmian, ile rzeczywiste wartości kursu po jednym lub dwu

krokach. Stąd zapewne błędy (co do kierunku) mają zdecydowanie mniejszy wpływ negatywny na wynik końcowy, niż pozytywny wpływ poprawnych predykcji. Te liczne przegrane są po prostu średnio mniejsze niż liczne wygrane. Np. dla badań przedstawionych na rys. 91 średnia przegrana wynosiła 6 pipsów a średnia wygrana aż 21 pipsów. Inaczej objaśniając dobre rezultaty można byłoby scharakteryzować je następująco – algorytm poprawnie wyłapuje prawie wszystkie istotne (silne) odchylenia od kursu bieżącego, przy odchyleniach słabszych często się myli.



Rys. 93. Skumulowany zwrot po 801 krokach systemu transakcyjnego realizowanego wg zasad klasycznej analizy technicznej przez Liviu Flesara [137]. Wskaźnik Sharpe'a równy 0.2312.

Podsumowując wyniki przeprowadzonych symulacji należy uznać je za bardzo obiecujące. Uzyskane są na największym pod względem obrotów światowym rynku, na najbardziej płynnej parze walutowej. Porównywalne są z najlepszymi publikowanymi w Internecie – np. z wynikami uzyskiwanymi przez Liviu Flesara [137], od lat skrupulatnie ujawnianymi, uzyskiwanymi w wyniku tradycyjnej percepcyjnej analizy formacji technicznych. Rezultaty publikowane przez tego rumuńskiego analityka przedstawiono na rys. 93. Uzyskany przez niego wskaźnik Sharpe'a wynoszący 0.2312 porównywalne jest z prezentowanym w niniejszym rozdziale (przy niemal identycznej liczbie próbkowań).



## 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 tyśiące alternatyw wobec bezpiecznych obligacji?

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

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

Rozwiązaniem dla wozu jest zmiana koni. Na bardzo szybkie.



# Bibliografia

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



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

## Bibliografia

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

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

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