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[1] 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[2] 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[3] i foreksowych[4] transakcjach to zdarzenia częstotliwe, niejednokrotnie z odcieniem chorobliwego hazardu i najczęściej bez happy-endu.

---

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

[2] Freddie Mercury – słynny brytyjski piosenkarz ub. wieku, którego jednym z największych przebojów był utwór „The show must go on".

[3] W pracy rozpatrywane będą predykcje walorów (akcji) giełdowych i kursów walut.

[4] Forex – to międzynarodowy rozproszony rynek walutowy udostępniany właśnie poprzez tysiące internetowych platform brokerskich.

Czego więc syntezą jest ostatecznie sukces na rynku[5] 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 sytemu 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 zawężać 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 odżegnując 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 -

[5] 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 teoria 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 kontynuacje 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 ideą przyświecającą 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[6] 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 wspominanych 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

[6] 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 [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 sie 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 selekcjonerów 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 konsensusem 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[7].

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

[7] 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 rezygnować z możliwości rozpoznania tej złożoności. Mało tego, bardzo często złożony model jest niespodziewanie precyzyjny i nie przeuczony[8].

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)[9] i możliwości aktualnych komputerów. Dodatkową obliczeniową komplikacją podejścia proponowanego przez Aleksieja Ivakhnenkę i naśladowców – Hema Madalę [73], Johanna Muellera [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] x TP[11] x Tdop (poziom akceptowalnej straty/poziom satysfakcjonującego zysku/ustalony maksymalny okres

---

[8] To twierdzenie o krok od teorii chaosu, której Ivaknenko nie przywołuje.

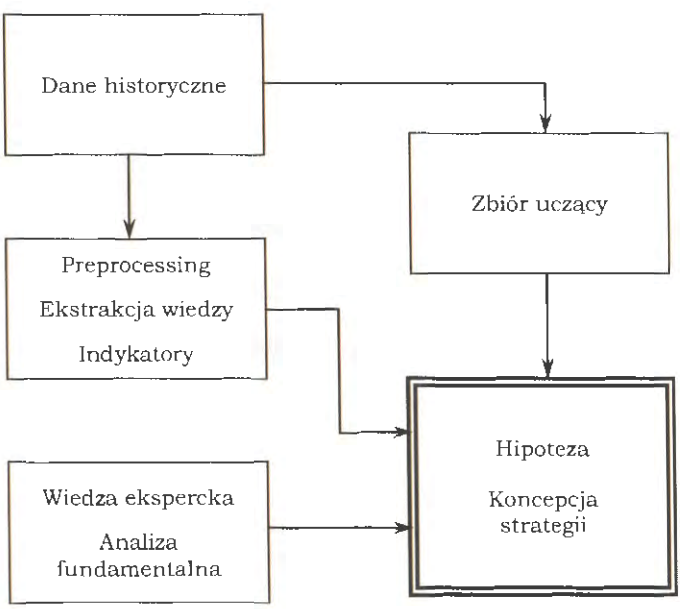
[9] Group Method of Data Handling – angielskie tłumaczenie metody opracowanej przez Ivakhnenkę w oryginale zwane *Metod Gruppovogo Uceta Argumetov* (ros.). W języku polskim nie ma dotąd jednoznacznego przyjętego powszechnie tłumaczenia. Właściwe wydaje się być – Metoda Grupowania Atrybutów czy Metoda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

[10] 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.

[11] 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 utrata 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.

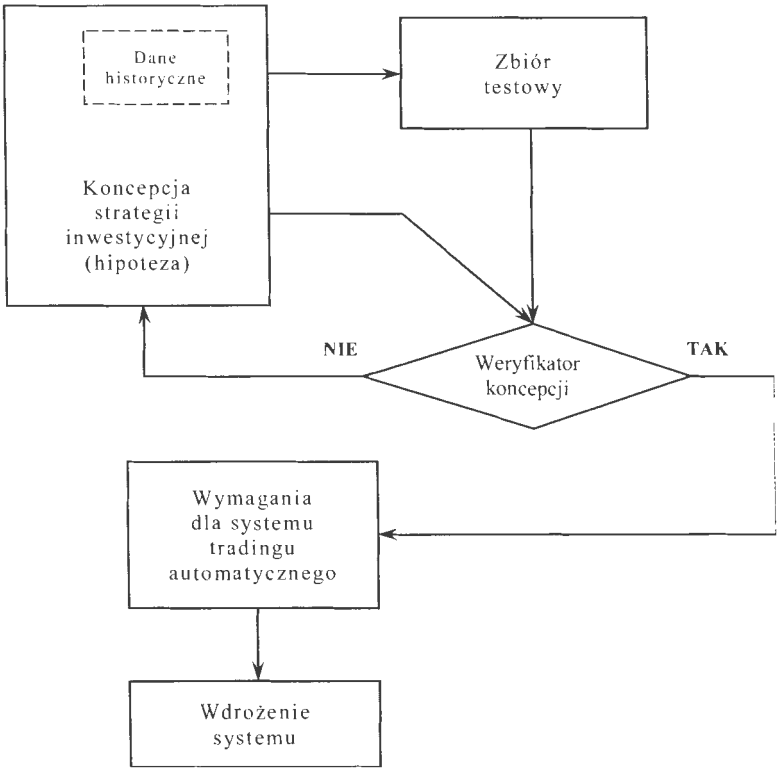


Rys. 1. Schemat blokowy procesu powstawania hipotez badawczych dotyczących koncepcji strategii inwestycyjnych.

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

algorytmami tej przestrzeni, stosowaniu wyrafinowanej filtracji, uznanych i nowych indykatorów 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[12] 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

[12] 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 implementacje 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.

3

Przestrzeń zmiennych dla systemu transakcyjnego wykorzystującego algorytmy predykcyjne GMDH

By istniało zwierciadło świata, świat musi mieć jakiś kształt
Umberto Eco (1932 -)

We wstępie do rozdziału 2 wspomniano o konieczności odrębnego opisania zmiennych wejściowych zbudowanych na rozpatrywanych tu świecach EUR/USD. Zmienne te są wynikiem przekształcenia danych surowych – tu, świecy OHLC (*Open, High, Low, Close*) wybranej pary. Były one tam oznaczane jako macierz $\mathbf{X}_j$, która można rozumieć jako M wektorów kolumnowych X_j dla j=1, 2, ... M . W każdej kolumnie tej macierzy znajduje się N wartości zmiennej obliczonej wg kolejno ustalanych lub wybieranych reguł. Sposób i wybór zmiennych jest wyłączną domeną człowieka, tego procesu nie da się zautomatyzować i nic nie jest tu oczywiste. Przy dużym stopniu złożoności formuły obliczeniowej lub sposobu pozyskiwania danych do obliczenia bieżącej wartości zmiennej X_j nazywana jest ona indykatorem (wskaźnikiem). Na rynkach finansowych stosuje się dziesiątki znanych indykatorów do gene-

rowania sygnałów otwarcia i zamknięcia pozycji. Najważniejsze to np. MACD, RSI, ATR, wstęgi Bollingera, średnie kroczące, punkty Pivota, poziomy Fibonacciego, indykator Williamsa, Stochastic i wiele innych. Ich znaczenie dla efektywności systemu transakcyjnego jest często dyskutowane, przeceniane, niedoceniane, a w istocie jest najczęściej wynikiem kreatywności inwestora stosującego te indykatory w swoim systemie z pewną dozą krytycyzmu i ostrożności lub z ich pominięciem. Tak czy inaczej, do dziś nie wynaleziono na szczęście uniwersalnego indykatora, który wskazywałby jednoznacznie, kiedy należy otworzyć pozycje a kiedy zamknąć, tak jak nie wynaleziono *perpetuum mobile* w formie robota zarabiającego pieniądze na rynkach finansowych bez żadnych ograniczeń, uwarunkowań i ryzyka (gdyby wyedukowany czytelnik pomyślał o arbitrażu, to ten też ma te ograniczenia). Istnieją też tysiące mniej znanych, niepublikowanych powszechnie indykatorów autorskich tworzonych przez rozmaitych inwestorów (graczy, spekulantów, traderów[24]) na swoje potrzeby [140]. Niektóre z nich są w przypadku sukcesu opisywane i publikowane, inne zaś nie. Pozostają tajemnicą ich wynalazców [66].

W rozdziale przedstawione będą indykatory, które, użyte równocześnie, utworzyły przestrzeń danych o dość wysokiej efektywności transakcyjnej.

Z matematycznego punktu widzenia indykator jest pewną funkcją charakterystyczną zbioru wartości zmiennej X_j. Niech funkcja ta będzie zdefiniowana jako:

$$f(x_j) = \begin{cases} 1 & \text{gdy x należy do } B_j^l \\ 0 & \text{gdy x nie należy ani do } B_j^l \text{ ani do } B_i^s \\ -1 & \text{gdy x należy do } B_j^s \end{cases} \qquad (36)$$

gdzie zbiór B_j^l jest zawarty w X_j i oznacza zbiór takich wartości x_j, dla których generowany jest sygnał otwarcia pozycji długiej w systemie transakcyjnym, a B_j^s zawarty w X_j oznacza zbiór takich wartości x_j, dla których generowany jest sygnał otwarcia pozycji krótkiej w tym samym systemie transakcyjnym.

[24] Określenie osoby czy podmiotu wykonującego akcje na rynku ma tu znaczenie drugorzędne i najczęściej odnosi się nie do obiektywnych i technologicznych aspektów metodyki działania, lecz do subiektywnie ocenianych motywacji i celów do osiągnięcia. Określenia te ujawniają także często zawoalowaną relację osoby stosującej to określenie do osoby ocenianej.

Funkcja f(x) przybiera najczęściej postać funkcji logicznej w postaci zdania:

$$\text{czy } f(x_j) > g_j,$$

gdzie g_j jest charakterystyczną wartością graniczną j-tej zmiennej wejściowej.

Należy przy tym pamiętać, że X_j jest przetransformowaną funkcją bazową dla danego indykatora – najczęściej jest wynikiem przetworzenia danych z zakresu świecy OHLC, na bieżących i wybranych z historii wartościach tych świec.

Dla przykładu – jeżeli indykator miałby być zbudowany na różnicy średnich kroczących MA_5 i MA_{10} co oznacza odpowiednio średnią z pięciu i dziesięciu ostatnich świec (np. wartości Close) a różnica (funkcja bazowa indykatora) byłaby określona jako RMA= MA_5 - MA_{10}, to funkcja charakterystyczna mogłaby mieć np. następującą postać:

$$f(x_j) = \begin{cases} 1 & \text{gdy x należy do } B_j^l, \text{czyli np. gdy RMA} > 0.0010 \text{ (10 } \textit{pipsów}) \\ 0 & \text{gdy x nie należy ani do } B_j^l \text{ ani do } B_j^s \\ -1 & \text{gdy x należy do } B_j^s, \text{czyli np. gdy RMA} < -0.0015 \text{ (15 pipsów)} \end{cases}$$

co można byłoby wyjaśnić językiem systemu transakcyjnego następująco – jeżeli szybsza średnia krocząca będzie wyższa o 10 pipsów od wolniejszej średniej kroczącej, to otwórz pozycję długą (to oczywiście zupełnie wyimaginowana przykładowa decyzja typu – zlecenie *Buy Stop*), a jeżeli będzie mniejsza niż 15 pipsów, to otwórz pozycję krótką. Decyzje takie muszą wynikać ze sprawdzonej na danych historycznych strategii tradingowej i mogą być zupełnie inne i przy całkowicie innych, niż użyte w powyższym przykładzie progach. Najistotniejsze w tym przykładzie jest zrozumienie istoty przetworzonej wstępnie funkcji bazowej dla indykatora. Zmienna X_j to w danym przykładzie zmienna RMA (różnica średnich kroczących) transformowana do wartości z dziedziny trójwartościowej {-1 , 0 , 1} w zależności od jej relacji względem ustalonych w strategii wartości progowych.

W rozpatrywanych w niniejszej pracy strategiach transakcyjnych stosowana była zasada koniunkcji, czyli równoczesności występowania (wzajemnego potwierdzania) wielu warunków związanych z poszczególnymi indykatorami.

Poniżej przedstawione zostaną wyniki badań efektywności granic g(x) dla różnych par atrybutów z przestrzeni zmiennych X_j dla j=1,2...M.

Dane stosowane w badaniach

W przedstawionych w niniejszym rozdziale badaniach użyte zostały świece EUR/USD 15 min w liczbie ok. 31000 z okresu od końca 2005 roku do ok. początku kwietnia 2007 roku pozyskane z platformy brokerskiej cmsfx.com. Rok tradingowy w tym zbiorze danych liczy ok. 250 dni (okresy weekendowe pozostawały niewypełnione).

Nazwa pliku danych: cms15m0604

Charakterystyka strategii

W największym uproszczeniu strategia sprowadza się do otwierania pozycji (długich i krótkich) na rynku spot EUR/USD w odpowiednich warunkach, a także zamykania tych otwartych pozycji w innych, także odpowiednich warunkach. Długie lub krótkie pozycje otwierane będą w chwili równoczesnego spełnienia warunków – osiągania przez indykatory wartości granicznych. Indykatory te i warunki z nich wynikające będą w przedstawianych tu badania rozpatrywane parami w celu ułatwienia wizualizacji oddziaływania warunków na kryteria.

Zazwyczaj uważa się, że dla bezpieczeństwa inwestora ważniejsza jest umiejętność właściwego zamknięcia pozycji niż jej otwarcia. Jeżeli chodzi o sposób zamykania otwartych pozycji to będą one zamykane alternatywnie po osiągnięciu jednej z trzech następujących barier: TP (*Take Profit*), SL (*Stop Loss*), Tdop (maksymalny czas otwarcia pozycji – czas wygaśnięcia). Jest to przestrzeń charakterystyczna dla wielu zautomatyzowanych lub półzautomatyzowanych platform brokerskich. Gwarantuje ona zamkniecie każdej otwartej pozycji (co jest logiczną koniecznością zapewnienia inwestorowi możliwości zarządzania ryzykiem), nie gwarantując oczywiście sukcesu, który zależy od trafności przewidywania wartości stosowanych barier zamknięcia.

Reguły zamknięcia otwartych pozycji rozpatrywane są w kolejności „pesymistycznej" tzn. najpierw w każdym przedziale próbkowania (przy rozpatrywaniu kolejnej świecy) rozpatruje się ewentualna porażkę a dopiero potem ewentualny sukces. Takie „ostrożnościowe" podejście powoduje, że wynik symulacji jest oszacowaniem dolnym wyniku rzeczywistego (wynik rzeczywisty nie powinien być gorszy niż wynik symulacji). Dla pewnej konkretnej pozycji, np. długiej oznacza to, że rozpatruje się najpierw ewentualność jej zamknięcia przez barierę SL położoną poniżej ostatniego zamknięcia a dopiero potem ewentualność zamknięcia przez barierę TP. Gdyby w obrębie jednej kolejnej świecy wystąpiły oba przypadki dotknięcia tych barier to zawsze nastąpi najpierw zamkniecie wskutek porażki.

Przestrzeń sygnałów otwarcia, spośród której rozpatrywane będą kolejno pary zmiennych, które tworzą zmienne z następującego zbioru atrybutów:

- wektor rzędów kwantyli, zbudowany na rozkładzie statystycznym wartości zamknięć ostatnich *Sd* świec. Będzie on omówiony szczegółowo w dalszym podrozdziale.
- liczba świec *Sd*, wstecz od chwili bieżącej, którą uwzględnia się dla obliczenia momentów statystycznych i innych wskaźników (najczęściej w badaniach rozważano *Sd*=500);
- zmodyfikowany indykator Lane'a zwany Stochastic Oscillator;
- skośność, jako moment statystyczny trzeciego rzędu zbudowany na zbiorze zamknięć ostatnich *Sd* świec;
- kurtoza - jako moment statystyczny czwartego rzędu zbudowany na zbiorze zamknięć ostatnich *Sd* świec.

Przestrzeń sygnałów zamknięcia jest utworzona przez trzy wspomniane bariery zapewniające bezwarunkowe zamkniecie każdej otwartej pozycji poprzez osiągnięcie jednej z tych barier:

- bariery polegającej na wystawieniu zlecenia ograniczającego zysk po osiągnięciu satysfakcjonującego inwestora poziomu *Take Profit* (TP);
- bariery polegającej na wystawieniu zlecenia ograniczającego stratę w przypadku zmiany wartości instrumentu finansowego w kierunku przeciwnym do oczekiwanego przez inwestora – zlecenia zwanego *Stop Loss* (SL);
- zlecenia automatycznego zamknięcia każdej otwartej pozycji niezależnie od jej aktualnej wartości po upływie zadanego czasu ważności (wygaśnięcia) T_{dop}, tu mierzonego najczęściej liczbą świec S_p wprzód do zamknięcia pozycji.

Kryteria stosowane w badaniach:

W (Wynik) – to suma zwrotów po zamknięciu pozycji przez bariery TP i SL (od wyniku zamknięć na barierze TP - zysków, odejmowane są wyniki zamknięć na barierze SL - strat).

Wp (Wynpopr) – wynik poprawiony (skorygowany) o rezultaty zamknięć na barierze T_{dop} – spowodowane osiągnięciem czasu wygaśnięcia pozycji. Wyniki te mogą być zarówno dodatnie jak i zakończone stratą,

Lmax (Obsm) – maksymalna wielkość obsunięcia kapitału (lub badanej zmiennej obserwowanej będącej odpowiednikiem kapitału np. wyrażonej w pipsach) w badanym okresie;

K1 – stosunek obsunięcia *Lmax* do *Wp*;

K2 - kryterium uwzględniające jednocześnie wynik poprawiony i ryzyko porażki wyrażone poprzez 2-krotne zwiększenia maksymalnego obsunięcia, czyli K2= wynpopr-2*obsm;

K10 – kryterium uwzględniające jednocześnie wynik poprawiony i ryzyko porażki wyrażone poprzez 10-krotne zwiększenia maksymalnego obsunięcia, czyli K10= wynpopr-10*obsm;

K20 – kryterium uwzględniające jednocześnie wynik poprawiony i ryzyko porażki wyrażone poprzez 5-krotne zwiększenia maksymalnego obsunięcia, czyli K10= wynpopr-5*obsm; to kryterium dla bardziej ryzykownego inwestora niż K10.

Wymienione kryteria tworzą więc specyficzna mapę zysku i ryzyka, przy czym ryzyko wyrażane jest tu najczęściej poprzez awersję inwestora do dużych obsunięć kapitału. W dalszych rozdziałach rozważany będzie także typowe dla np. analizy portfelowej rozważanie ryzyka wyrażonego poprzez wariancję rezultatów (zwrotów) na tle średnich korzyści (np. w sposób stosowany przez Sharpe'a i Sortino).

3.1. Badania skuteczności indykatorów opartych na wyższych momentach statystycznych

Pierwszą rozpatrywaną parę zmiennych wejściowych stanowią momenty statystyczne trzeciego i czwartego rzędu – skośność i kurtoza.

Bieżąca skośność i kurtoza obliczana dla każdych ostatnich Sd=500 świec porównywane były z ustawianymi badaniach ich wartościami granicznymi. Momenty obliczane były względem średniej wartości ostatnich Sd wartości zamknięcia świec (badano jak zwykle rozkład kursu EUR/USD). Przekroczenie ustawionej badanej wartości granicznej przez bieżącą wartość skośności lub kurtozy stanowiło regułę otwarcia pozycji. Przyjmowane wartości graniczne w prezentowanych badaniach zmieniano odpowiednio w zakresie od -0.2 do 0.3 (co 0.1) dla **skośności,** wg reguły:

$$g_3 = -0.3+i_3*0.1 \tag{37}$$

dla $i_3 = 1, \ldots, 5$ oraz w zakresie od -1.3 do 1.5 dla **kurtozy**, wg reguły:

$$g_4 = -1.5+i_4*0.2 \tag{38}$$

dla $i_4 = 1, \ldots, 15$. Skośność i kurtoza obliczane były jak w skrypcie (Listing 6).

```
    sum3(i)=0;
    sum4(i)=0;
    for j=1:sd
        sum3(i)=sum3(i)+(C((i-sd+j),4)-mi(i))^3;
        sum4(i)=sum4(i)+(C((i-sd+j),4)-mi(i))^4;
    end
    skos(i)=sum3(i)/sd/s(i)^3;      %skosnosc
    kurt(i)=sum4(i)/sd/s(i)^4-3;    %kurtoza
```

Listing 6. Skrypt wyjaśniający sposób obliczania momentów statystycznych wyższych rzędów do testowania skuteczności strategii inwestycyjnej.

Zgodnie z zapisem w listingu skośność jest rozumiana tu jako wskaźnik asymetrii A_i rozkładu wartości zamknięcia świec wokół średniej x_{sr} z ostatnich Sd świec (w skrypcie zapisana jako mi(i)).

Wartość ta podzielona jest przez odchylenie standardowe podniesione do trzeciej potęgi:

$$A_i = \Sigma (x_k - x_{sr})^3 / s^3 / S_d \qquad (39)$$

gdzie:

i – indeks bieżącej chwili czasowej;

x_k – wartość zamknięcia k-tej świecy poprzedzającej chwilę bieżącą; k = 1, 2, ..., S_d;

x_{sr} – średnia arytmetyczna ostatnich S_d zamknięć świec;

s – odchylenie standardowe ostatnich S_d świec względem średniej.

Podobnie o kurtozie zapisanej w przedstawionym skrypcie należałoby powiedzieć, że to wskaźnik koncentracji rozkładu, gdyż podzielona jest przez czwarta potęgę odchylenia standardowego i dla porównania z rozkładem normalnym pomniejszona o 3.

$$K_i = \Sigma (x_k - x_{sr})^4 / s^4 / S_d \; - 3 \qquad (40)$$

Oznaczenie jak wyżej przy wzorze (39).

Decyzje inwestycyjne o otwarciu pozycji długich, krótkich lub zaniechaniu działania podejmowane były wówczas, gdy funkcja charakterystyczna zmiennej x^1 związanej z rozpatrywanymi wielkościami skośności i kurtozy przyjmowała odpowiednie wartości:

$$f(x_j) = \begin{cases} 1 \text{ gdy } A_i > g_{3(i3)} \text{ oraz } K_i > g_{4(i4)} \\ 0 \text{ w pozostałych przypadkach} \\ -1 \text{ gdy } A_i < -g_{3(i3)} \text{ oraz } K_i < -g_{4(i4)} \end{cases} \quad (41)$$

dla i = 1, 2, ... , S_d; i_3 = 1, ... , 6; i_4 = 1, ... , 15.

Pozycje długie otwierane były, gdy $f(x_i^1) = 1$ a krótkie, gdy $f(x_i^1) = -1$.

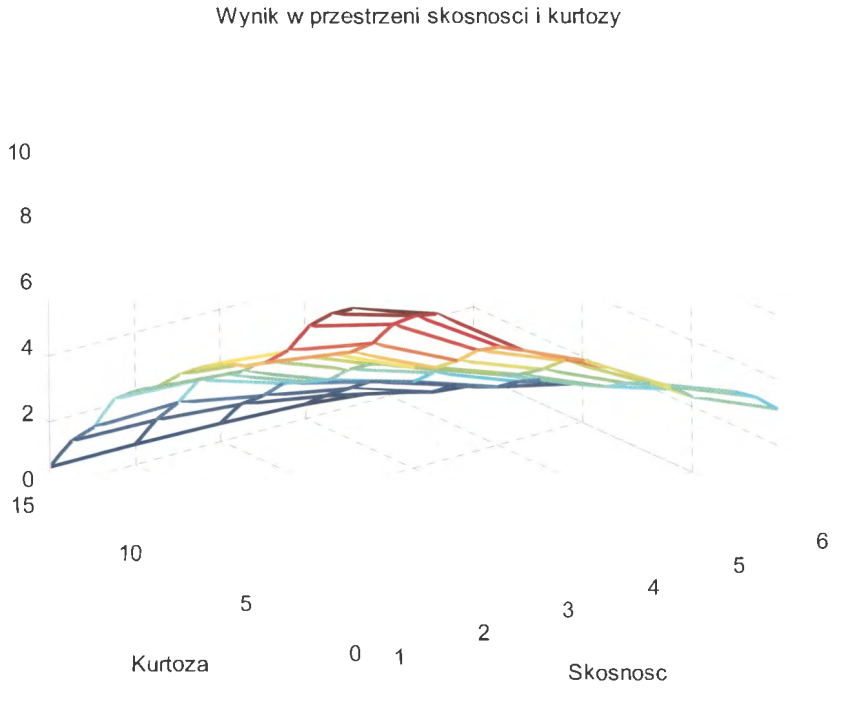
Badania przeprowadzono przy stałych pozostałych parametrach takich jak:

- liczba świec, dla których kolejno inicjowano badanie okna o liczności Sd=500 świec wynosiła ls=25000; (wartością początkową szeregu czasowego była świeca nr 5000, a końcową świeca o numerze 30000);
- wartość bariery *Stop Loss*, po osiągnięciu której zamykana była pozycja otwarta ze stratą SL= 0.0140 (140 pipsów);
- wartość bariery *Take Profit*, po osiągnięciu której zamykana była pozycja otwarta z zyskiem TP=0.0040 (40 pipsów);
- liczba świec symbolizująca dopuszczalny czas przetrzymywania otwartej pozycji opisana wyżej jako T_{dop}, tu przyjęta w formie Sp świec 15 min; przyjęto Sp=100;
- rząd kwantyla zewnętrznego, który rozstrzygał o zmianie kierunku otwarcia pozycji (będzie omówiony w podrozdziale o wektorze kwantyli); tu przyjęto $p(Q_{g2}) = 0.98$;
- rząd kwantyla wewnętrznego, który rozstrzygał o otwarciu pozycji (będzie omówiony w podrozdziale o wektorze kwantyli); tu przyjęto $p(Q_{g1}) = 0.56$;
- oscylator stochastyczny Lane'a; przyjęto jako barierę otwierającą pozycję L=35.;

Powyższe parametry czynników do tej pory nierozpatrywanych nie są dobrane przypadkowo, lecz wynikają z kilkuset wstępnych testów przeprowadzonych przed ostatecznym sukcesywnym objaśnieniem ich wpływu.

Wyniki badań przedstawione na rys. 46 dotyczą sytuacji, gdy kolejno każda z 25000 rozważanych świec powodowała otwarcie jakiejś pozycji lub bezczynność. Każda otwarta pozycja była zamykana tylko przez bariery TP i SL, nie uruchamiano zamykania poprzez wygaśnięcie pozycji po upływie zadanego czasu $T_{dop.}$ Wynik przedstawiany jest jako suma pipsów uzyskanych przy udanych zamknięciach na TP pomniejszona o straty na SL. Przy tak asymetrycznie i raczej nietypowo usta-

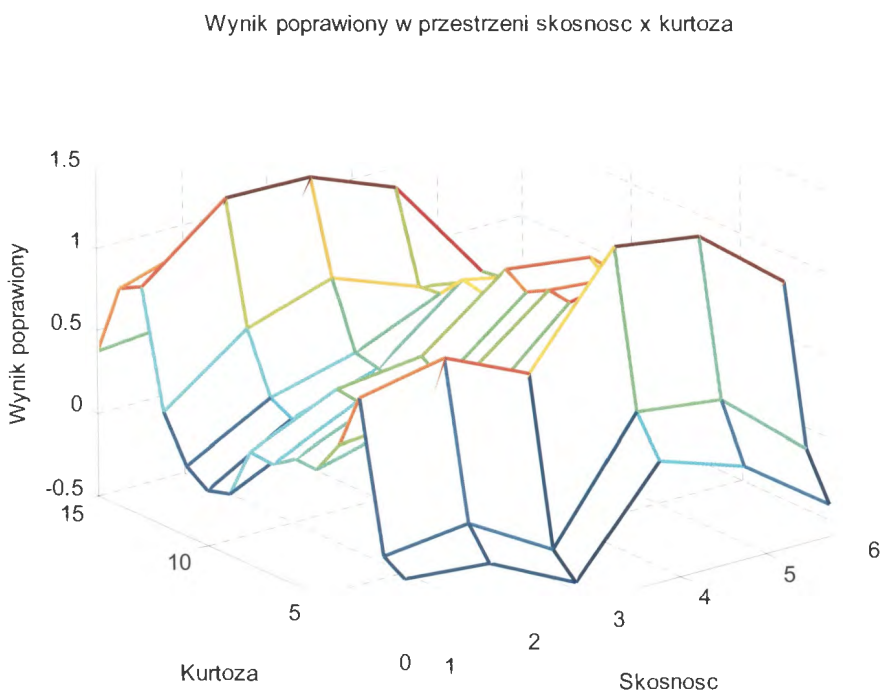
wionych barierach (TP=40 pipsów a SL=140 pipsów) uzyskać można, zwłaszcza dla początkowych wartości zakresów zmienności badanej skośności i kurtozy, dość dobre wyniki rzędu kilkudziesięciu tysięcy pipsów (wartość 1.0 na osi rzędnych rys. 46 to 10000 pipsów. Wyniki tak przedstawione nie ujawniają jednak ponoszonego przez inwestora ryzyka. Bardzo dobry rezultat może być obarczony dużą zmiennością wyników cząstkowych, w tym także porażek sekwencyjnie skupionych obok siebie, skutkujących dużymi obsunięciami kapitału.



Rys. 46. Wyniki symulacji w przestrzeni zmiennej skośności (1:6) i kurtozy (1:15). Rezultat – wynik końcowy (w skrypcie nazywany wynik) od zamknięć na barierach SL i TP oraz dopuszczalnego czasu otwarcia pozycji.

Pewnym zabezpieczeniem przed kilkoma dużymi kolejnymi porażkami na barierze SL jest wprowadzenie czasu wymuszonego zamknięcia $T_{dop.}$

Wyniki uzyskiwane po wprowadzeniu tej bariery zwane poprawionymi *wynpopr* przedstawione sa na kolejnym wykresie rys. 47.



Rys. 47. Wyniki symulacji w przestrzeni zmiennej skośności (1:6) i kurtozy (1:15). Rezultat – wynik końcowy poprawiony (wynpopr) od zamknięć na barierach SL i TP z uwzględnieniem zamknięć wskutek upływu zadanego czasu otwarcia pozycji.

Charakterystyczne dla wyników przedstawionych na rys. 47 są dwa odrębne obszary rozwiązań bliskich optymalnym – to obszar środkowych wartości skośności (dla indeksu 3 i 4 co odpowiada skośności 0 i 0.1) oraz dwóm obszarom kurtozy – bardzo niskiej i bardzo wysokiej (albo w okolicach wartości -1.0 albo 1.0). To ważna sugestia dla dalszych rozważań – kurtoza charakterystyczna dla rozkładu normalnego nie jest atrakcyjna.

Należy przypomnieć, że kurtoza obliczana jest jako odniesiona do kurtozy rozkładu Gaussa (jako wskaźnik koncentracji rozkładu) i pokrywa się z kurioza gaussowską przy wartości K=0 (wg 40).

Porównanie wyników na wykresach 46 i 47 nasuwa wniosek o znacznym pogorszeniu rezultatów po wprowadzeniu T_{dop}. Przecież na wykresie wyniku niepoprawionego (rys. 44) najlepsze rezultaty sięgają 80 tys. pipsów na wykresie 47. zaledwie kilkunastu tysięcy. Na obu wykresach nie przedstawiono jednak wpływu ryzyka na wiarygodność

tych rezultatów. Zgodnie z naturą otaczającego nas świata oczekiwać należy, że przy wysokich zyskach pojawi się wysoka zmienność składników, z których zyski te powstają.

Miarą tego ryzyka powszechnie przyjętą w analizie szeregów czasowych zwrotów z inwestycji jest odchylenie standardowe (wyrażane np. poprzez wskaźnik Sharpe'a [106]) lub zestaw pewnych miar ryzyka, charakterystyczny dla rynku Forex [141]. Jedną z najważniejszych w tym zestawie miar jest maksymalne obsuniecie kapitału, wyrażane w pieniądzach lub mierze względnej, często także jako procent uzyskiwanego profitu. Ważne jest i często także mierzone miejsce występowania tego obsunięcia (na początku gromadzenia kapitału – to źle, lub pod koniec – zdecydowanie lepiej), mierzone są prawdopodobieństwa wystąpienia zysku i porażki, średnia wielkość zysku i porażki i temu podobne.

Na rys. 48 wybrano i przedstawiono jedną z tych miar – maksymalna wielkość obsunięcia, mierzona tak jak zysk – w pipsach, w rozpatrywanej tu przestrzeni różnych zadawanych granic skośności i kurtozy.

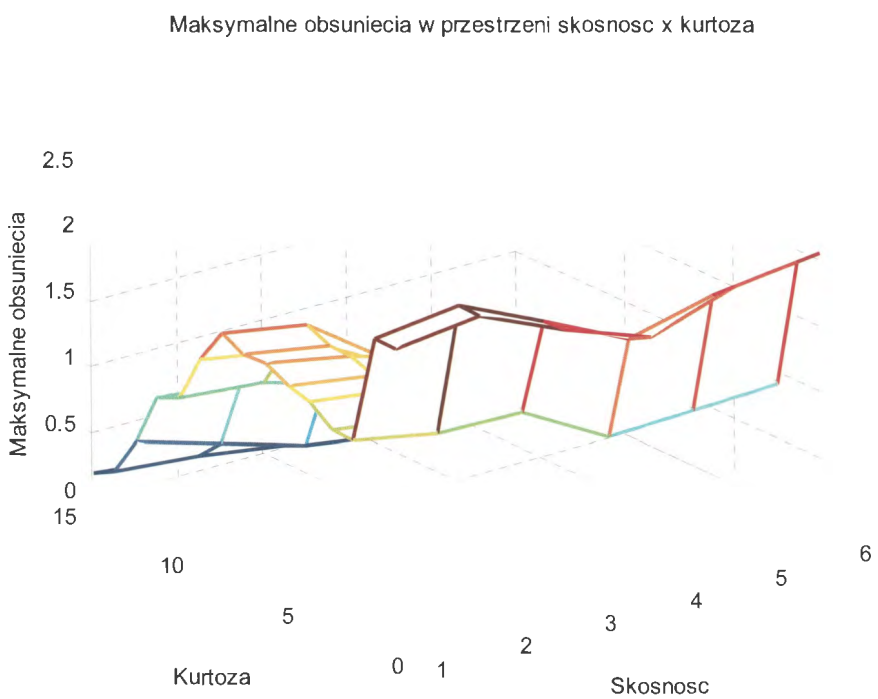
Wykres przedstawiający obsunięcia maksymalne w rozpatrywanej przestrzeni wyjaśnia, dlaczego wynik uwzględniający ryzyko nie będzie miał charakteru unimodalnego. Podobnie bowiem jak na rys. 46 wysokim wartościom wyniku niepoprawionego towarzyszą duże obsunięcia, a tam gdzie obsunięcia są atrakcyjnie małe, tam także *wynik* jest niższy.

Najlepsze wyniki dla kryterium *wynpopr* uzyskano dla granicy skośności wynoszącej $g_3 = 0.1$ oraz dla granicy kurtozy $g_4 = 1.1$.

Najlepsze wartości wg kryterium minimum ryzyka K1 (minimalizacji maksymalnego obsunięcia) wynoszą kolejno *wynik, wynpopr, obsm, k1*. Kryterium K1 to stosunek obsuniecia maksymalnego do wyniku poprawionego. W rozpatrywanym tu przypadku najlepszy wynik kształtuje się na poziomie 10%.

Tab. 5. Wyniki symulacji w przestrzeni skośność x kurtoza.

Badane kryterium	Wynik bez T_{dop} *wynik*	Wynik poprawiony *Wynpopr*	Maksymalne obsunięcie kapitału *obsm*	Stosunek obsunięcia do wyniku K1
Wynik [pips x 10^4]	1.5059	1.0753	0.1180	0.1097



*Rys. 48. Wyniki symulacji w przestrzeni zmiennej skośności (1:6) i kurtozy (1:15). Rezultat – wielkości maksymalnych obsunięć (*obsm*).*

W powyższej tabeli występują dwa charakterystyczne kryteria tego rynku (i zapewne wszystkich rynków, a może nawet wszystkich zachowań inwestorów). To kryterium maksymalizacji zysku (tu – *wynpopr*) i minimalizacji ryzyka (maksymalnego obsunięcia – *obsm*). To także kryterium ujarzmienia chciwości i opanowania strachu – stąd dygresja o ich ponadrynkowym znaczeniu.

Powstał dylemat, które kryterium jest ważniejsze lub jak je ważyć. W ten sposób powstała idea kryterium wiążących oba wyniki – K10 i K20. Kryteria te zostały zdefiniowane na wstępie do niniejszego rozdziału.

W badaniach stosowano zasadę ostrożnego podejścia do symulowanych wyników. Wyrażało się ono m.in. w „pesymistycznym" traktowaniu zmian kursu waluty wewnątrz rozpatrywanej świecy. Poniższy fragment skryptu (Listing 7) wyjaśnia, na czym polega wariant „pesymistyczny" rozpatrywania reguł otwarcia pozycji. Rozpatrywane są kolejno dwa warunki – w pierwszym sprawdza się, czy kolejno rozpatrywana

l-ta wartość świecy (wprzód) nie spowoduje przegranej a dopiero potem, gdy warunek ten nie jest spełniony (zmienna *stop1* dalej równa jest 0), rozpatrywane jest pytanie, czy nie została osiągnięta bariera TP. Ta kolejność rozpatrywania barier powoduje, że wynik rzeczywisty nigdy nie może być gorszy niż pesymistyczny, ten ostatni stanowi więc dolne oszacowanie wyniku końcowego (wg zasady „gorzej nie może być"). We wszystkich przeprowadzanych symulacjach uwzględniany jest 3-pipsowy spread na parze EUR/USD.

```
for l=1:sp
    if (C(k+l,2)-C(k,4))>SL & loss(k)==0 & stop1==0 &
    plon(k)>klg2 & pDslow(k)>100-Lanerec
    %& ~ww %otwieramy short
        loss(k)=1;
        stop1=1;
        wynk(k)=-SL-spread;
        klasa(1,2)=klasa(1,2)+1;
    end
    if (-C(k+l,3)+C(k,4))>TP & win(k)==0 & stop1==0 &
    plon(k)>klg2 & pDslow(k)>100-Lanerec
    %& ~ww %otwieramy short
        win(k)=1;
        stop1=1;
        wynk(k)=TP-spread;
        klasa(1,1)=klasa(1,1)+1;
    end
```

Listing 7. Skrypt wyjaśniający kolejność rozpatrywania zmian kursu wewnątrz świecy.

3.2. Badanie wpływu oscylatora Lane'a na skuteczność inwestowania

Indykator Lane'a [129, 139, 141] zwany najczęściej Stochastycznym (ang. *Stochastic*), zdefiniowany przez Georga Lane w latach 50-tych ub. wieku, wybrany został spośród wielu innych popularnych wskaźników stosowanych w szczególności w platformach Foreksu. Rezygnacja w niniejszych badaniach z kilku innych indykatorów poprzedzona była wspomnianymi już wstępnymi testami ich przydatności w rozważanej przestrzeni sygnałów. Oscylator stochastyczny obliczany był, po kilku próbach na świecach 15 min, dla l_{tyl} = 50 świec wstecz (nie, jak najczęściej

zalecają źródła literaturowe dla 14 okresów wstecz). Składniki %K i %D charakterystyczne dla tego oscylatora obliczane były jak w listingu 8.

Podobnie jak dla skośności i kurtozy wprowadzono dobieraną granicę indykatora g_{Lane}. Badania przeprowadzono przy zmiennej granicy funkcji charakterystycznej g_{Lane} i stałych pozostałych parametrach takich jak w poprzednich badaniach, z tym, że granice skośności g_3 i kurtozy g_4 ustawione już zostały na poziomie ustalonym w poprzednim podrozdziale jako optymalny. Tak więc kolejno:

- liczba świec, od których inicjowano badanie okna o liczności S_d = 500 świec wynosiła ls=25000;
- bariera SL= 0.0140;
- barieraTP=0.0040;
- rząd kwantyla zewnętrznego $p(Q_{g2})$ = 0.98;
- rząd kwantyla wewnętrznego $p(Q_{g1})$ = 0.56;
- granica skośności g_3 = 0.1;
- granica kurtozy g_4 =1.1;
- zas dopuszczalny otwarcia pozycji wyrażony liczba świec wprzód S_p =120;
- l_{tyl} = 50; liczba świec wstecz, dla której budowano wskaźniki Lane'a.

Wskaźnik Lane'a jako pDslow obliczany był wg skryptu (w pętli dla każdej świecy spośród ls=25000).

```
ltyl=50;
minn(i)=min(C(i-ltyl:i,3));
maxn(i)=max(C(i-ltyl:i,2));
pKfast(i)=100*(C(i,4)-minn(i))/(maxn(i)-minn(i));
if i>ipocz+3
      pDfast(i)=(pKfast(i)+pKfast(i-1)+pKfast(i-2))/3;
end
pKslow(i)=pDfast(i);
if i>ipocz+3
      pDslow(i)=(pKslow(i)+pKslow(i-1)+pKslow(i-2))/3;
end
```

Listing 8. Skrypt wyjaśniający sposób obliczania indykatora stochastycznego Lane'a.

W regułach otwierających pozycje (sygnałach) porównywano pDslow (%Ds) z barierą g_{Lane} zmienianą w tym badaniu w pętli j=1:20 od wartości 5 do 100 (%).

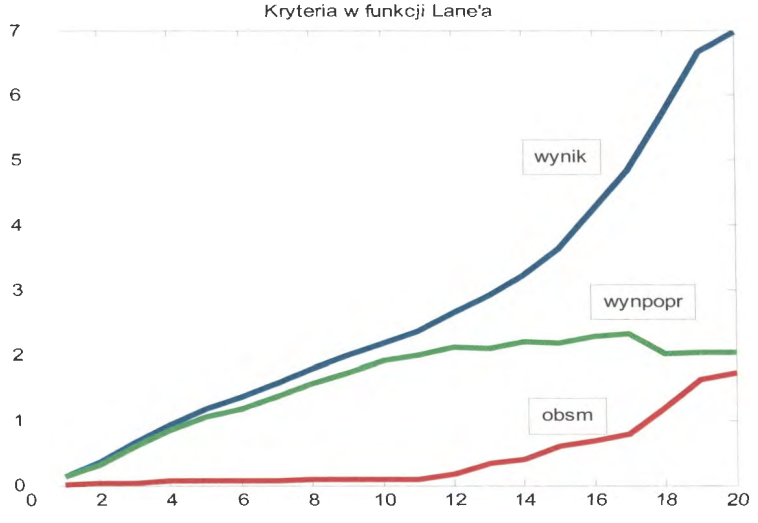
Stosowano regułę decyzyjną:

$$f(x_j) = \begin{cases} 1 & \text{gdy } \%Ds_i > g_{Lane(j)} \\ 0 & \text{w pozostałych przypadkach} \\ -1 & \text{gdy } \%Ds_i < 100 - g_{Lane(j)} \end{cases} \quad (42)$$

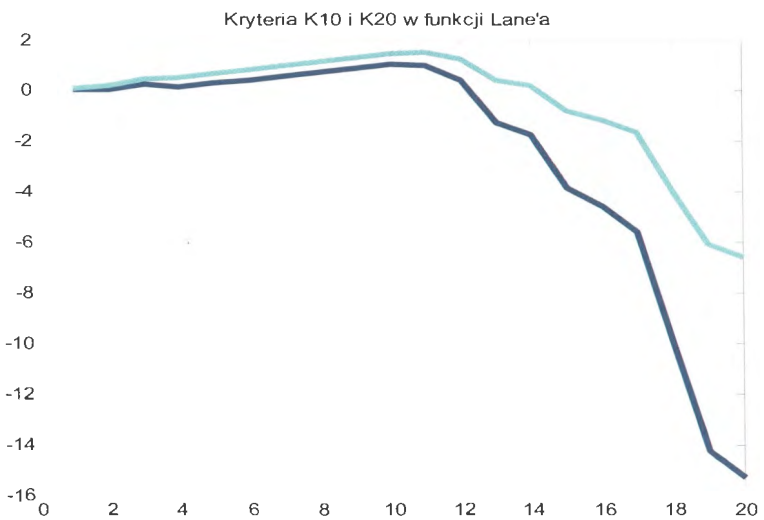
Pozycje długie otwierane były, gdy $f(x_i^2) = 1$ a krótkie, gdy $f(x_i^2) = -1$.

Z uwzględnieniem poprzednich barier skośności i kurtozy pozycje otwierane były w sytuacji jednoczesnego spełnienia warunków określonych przez (41) i (42). Kolejne koniunktywnie uwzględniane warunki wymieniane będą w kolejnych rozpatrywanych wymiarach przestrzeni sygnałów rekomendujących otwarcie długich i krótkich pozycji rozpatrywanego waloru walutowego.

Uzyskano wyniki przedstawione na rys. 49 i 50. Najlepsze wyniki uzyskano dla wartości indykatora wynoszącej 50. To zupełnie inna wartość niż sugerowana przez wiele platform brokerskich – zazwyczaj 20-30 [141]. Wg kryterium minimaksowego K1 uzyskano wyniki jak poniżej:



Rys. 49. Wyniki wynik (nieb) i wynpopr (ziel) oraz obsunięcie w funkcji indykatora Lane'a dla wartości indykatora od 5 do 100.



Rys. 50. Kryteria K10 (niebieski) i K20 (cyjan) w funkcji indykatora Stochastic Oscillator Lane'a zmieniającego się od 5 do 100. Maksimum występuje dla obu kryteriów przy wartości Lane=50 (10x5).

Tab. 6. Wyniki symulacji dla optymalnej wartości indykatora Lane'a.

Badane kryterium	Wynik bez T_{dop} *wynik*	Wynik poprawiony *wynpopr*	Maksymalne obsunięcie kapitału *obsm*	Stosunek obsunięcia do wyniku K1
Wynik [pips x 10^4]	2.1872	1.9081	0.0899	0.0471

Są to wyniki ostatnie lepsze niż w poprzednim badaniu w przestrzeni dwóch wyższych momentów statystycznych. Obsunięcie w skali całego niemal roku na poziomie 4-5% jest znakomitym wynikiem. Wynik bezwzględny rzędu 20 tys. pipsów także byłby znakomitym dokonaniem.

Badania te przeprowadzono przy założeniu, że indykator Lane'a będzie obliczany dla liczby świec rozpatrywanych wstecz $l_{tyl} = 50$. Oznacza to, że każdy kolejny komponent algorytmu Lane'a pKfast i pKslow i następne pochodne obliczone zostały z uwzględnieniem ostatnich 50 świec. Po zmianie tego parametru kolejno na 40 i 60 uzyskano wyniki przedstawione w tab. 7.

Tab. 7. Wyniki symulacji dla sąsiednich wartości indykatora Lane'a.

Badane kryterium	Wynik bez T_{dop} *wynik*	Wynik poprawiony *wynpopr*	Maksymalne obsunięcie kapitału *obsm*	Stosunek obsunięcia do wyniku - K1
Wynik dla l_{tyl}=40	0.1184	0.1066	0.0063	0.0591
Wynik dla l_{tyl}=60	1.9763	1.7267	0.0783	0.0453

Pierwszy z tych wyników jest ewidentnie gorszy od wymienionego jako bazowy dla l_{tyl} = 50. Drugie badanie daje gorszy wynik i nieco lepsza wartość kryterium obsunięcia. Uznając, że zmiany te są mało istotne, do dalszych badań pozostawia się wartość l_{tyl} = 50 świec.

3.3. Badanie wpływu wektora rzędów kwantyli

To jedna z najbardziej innowacyjnych i wrażliwych podprzestrzeni w rozpatrywanej przestrzeni sygnałów. Zgodnie z opublikowaną wcześniej autorską koncepcją [126, 128] stawiana jest tu teza, że bieżąca zmiana ceny zależy od specyficznie określonej przeszłości. Na podstawie wstępnych badań okazuje się, że pamięć niezbędna do zbudowania stosunkowo skutecznego systemu transakcyjnego jest dość długa i sięga kilkuset zdarzeń (świec) wstecz. Przyjęto, że, podobnie jak w dotychczasowych podprzestrzeniach, rozpatrywany będzie szereg S_d = 500 wartości wstecz od chwili bieżącej. Podobnie jak w (39) x_k oznacza k-tą świecę wstecz od chwili bieżącej xi, a x_{sr} średnią z ostatnich S_d = 500 wartości zamknięć świec.

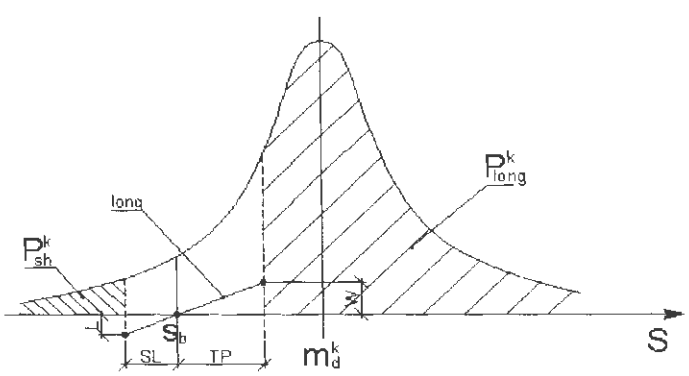
Rozpatrując wszystkie ostatnie S_d zamknięć można zdefiniować dla każdej chwili bieżącej pewien umowny rozkład normalny odchyleń ceny wokół jej średniej o parametrach x_{sr} i s (średnia i odchylenie standardowe jak w (39)). Dokładniej, pamiętając o oznaczeniu przez indeks *i* chwili bieżącej należałoby zapisać x^i_{sr} oraz s^i; dla uproszczenia zapisów indeksy górne są pomijane. Oczywiście rozkłady rzeczywiste często dalekie są od normalnych jednak te ostatnie traktowane są tu jako pewien szczególny indykator rozkładu. Na ich tle budowane są pewne klasy zachowań inwestycyjnych.

Umieszczając wartość bieżącą ceny rozpatrywanego waloru (tu kursu EUR/USD) na tle rozkładu normalnego o parametrach estymowanych przez rzeczywiste wartości średniej i wariancji, można próbować wyciągnąć wniosek o prawdopodobieństwie zmiany ceny. Gdyby ob-

serwowane procesy zmiany ceny były markowskimi procesami stacjonarnymi tj. takimi, dla których stały byłby rozkład prawdopodobieństwa osiągnięcia jednego z dopuszczalnych stanów ze stanu bieżącego, problem ten byłby stosunkowo prosty i dawno byłoby zauważone jego rozwiązanie. Tak jednak, na szczęście (dla rozwoju badań), nie jest. Rzeczywistość nie daje żadnych gwarancji stacjonarności i ciągle czyni rozpatrywane tu zadania predykcyjne jednymi z bardziej fascynujących problemów ludzkości.

Potwierdzeniem takiej niekonsekwencji w zachowaniu rynków są wielokrotnie obserwowane zdarzenia „odstające" od „prawdopodobnych" uzewnętrzniające się w rozkładach rzeczywistych pojawianiem się tzw. „grubych ogonów". Np. po bardzo silnych wzrostach kursu zdarzają się jeszcze silniejsze, a po bardzo silnych spadkach – dalsze spadki. Zachowanie takie narusza naturalny obraz, także zauważalny, zachowania się rynku w sposób przypominający zachowanie się pijanego mężczyzny [17], który ma gdzieś w świadomości ogólny kierunek ruchu jednak pozwala sobie na permanentne, losowe odchylenia od tej trajektorii. Rozpatrywaną tu koncepcje porównać można, pozostając przy metaforze pijaka, raczej do takiego jego zachowania, gdy co pewien czas w jego świadomości kiełkuje nowa idea wędrówki w całkowicie odmiennym kierunku. Nie można było przy tym przewidzieć, co też niespodziewanego pojawi się w tym umyśle. Oto istota niestacjonarności procesu i takie są rzeczywiste zachowania, także rozpatrywanego tu najczęściej rynku par walutowych.

Istotę prawdopodobieństwa kolejnej zmiany kierunku ruchu rynku na tle niedawnej historii oddaje rys. 51.



Rys. 51. Hipotetyczny rozkład normalny ostatnich S_d odchyleń zamknięć od średniej m_d.

Niech będzie dany rozkład normalny odchyleń ostatnich S_d zamknięć od średniej z tych zamknięć oznaczonej m_d zastępujący rozkład rzeczywisty tych odchyleń.

Bieżąca wartość obserwowanej zmiennej niech będzie równa S_b. Na wykresie jest ona położona poniżej średniej (z lewej strony są niższe wartości zmiennej). Gdyby rozkład prawdopodobieństwa nie zmienił się w kolejnym kroku, czyli gdyby obserwowany proces był procesem markowskim, to prawdopodobieństwo wzrostu rynku jest oczywiście większe niż prawdopodobieństwo spadku. Na rysunku dodatkowo zaznaczono dwie bariery, typowe dla handlu na platformach brokerskich – barierę TP i SL. Gdyby kierować się sugestią utrzymania rozkładu prawdopodobieństwa to naturalne byłoby otwarcie pozycji długiej w chwili bieżącej. Gdyby przy tym, tak otwartej pozycji towarzyszyły odpowiednie bariery o wartościach jak na rys. 51, to można zauważyć, że prawdopodobieństwo osiągnięcia TP (oznaczone jako pole P^k_{long}) jest wielokrotnie większe niż prawdopodobieństwo osiągnięcia SL (oznaczone jako pole P^k_{sh}).

Oczywiście zakłada się, że bariery SL i TP skojarzone są tzw. zleceniem OCO (*One Cancel Other*), czyli zamkniecie pozycji w którymkolwiek miejscu powoduje unieważnienie drugiej bariery. Znaczenie ma więc, w którą stronę rynek zmieni się najpierw.

Zakładając, że nastąpi zwrot kursu (*reversal*) w sytuacji S_b prawdopodobieństwo osiągnięcia co najmniej wzrostu o TP jest równe P^k_{long}. Jeżeli prawdopodobieństwo to jest dostatecznie duże warto zaryzykować taką inwestycję. Zapewne jest tak, że wartość S_b położona bliżej wartości średniej m^k, ma niewielką przewagę P^k_{long} nad P^k_{sh}, lub po prostu nie ma jej wcale. Logiczną reakcją na taki stosunek prawdopodobieństw P^k_{long} / P^k_{sh} będzie zaniechanie inwestycji. Może to zdarzyć się stosunkowo często w sytuacjach, gdy wartość bieżąca leży w okolicach wartości średniej tj. być może bardzo często. Zgodne to będzie jednak z przyjęta strategia, by inwestować w sytuacji wyraźnej przewagi jednego z prawdopodobieństw P^k_{long} lub P^k_{sh}. Podobna decyzja powinna być bowiem podjęta także po drugiej stronie wartości średniej – gdy dominuje wyraźnie P^k_{sh} nad P^k_{long}.

Można także analizując wykres na rys. 51 zauważyć, że gdyby zbyt optymistycznie przesunąć barierę TP mocno w prawo to istnieje gdzieś, statystycznie uzasadniony, kres takiego optymizmu. Każde bowiem zwiększenie bariery TP zmniejszy prawdopodobieństwo jej osiągnięcia. Podobne wnioski można sformułować w odniesieniu do SL

Wracając jednak do uwag o niestacjonarności procesu zmiany ceny instrumentu bazowego można postawić tezę, że zapewne nie

zwiększa się ciągle i nieuchronnie prawdopodobieństwo ruchu ceny w górę w miarę oddalania się wartości bieżącej od średniej w dół (w lewo). Gdyby tak było, to znaczy gdyby rosło prawdopodobieństwo ruchu powrotnego w funkcji oddalenia wartości bieżącej od średnie tzn. gdyby proces miał rozkład gęstości prawdopodobieństwa gaussowski, to bardzo prosta strategia inwestycyjna zapewniałaby zysk. Wystarczyłoby np. czekać wyłącznie na duże odchylenia ceny i otwierać pozycje w kierunku średniej. Skoro tak się nie dzieje, to być może istnieje jakaś bariera na rozkładzie odchyleń, po przekroczeniu której prawdopodobieństwo ruchu powrotnego przestaje mieć związek z dotychczasowym rozkładem zdarzeń (tych odnotowanych w S_d historii).

Niech Q_{gb} oznacza pewien kwantyl wyznaczony przez bieżąca wartość na osi ceny obserwowanej zmiennej umieszczony gdzieś w lewej części rozkładu gęstości a $p(Q_{gb})$ rząd tego kwantyla. Wracając do rys. 51 $p(Q_{gb})$ oznacza pole na wykresie rozkładu gęstości położone z lewej strony wartości bieżącej zmiennej obserwowanej.

Niech pole pod rozkładem gęstości będzie podzielone umownie na 5 części (rys. 52) oznaczonych odpowiednio C_b^i i=1,2,...,5 . Niech każda z tych części (klas) będzie wyznaczona przez dwie granice – kwantyl dolny (położony z lewej strony pola) i górny. Granice te będą odpowiednio oznaczone (S_d^i, S_g^i).

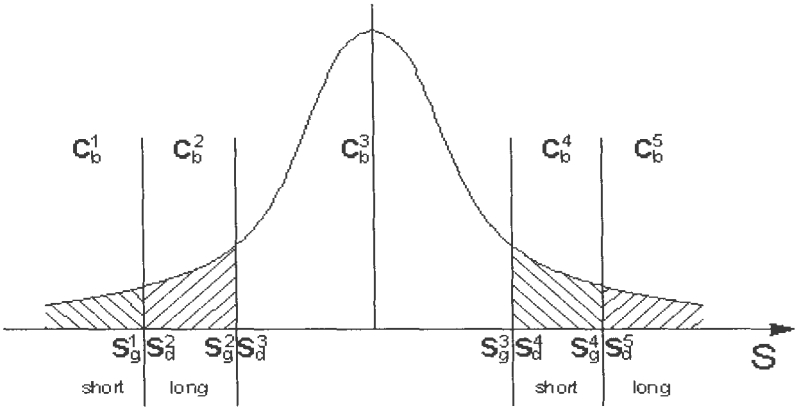
Stawia się tezę, że: (43)

1. jeżeli Q_{gb} należy do C_b^1 to korzystniej jest otworzyć pozycję krótką;
2. jeżeli Q_{gb} należy do C_b^2 to korzystniej jest otworzyć pozycję długą;
3. jeżeli Q_{gb} należy do C_b^3 to korzystniej jest zaniechać działania;
4. jeżeli Q_{gb} należy do C_b^4 to korzystniej jest otworzyć pozycję krótką;
5. jeżeli Q_{gb} należy do C_b^5 to korzystniej jest otworzyć pozycję długą;

Ponieważ w ustawicznym procesie zmieniającej się ceny zdecydowanie wygodniejsze jest operowanie nie kwantylami, lecz ich rzędami (miara w postaci wielkości pola z prawej strony kwantyla na wykresie rozkładu gęstości), to wprowadzono pojecie wektora rzędów kwantyli determinujących podział na wymienione w (43) klasy.

Na rys. 52 widać, że wystarczą 4 wartości w wektorze rzędów kwantyli, by podzielić przestrzeń pod rozkładem gęstości na 5 klas. Wartości te odpowiednio oznaczono:

$$P_q = [k_{lg2}\ k_{lg1}\ k_{ld1}\ k_{ld2}] \tag{44}$$



Rys. 52. Podział rozkładu gęstości odchyleń na klasy rekomendujące decyzje o otwarciu pozycji

Np. rząd kwantyla górnego zewnętrznego k_{lg2} odpowiada wielkości pola (prawdopodobieństwa) umieszczonego z prawej strony kwantyla wyznaczonego przez barierę S_g^1 i S_d^2.

Dokładniej, wartości k_{lg2} i dalej to rzędy kwanty li, których przekroczenie wywołuje decyzję. Rzędy kwanty li porównywane sa z prawdopodobieństwem wzrostu kursu od wartości bieżącej określonym na podstawie cech statystycznych historii ostatnich 500 świec. Przeprowadzono badania w kierunku określenia wartości wektora P_q, takich, dla których maksymalizowany jest zysk z inwestycji.

Badania przeprowadzono przy stałych pozostałych parametrach takich jak:

- liczba świec, od których inicjowano badanie okna o liczności S_d = 500 świec wynosiła ls=25000;
- bariera SL= 0.0140;
- barieraTP=0.0040;
- granica skośności g_3 = 0.1;
- granica kurtozy g_4 =1.1;
- czas dopuszczalny otwarcia pozycji wyrażony liczba świec wprzód S_p =120;
- l_{tyl} = 50; liczba świec wstecz, dla której budowano wskaźniki Lane'a;
- wskaźnik Lane'a – 50.

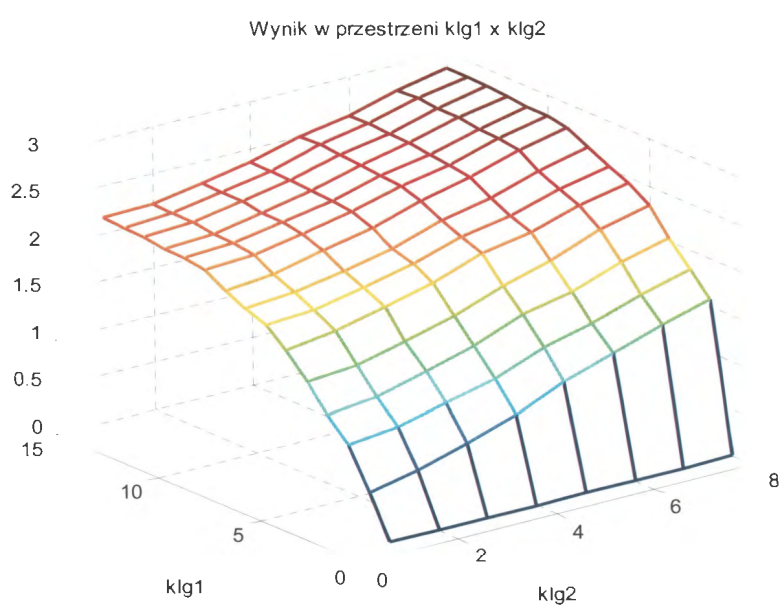
Badania przeprowadzono w przestrzeni k_{lg2} x k_{lg1}, gdzie k_{lg2} to rząd kwantyla zewnętrznego górnego klasy rekomendującej otwarcie pozycji short (występujący przy niskich położeniach wartości bieżącej kursu w stosunku do wartości średniej) a k_{lg1} to rząd kwantyla dolnego tej klasy. Podobnie odpowiednio dla k_{ld2} i k_{ld1} (symetrycznie).

Zmieniano k_{lg2} w pętli 1:8 od 1.00 co 0.005 do 0.96 oraz k_{lg1} w zakresie od k_{lg2} do wartości mniejszej od k_{lg2} o jj*0.03, gdzie jj zmieniano w pętli 2:15.

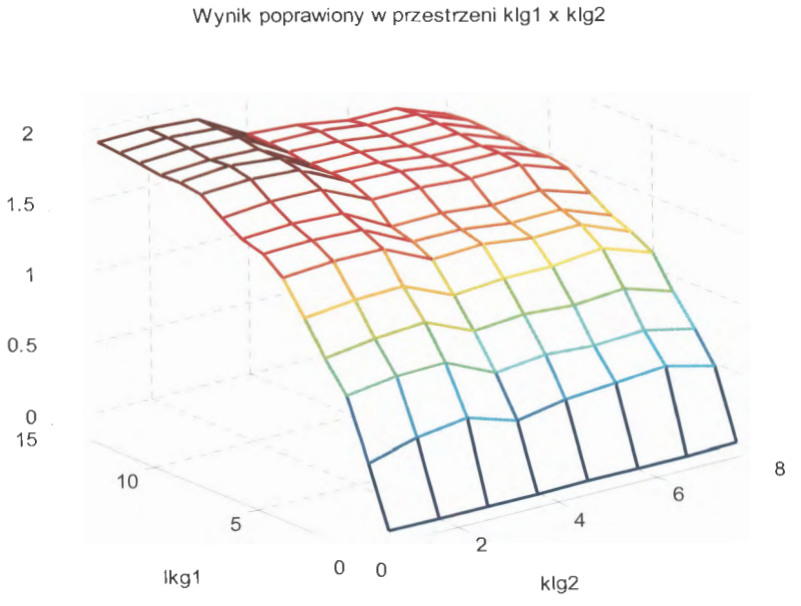
Uzyskano wyniki przedstawione na kolejnych wykresach – rys. 53, 54 i 55.

Mniejsze wartości klasy k_{lg2} wynikające z rys. 55 oznaczają preferencje dla małych różnic pomiędzy górnym i dolnym kwantylem. Oznacza to zapewne mniejsza częstość zdarzeń (sygnałów otwarcia), za to większe zapewne bezpieczeństwo transakcji.

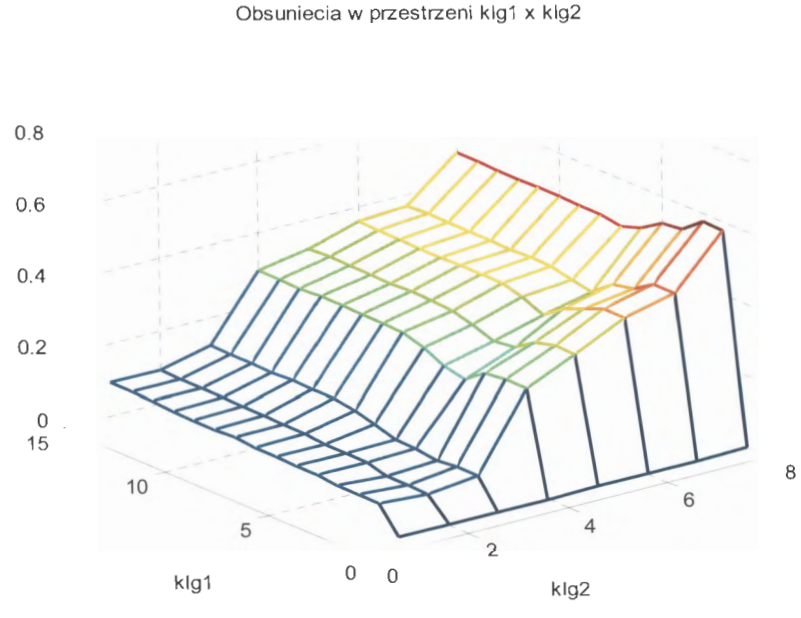
Widać to natychmiast na kolejnym wykresie kryterium uwzględniającego umownie obie wzajemne przeciwstawne cechy transakcji – wysoki wynik bezwzględny i niskie ryzyko. Właśnie w obszarze niskich obsunięć i stosunkowo dobrego wyniku poprawionego (rys. 56) występują maksymalne wartości K10.



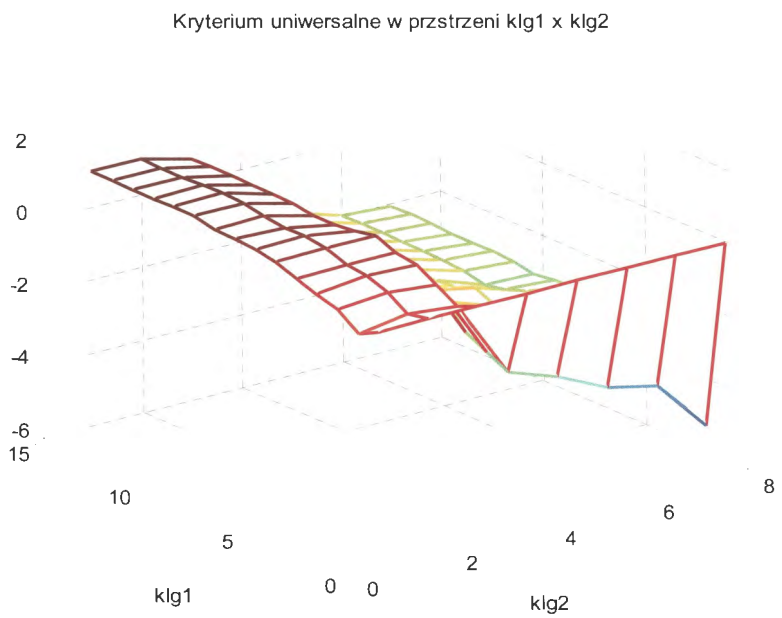
Rys. 53. Wykres przedstawia wynik w przestrzeni rzędów kwantyli k_{lg1} x k_{lg2}. Maksimum występuje dla maksymalnych wartości obu parametrów.



Rys. 54. Wykres przedstawia wynpopr w przestrzeni rzędów kwantyli k_{lg1} x k_{lg2}. Maksimum występuje dla minimalnej wartości k_{lg1} (k_{lg1} obliczane jest poprzez odejmowanie od k_{lg2}) oraz dla k_{lg2} rzędu 0.99.



Rys. 55. Wykres obsunięć w przestrzeni k_{lg1} x k_{lg2}. Widać wyraźnie korzystniejszy obszar dla mniejszych wartości k_{lg2}.



Rys. 56. Wykres kryterium K10 w przestrzeni rzędów kwantyli k_{lg1} x k_{lg2}. Maksimum występuje przy wartości k_{lg2}=0.99 i największej wartości k_{lg1}.

Tab. 8. Wyniki symulacji dla optymalnej wartości kryterium K10 w przestrzeni k_{lg1} x k_{lg2}.

Badane kryterium	Wynik bez T_{dop} *wynik*	Wynik poprawiony *wynpopr*	Maksymalne obsunięcie kapitału *obsm*	Stosunek obsunięcia do wyniku - K1
Wynik [pips x 10^4]	2.1354	1.8855	0.0855	0.0454

Wektor ostatecznych wyników tego badania przedstawiono w tab. 8. Wynik ten zbliżony jest do poprzedniego z tab.6. Ważnym rezultatem badań jest określenie interesujących przedziałów zmienności rzędów kwantyli.

3.4. Badanie wpływu barier zamknięcia pozycji TPxSL

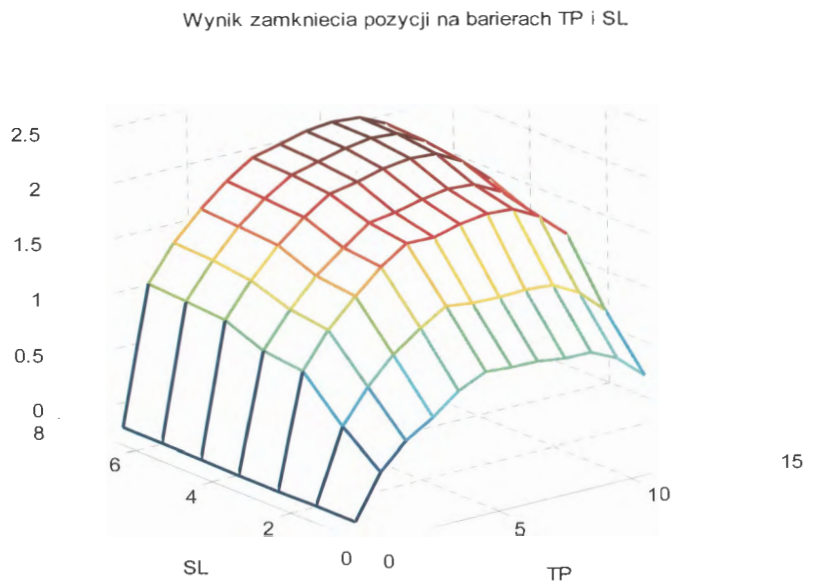
To szczególna dziedzina czynników oddziaływujących na rezultat inwestycji z uwagi na ich związek z technologią platform transakcyjnych. Są to bowiem czynniki determinujące sposób zamykania pozycji najczęściej ściśle związane z mechanizmami automatycznego lub półautomatycznego zarządzania transakcją. To podstawowa różnica w stosunku do dotychczas rozpatrywanych czynników związanych z warunkami otwarcia pozycji. Bariery ograniczania straty SL (*Stop Loss*) i zabezpieczenia zysku TP (*Take Profit*) są podstawowymi i powszechnie stosowanymi w większości platform internetowych.

Ta szczególna okoliczność pozwoli na wyjątkowo odrębne potraktowanie tych dwu czynników. Okaże się bowiem, że przeprowadzony eksperyment, polegający na poszukiwaniu najlepszych wartości zamknięć TP i SL nawet przy irracjonalnym założeniu, że przyszłość, w zakresie najlepszych wartości SL i TP, byłaby znana, prowadzi do bardzo dobrych rezultatów tradingu. Powstała w ten sposób koncepcja predykcji tych czynników, zamiast, o wiele trudniejszej, predykcji zmiany wartości badanego waloru.

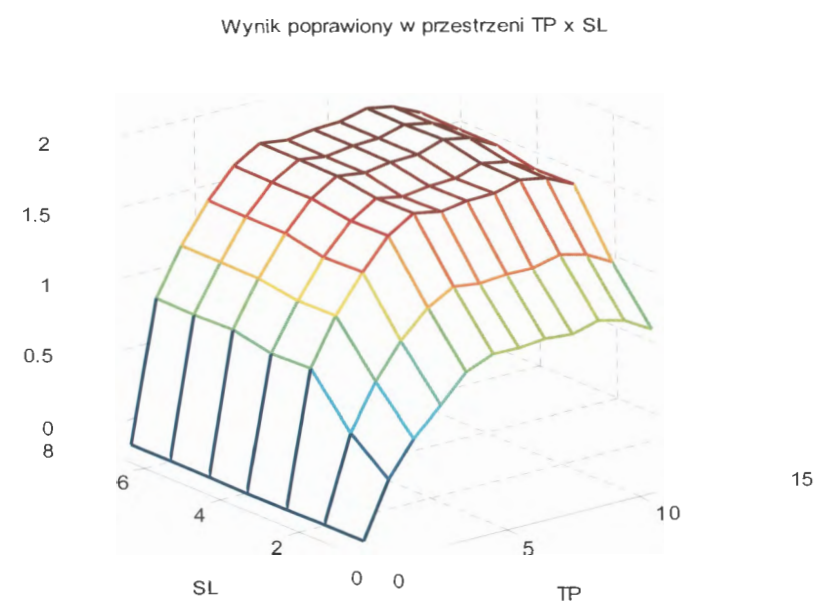
Badania przeprowadzono w ustalonej na podstawie dotychczasowych wyników przestrzeni parametrów, następująco ustalając ich wartości:

- liczba świec, od których inicjowano badanie okna o liczności $S_d = 500$ świec wynosiła ls=25000;
- granica skośności $g_3 = 0.1$;
- granica kurtozy $g_4 = 1.1$;
- czas dopuszczalny otwarcia pozycji wyrażony liczba świec wprzód $S_p = 120$;
- $l_{tyl} = 50$; liczba świec wstecz, dla której budowano wskaźniki Lane'a;
- wskaźnik Lane'a – 50;
- rząd kwantyla zewnętrznego $k_{lg2} = 0.99$;
- rząd kwantyla wewnętrznego $k_{lg1} = 0.66$.

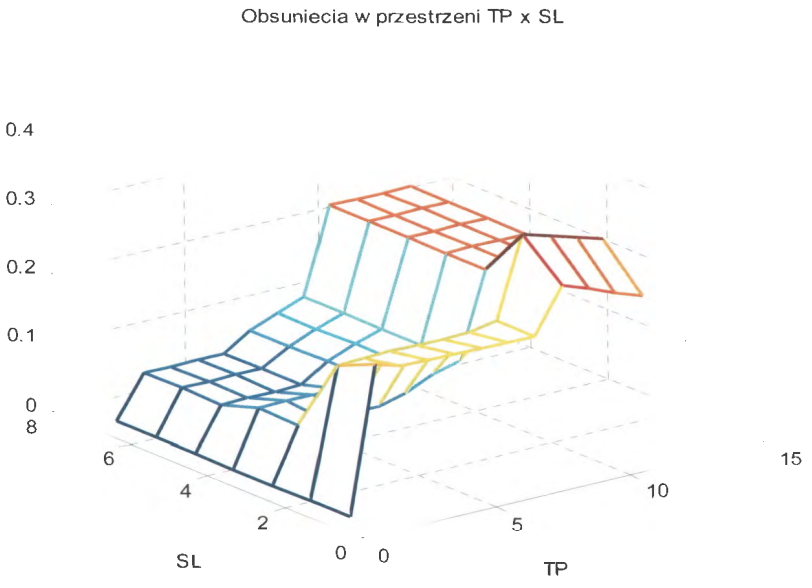
Poszukując najlepszych wartości SL i TP badania przeprowadzono dla SL od 0.0120 (120 pipsów pary walutowej EUR/USD) co 0.0020 (20 pipsów) do 0.0220 (220 pipsów) oraz dla TP od 0.0040 co 0.0005 do 0.0085. Wyniki przedstawiono na kolejnych wykresach.



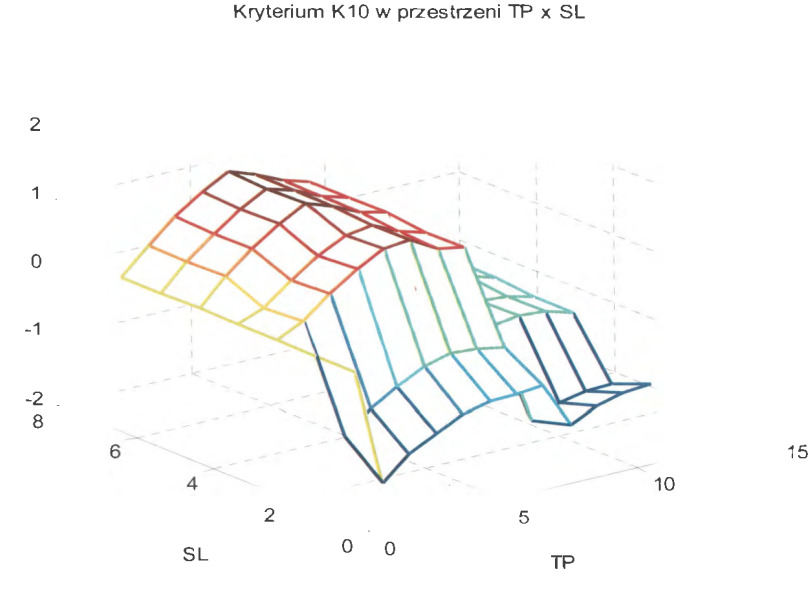
Rys. 57. Wykres wyniku niepoprawionego w przestrzeni SLxTP, gdzie SL zmienia się od 1:7 a TP – 2:12. Najwyższe wartości widać w obszarze SL=120-160 pipsów i TP=40-55 pipsów.



Rys. 58. Wykres wynpopr w przestrzeni SLxTP, gdzie SL zmienia się od 1:8. Najwyższe wartości widać podobnie jak poprzednio w obszarze najwyższych wartości SL oraz dla TP w obszarze 35-55 pipsów.



Rys. 59. Wykres obsunięć w przestrzeni SLxTP, gdzie SL zmienia się od 1:8. Najwyższe wartości widać dla początkowych wartości SL i dużych wartości TP.



Rys. 60. Wykres k10 w przestrzeni SLxTP, gdzie SL zmienia się od 1:8. Najwyższe wartości widać podobnie jak poprzednio w obszarze najwyższych rozpatrywanych wartości SL i TP=40 pipsów.

Tab. 9. Wyniki symulacji dla optymalnej wartości kryterium K10 w przestrzeni SL x TP.

Badane kryterium	Wynik bez T_{dop} *wynik*	Wynik poprawiony *wynpopr*	Maksymalne obsunięcie kapitału *obsm*	Stosunek obsunięcia do wyniku - K1
Wynik [pips x 10^4]	1.9444	1.7728	0.0609	0.0344

W wektorze wyników przedstawionym w tab. 9 występuje dotychczasowe minimum względnego obsunięcia K1=0.0344. Wynik ten uzyskano dla TP=0.0035 oraz SL=0.0120.

3.5. Badanie wpływu wielkości diagnostycznego okna czasowego – liczby świec badanych wstecz dla określenia rozkładu

Badania przeprowadzono dla najlepszych dotychczas obliczonych parametrów:

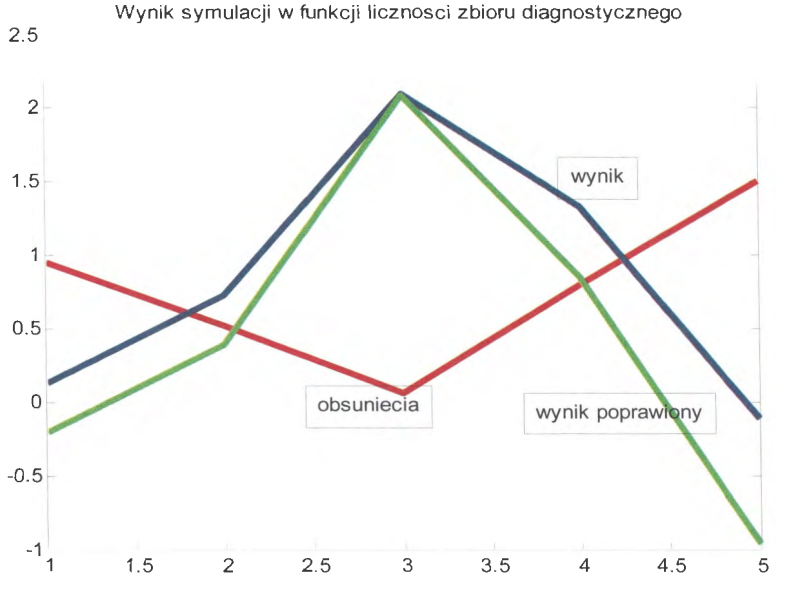
- liczba świec, od których inicjowano badanie okna o liczności S_d = 500 świec wynosiła ls=25000;
- granica skośności g_3 = 0.1;
- granica kurtozy g_4 =1.1;
- czas dopuszczalny otwarcia pozycji wyrażony liczba świec wprzód S_p =120;
- l_{tyl} = 50; - liczba świec wstecz, dla której budowano wskaźniki Lane'a;
- wskaźnik Lane'a – 50;
- rząd kwantyla zewnętrznego k_{lg2} = 0.99;
- rząd kwantyla wewnętrznego k_{lg1} = 0.66;
- bariera Stop Loss ustawiona została na poziomie SL=0.0120;
- bariera *Take Profit* ustawiona została na poziomie TP=0.0035.

Zmieniano liczność świec okna diagnostycznego, standardowo dla pozostałych badań w tym raporcie przyjętego sd=500 świec. Liczność te zmieniano dla *i* z przedziału (1:5) wg reguły 200+*i**100, czyli od 300 co 100 do 700. Wyniki przedstawione na rys.61 wskazują na dość mocną pozycję wybranego horyzontu 500 świec. Liczność pozostawiono

więc bez zmian. To kolejne poprawienie wyniku mierzonego wielkością obsunięcia uzyskane poprzez zmianę wielkości zbioru testowego (uczącego) tu zmienionego z dotychczasowej wartości 120 na 170 świec. Wyższa jest tu także wartość wyniku poprawionego, a więc w sensie równoczesnego uwzględnienia obu tych kryteriów parametry określone przez parę tych wyników są dominujące dla całego dotychczasowego zbioru wyników.

Tab. 10. Wyniki symulacji dla optymalnej wartości liczności okna diagnostycznego wg kryterium K1.

Badane kryterium	Wynik bez T_{dop} *wynik*	Wynik poprawiony *wynpopr*	Maksymalne obsunięcie kapitału *obsm*	Stosunek obsunięcia do wyniku - K1
Wynik [pips x 10^4]	2.0973	2.0886	0.0615	0.0294



*Rys. 61. Wyniki badań wpływu liczności okna diagnostycznego na jakość kryteriów. Przedstawiono wynik (niebieski), wynik poprawiony (zielony) i obsunięcia (czerwony). Optymalna wielkość zbioru występuje dla i=3, co oznacza liczność 200+i*100=500 świec.*

3.6. Badanie wpływu wielkości prognostycznego okna czasowego – odległości do bariery wygaśnięcia pozycji

To trzecia (po TP i SL) podstawowa bariera przestrzeni zamknięcia otwartych pozycji. Badania, podobnie jak dotychczas, przeprowadzono dla najlepszych obliczonych parametrów, takich jak:

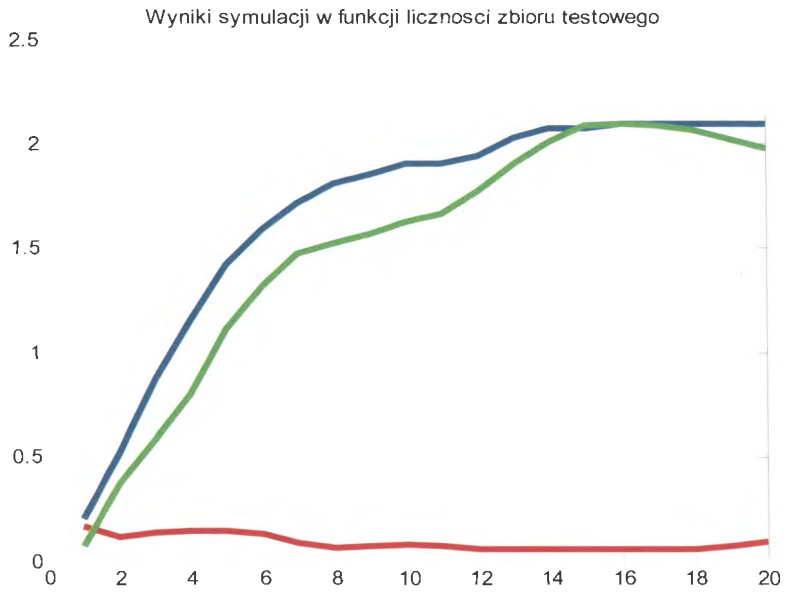
- liczba świec, od których inicjowano badanie okna o liczności S_d = 500 świec wynosiła Is=25000;
- granica skośności g_3 = 0.1;
- granica kurtozy g_4 =1.1;
- czas dopuszczalny otwarcia pozycji wyrażony liczba świec wprzód S_p =120;
- l_{tyl} = 50; - liczba świec wstecz, dla której budowano wskaźniki Lane'a;
- wskaźnik Lane'a – 50;
- rząd kwantyla zewnętrznego k_{lg2} = 0.99;
- rząd kwantyla wewnętrznego k_{lg1} = 0.66;
- bariera Stop Loss ustawiona została na poziomie SL=0.0120;
- bariera *Take Profit* ustawiona została na poziomie TP=0.0035.

Badania przeprowadzono dla liczby świec rozpatrywanych wprzód od chwili bieżącej sp=10 co 10 do 200 (1:20). Uzyskano poniżej przedstawione wyniki.

Tab. 11. Wyniki symulacji dla optymalnej wartości liczności okna testującego (prognostycznego) wg kryterium K1.

Badane kryterium	Wynik bez T_{dop} *wynik*	Wynik poprawiony *wynpopr*	Maksymalne obsunięcie kapitału *obsm*	Stosunek obsunięcia do wyniku - K1
Wynik [pips x 10^4]	2.0941	2.0958	0.0615	0.0293

Wyniki te są zbliżone do dotychczas uzyskiwanych. Jest to zrozumiałe z uwagi na płaski przebieg krzywych ilustrujących wyniki. Stosunkowo duża zmiana okna testującego z Sd=120 na Sd=160 zmienia dość w niewielkim zakresie obserwowaną odpowiedź modelu.



Rys. 62. Zmiany podstawowych wyników w zależności od liczby świec, po której następuje zamknięcie pozycji niezamkniętych przez bariery TP i SL. Tdop zmienia się od 10 do 200 świec. Maksimum występuje przy 16 cyklu obliczeń, czyli dla wartości Tdop=160. Wyniki to wynik (niebieski), wynpopr (zielony) i obsm – obsuniecie maksymalne (czerwony).

3.7. Wpływ dopuszczalnej liczby otwartych pozycji na efektywność i ryzyko transakcji

Po zbadaniu wrażliwości kolejno wybranych zmiennych, lub ich par zbadano znaczenie liczby otwieranych pozycji długich lub krótkich na skuteczność i ryzyko transakcji. Ryzyko do tej pory wyrażano poprzez wskaźnik maksymalnego obsunięcia. Zakładano także, że sukcesywnie otwierane były tylko pojedyncze pozycje w sytuacji, gdy spełnione były warunki otwarcia i zamykane były, gdy spełnione były warunki zamknięcia. Ponieważ zakładano kolejne sekwencyjne rozpatrywanie każdej kolejnej świecy mogło to doprowadzać do sytuacji wielokrotnego nakładania się w czasie otwartych równocześnie pozycji, i długich i krótkich, lub jednocześnie jednych i drugich. W sytuacji, gdy wyraźnie dominowała liczba pozycji tego samego typu (długich lub krótkich) stwarzało to zagrożenie dużej porażki przy niekorzystnej zmianie

rynku. Jednocześnie, przy korzystnym trendzie duża liczba poprawnie otwartych pozycji stwarzała szanse na duże zyski.

Na internetowych platformach brokerskich z reguły podaje się ten stan nierównowagi rynkowej wskazując różnicę pomiędzy np. sumą wartości otwartych długich pozycji i sumą wartości pozycji krótkich. Niekorzystny ruch rynku w kierunku zagrożonym (takim, w którym otwarta jest duża liczba przegrywających pozycji) zagrożenie to powiększa. Istnieją też platformy brokerskie, na których uniemożliwia się jednoczesne posiadanie odwrotnych pozycji. Tłumaczy się to najczęściej troską o bezpieczeństwo niedoświadczonego inwestora, jest jednak także zabezpieczenie brokera przed często uprawianym i dość skutecznym scalpingiem. Zasada ta (niedopuszczalności posiadania pozycji przeciwnych) jest przestrzegana także np. na rynku kontraktów terminowych na warszawskiej GPW. Na takich platformach pozwala się oczywiście na otwieranie, w jednym kierunku, wielu pozycji.

Niezależnie od stosowania tej zasady badanie zorientowane na optymalizację dopuszczalnej liczby otwartych pozycji ma więc sens.

Przeprowadzane tu rozważania dotyczące przestrzeni indykatorów mają luźny związek z zasadniczą ideą predykcyjnego wykorzystania GMDH. Związek ten polega mianowicie, po pierwsze, na wyjaśnieniu, jakie przetworzone zmienne (indykatory) mogą być użyte w roli zmiennych wejściowych, i po drugie, czy nie można dokonać kolejnych istotnych transformacji generujących nowe zmienne wejściowe dla metody.

Na podstawie dotychczas uzyskanego zbioru najlepszych otwarć w pewnej zdefiniowanej przestrzeni zmieniających się indykatorów rozpatruje się kolejne ograniczenie polegające na badaniu i zezwalaniu otwarcia pozycji tylko wtedy, gdy nie rośnie zagrożenie ponad założoną miarę. Ten kolejny indykator to zbiór najkorzystniejszych dopuszczalnych liczb otwarć długich i krótkich pozycji.

Sprawdzone i opisane w niniejszym rozdziale indykatory z ich najlepszymi wartościami to kolejno:

- liczba symulacji ls, dla których inicjowano badanie okna o najkorzystniejszej liczności $S_d = 500$ świec; liczba ta wynosiła ls=10;
- granica skośności z najlepszym wynikiem $g_3 = 0.1$;
- granica kurtozy z najlepszym wynikiem $g_4 = 1.1$;
- czas dopuszczalny otwarcia pozycji wyrażony liczba świec wprzód z najlepszym wynikiem $S_p = 160$;
- $l_{tyl} = 50$; - liczba świec wstecz, dla której budowano wskaźniki Lane'a;

- wskaźnik Lane'a z najlepszą wartością – 50;
- rząd kwantyla zewnętrznego z najlepszym wynikiem $k_{lg2} = 0.99$;
- rząd kwantyla wewnętrznego z najlepszym wynikiem $k_{lg1} = 0.66$;
- bariera Stop Loss z najlepszym wynikiem na poziomie SL=0.0120;
- bariera *Take Profit* z najlepszym wynikiem na poziomie TP=0.0035.

Oczywiście te najlepsze wartości indykatorów odnoszą się do całej wcześniej opisanej dziedziny zmienności czasu (indeksów świec) i mogą ulegać pewnym (niekiedy istotnym jak się okaże) zmianom w trakcie wykonywania długiej serii symulacji. Potwierdza to wielokrotnie wyrażane tu przekonanie autora, że jedyne co jest trwałe na rynkach finansowych to ciągłe zmiany. Przekonanie to nie jest zresztą szczególnie oryginalne.

W związku z tym dla przeprowadzenia badań sprowadzających się do określenia optymalnej, zmieniającej się dynamicznie hiperprzestrzeni indykatorów określono dodatkowo pewne dziedziny ich zmienności wokół średnio najlepszej, tej, wyżej wymienionej. Następnie przy wykonywaniu pewnej liczby symulacji w określonym miejscu macierzy danych szukano lokalnie najlepszych wartości indykatorów.

Algorytm ten oznaczony dalej jako A^1 będzie następnie skojarzony z innymi, A^2, A^3 i A^4 przedstawionymi w dalszej kolejności.

W tym rozpatrywanym algorytmie A^1 poszukiwania najlepszej strategii otwierania i zamykania pozycji w pewnej przestrzeni indykatorów stosowano następujące zasady przeszukiwania tej przestrzeni:

1. Na wstępie przyjęto, że najlepsze dla dalszych rozważań będą średnio najlepsze wartości indykatorów, czyli takie jak wymieniono powyżej;
2. Zauważono, że ważna jest kolejność, w jakiej dokonuje się korekt indykatorów dla rozpatrywanego odcinka badań (sekwencji świec) i po kilku wstępnych badaniach ustalono, że dość korzystna jest ta zaproponowana poniżej;
3. W pierwszej kolejności zmieniano rząd kwantyla zewnętrznego k_{lg2} poszukując najlepszej wartości w zbiorze $K_{lg2} = \{0.950, 0.955, 0.960, ..., 0.995, 1.000\}$. Z tego zbioru określano najlepszą w danym przedziale czasu wartość indykatora i przyjmowano ją jako przez chwile stałą w procedurze poszukiwania innych optymalnych wartości;

4. Następnym rozpatrywanym indykatorem był rząd kwantyla wewnętrznego k_{lg1}. Tu poszukiwano najlepszej wartości w zbiorze K_{lg1} = {0.48 0, 0.52, 056, 0,60 , ..., 0.80, 0.84}. Podobnie jak w poprzednim punkcie po znalezieniu tej wartości przyjmowano ja jako stałą w dalszych poszukiwaniach.
5. Kolejnym rozpatrywanym indykatorem był indykator Lane'a. Jego wartości wybierano ze zbioru Lane = {10, 20, 30, ..., 90}. Postępowano jak wyżej;
6. Następnie rozpatrywana była skośność wybierana ze zbioru G_3 = {-0.4, -0.2, 0, 0.2, ..., 1.6, 1.8};
7. Kolejnym indykatorem była kurtoza sprawdzana w zbiorze G_4 = {-0.8, -0.6, ..., 1.8, 2.0};
8. W dalszej kolejności sprawdzano liczbę świec w zbiorze testującym wprzód od chwili bieżącej Sp = {20, 30, 40, ..., 180};
9. Dwoma ostatnimi sprawdzanymi atrybutami były bariery zamknięcia. Barierę SL poszukiwano w zbiorze SL = {0.0010, 0.0020, ..., 0.0160};
10. Barierę TP rozpatrywano w zbiorze TP = {0.0005, 0.0010, 0.0015, ... , 0.0080};
11. Nie zmieniano wartości Sd – liczby świec testowanych wstecz od chwili bieżącej oraz l_{tyl} - liczby świec badanych wstecz dla ustalenia wartości indykatora Lane'a.

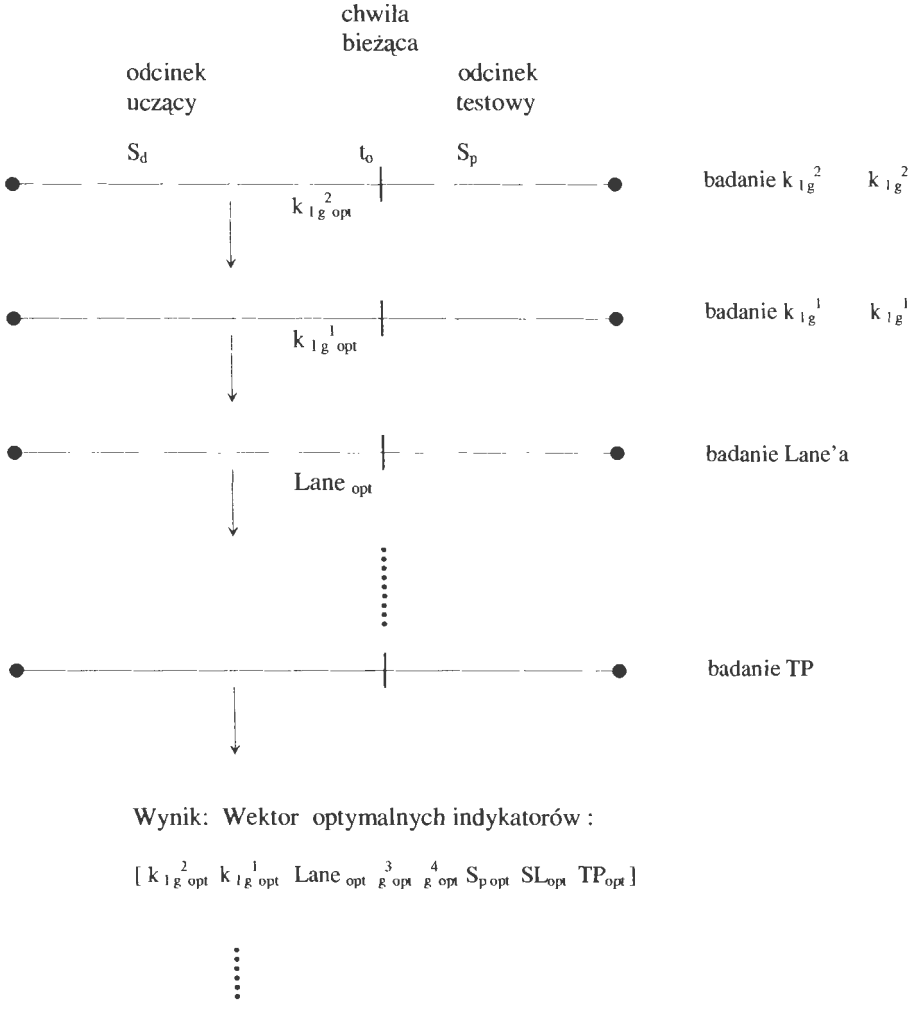
Schemat tego algorytmu wskazujący wielkość zbioru uczącego Sd świec poniżej wartości bieżącej, wielkość zbioru testującego Sp oraz sposób ekstrahowania najlepszych wartości indykatorów przedstawiono na rys. 63. Wyjściem z tego algorytmu jest wektor optymalnych wartości omówionych wyżej indykatorów, inny oczywiście dla każdej następnej chwili czasu (kolejnej świecy).

Rezultatem badań realizowanych wielokrotnie na tym samym fragmencie danych uczących Sd i testów wykonywanych w innym, umieszczonym w umownej przyszłości Sp jest wektor optymalnych indykatorów.

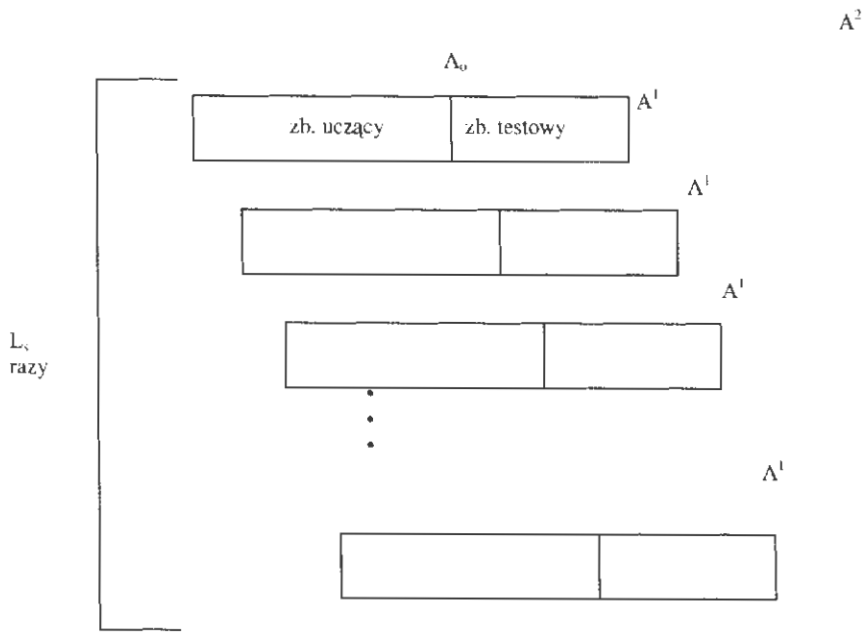
Powyższy algorytm był uruchamiany Ls razy. Liczba symulacji Ls wskazuje ile razy w przeszłości wykonywany był algorytm A^1 w sekwencji co jedną świecę. Schemat ten określony jako A^2 przedstawiony został na rys. 64. Każda sekwencja badań wg A^1 przesunięta jest względem poprzedniej o jedna świecę. W ten sposób szczelnie wypełniona jest cała przestrzeń dostępnych danych historycznych. Po każdorazowym wyszukaniu najlepszych wartości indykatorów w części uczącej

o długości Sd następowało sprawdzenie ich przydatności w testowej części przestrzeni. Przestrzegano tu zasady wykorzystywania jedynie danych, wcześniej, w części uczącej, nie wykorzystywanych.

W każdej części testowej algorytmu A^1 sprawdzano, czy kolejno rozpatrywane pozycje otwierane i zamykane zgodnie z wartościami indykatorów przynoszą zysk, czy stratę. Sumę kolejno tak wykonanych Ls algorytmów uznawano za wynik algorytmu A^2.



Rys. 63. Schemat graficzny objaśniający wykorzystanie algorytmu A^1.



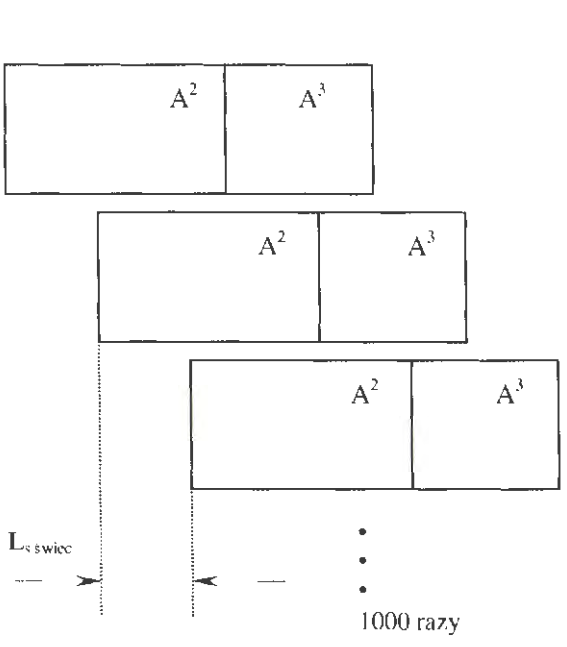
Rys. 64. Schemat graficzny algorytmu A^2. Składa się on z kolejno wykonywanych, przesuniętych o jedna świecę algorytmów A^1.

Na wykresach przedstawionych dalej każdy wynik algorytmu A^2 przedstawiany jest jako jeden punkt spośród 1000 przeprowadzonych symulacji. Ogólny obraz wszystkich realizowanych symulacji uogólnia schemat przedstawiony na rys. 65.

Schemat ten łączący algorytmy A^1 i A^2 nazwano tu algorytmem A^4. Na schemacie tym zaznaczono także algorytm A^3 jako wykonywany po każdorazowym wykonaniu algorytmu A^2. Istotą A^3 jest wykorzystanie najlepszych określonych w A^1 indykatorów i ustalenie, przy jakich ograniczeniach liczby otwartych pozycji długich i krótkich oraz przy jakim poziomie dopuszczalnej nierównowagi pomiędzy otwartymi pozycjami uzyskuje się najlepszy wynik.

Pierwsze symulacje przeprowadzono na tradycyjnie wykorzystywanym tu zbiorze 25000 świec 15 minutowych pary walutowej EUR/USD. Celem było sprawdzenie, przy jakich wartościach ograniczeń liczby otwieranych pozycji uzyskuje się najlepsze wyniki transakcyjne. Pozycje długie i krótkie otwierane były tylko wtedy, gdy spełnione były warunki wynikające z przestrzeni indykatorów. Najlepsze dopuszczalne liczby jednocześnie otwartych długich i krótkich pozycji określane były poprzez badanie w dodatkowej pętli z zadanymi wartościami tych liczb.

Nie była to wiec typowa symulacja predykcyjna a raczej weryfikacja znaczenia tego czynnika (dopuszczalnej liczby otwartych pozycji) dla jakości wyniku końcowego. Liczby te okazały się być czynnikiem niezwykle istotnym.

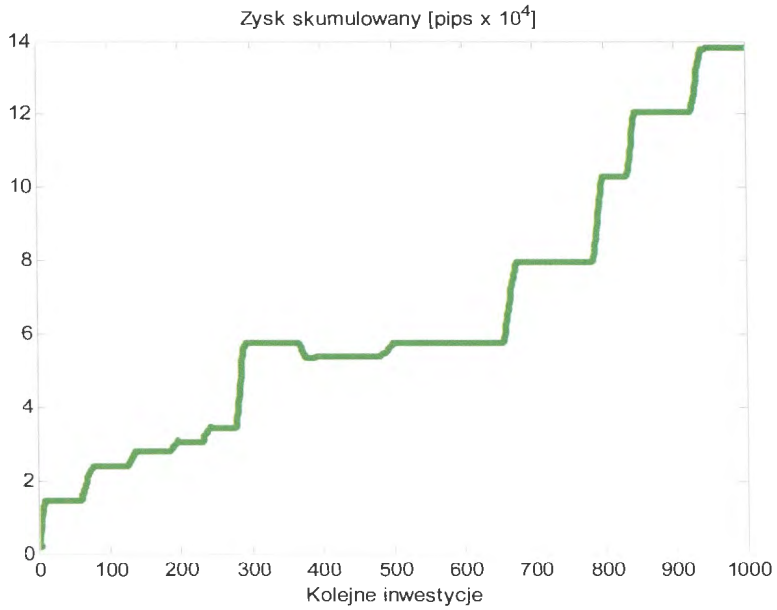


Rys. 65. Schemat organizacji obliczeń symulacyjnych. Wskazany inkrement Ls świec powoduje pełne wykorzystanie wszystkich danych z przeszłości.

Dla zadanych liczności zbiorów – uczącego Sd=500, testowego Sp=160, liczby symulacji wykonywanych w A^2 – Ls=10 uzyskano następujące wyniki (rys. 66).

Kapitał uzyskany po upływie 10000 rozpatrywanych świec (w seriach po 10 świec wg algorytmu A^1) – wynosił 138135 pipsów. To średnio na jedną pozycję ok. 13 pipsów zysku. To wynik znakomity przy dość dużej stabilności opisanej poniżej, uzyskany jednak nie dla prognozowanych danych, lecz dla najlepszych dopuszczalnych liczb otwartych pozycji.

Maksymalne obsunięcie w trakcie wykonania całego algorytmu A^4 – 369 pipsów, a odchylenie standardowe wynosiło ok. 42 pipsy (przy ok. 13 pipsach średniego zysku). Wykres narastania tego kapitału przedstawiono na rys. 66. Współczynnik Sharpe'a dla tego wykresu wyniósł SR = 0.3258.

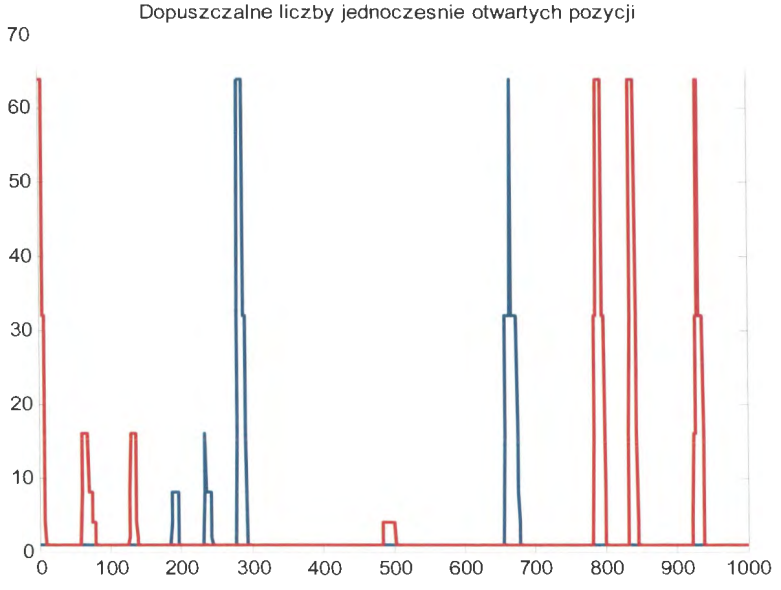


Rys. 66. Wykres narastania kapitału w trakcie wykonywania algorytmu A4 złożonego z 1000 realizacji A2 po 10 świec z jednoczesnym badaniem optymalnej liczby otwieranych pozycji.

W rozpatrywanej przestrzeni poszukiwano najlepszych wartości dla ograniczeń w postaci liczby dopuszczalnych jednocześnie otwartych pozycji długich i krótkich. Wyniki tych badań przedstawiono na wykresie 67. Dopuszczano tu otwarcie kolejnej pozycji nawet w sytuacji, gdy aktualny wynik na wszystkich otwartych pozycjach był ujemny, ale nie mniejszy niż minus 450 pipsów (to najlepszy wynik dla kilku prób także z innymi dopuszczalnymi ujemnymi wartościami bieżącego wyniku). Liczby otwartych pozycji rozpatrywano jako elementy pewnego zbioru stanowiące kolejne potęgi szeregu 2^n, n=0,1,2, ...

Na rys. 67 widać stosunkowo niewiele zdarzeń pozwalających na zwiększenie liczby otwartych pozycji powyżej jednej, ale mają one jednak, jak wspomniano, istotne znaczenie dla wyniku końcowego. W tych bowiem miejscach szeregu czasowego następowały istotne przyrosty zysku skumulowanego (rys. 66). Takich miejsc w długim szeregu czasowym (w rozpatrywanym jego fragmencie) jest ok. 10. Są to fragmenty szeregu, w których celowość jednoczesnego otwarcia wielu pozycji występuje kolejno w kilku tuż obok siebie położonych chwilach czasowych (początkach rozpatrywanych okresów 15 minutowych). Nasuwa to po-

mysł zastosowania prostej strategii, polegającej na wykorzystaniu w kolejnym kroku dopuszczalnej liczby otwartych pozycji takiej, jaka tuż przed tym krokiem była ustalona jako optymalna. W ten sposób zakłada się niejako poślizg z wykorzystaniem optymalnej liczby otwartych pozycji o jeden krok w szeregu czasowym.



Rys. 67. Liczby otwartych jednocześnie długich (niebieskie) i krótkich (czerwone) pozycji.

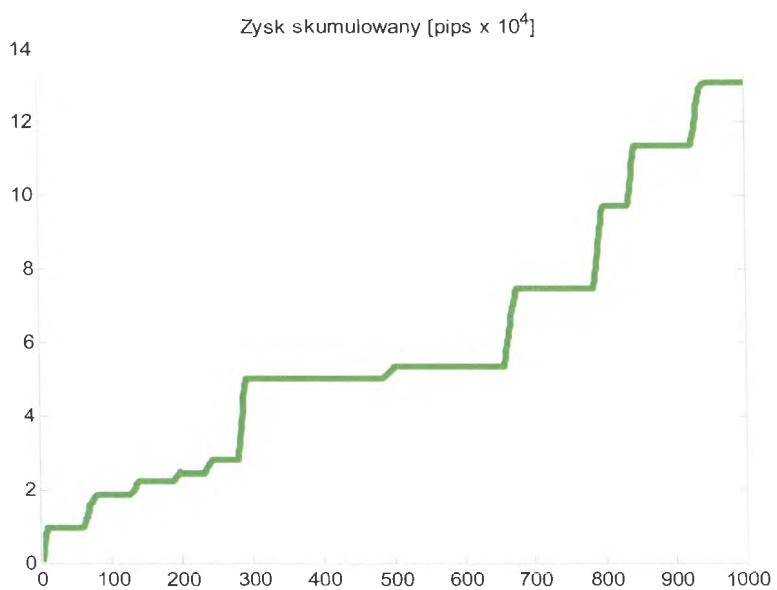
Wykorzystując wyniki wyżej przedstawionych badań przeprowadzono kolejną symulację w tym samej przestrzeni danych, jednak bez najmniejszej ingerencji w przyszłość. W każdym kroku symulacji wykorzystywano wyłącznie omówione wcześniej indykatory i po upływie okresu testowego, w którym poszukiwano ich najlepszych wartości dokonywano typowej predykcji (otwierano pozycje) w niezbadanej przyszłości.

Wyniki tej symulacji przedstawiono na kolejnych rysunkach – rys. 68 i 69. Wyniki te są bardzo podobne do tych zaprezentowanych wcześniej, osiągniętych w sytuacji kolejnego testowania różnych dopuszczalnych liczb otwartych pozycji i wyboru najlepszej z tych liczb. Wynik osiągnięty po ograniczeniu liczby otwartych pozycji do wartości optymalnej uzyskanej w kroku poprzedzającym decyzję wyniósł w ba-

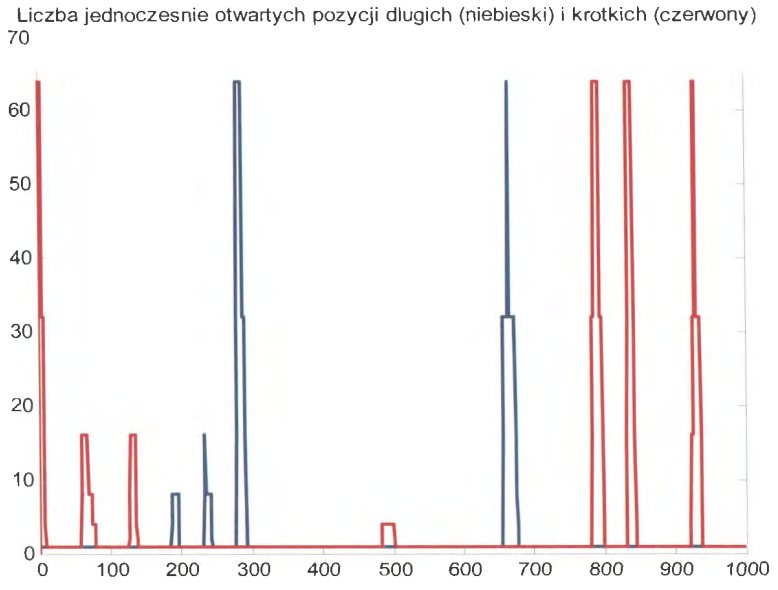
danym horyzoncie czasowym 130617 pipsów, a więc praktycznie tyle samo, co poprzednio. Największe obsuniecie kapitału wyniosło 153 pipsy, a średni zysk wyniósł także ok. 13 pipsów. Współczynnik Sharpe'a dla tego wykresu także był bardzo zbliżony do poprzedniego i wyniósł SR = 0.3252.

Wyniki przedstawione na wykresie 69 przedstawiające liczby jednocześnie otwartych pozycji długich i krótkich są w przybliżeniu takie same jak na rys. 67.

To pośrednie potwierdzenie słuszności hipotezy, że niewielkie, jednookresowe przesuniecie liczby otwartych pozycji nie pogarsza istotnie wyniku przedstawionego na rys. 66. Podsumowując ten fragment badań można powiedzieć, że stabilność wyników (mierzonych np. za pomocą współczynnika Sharpe'a) to bardzo interesująca właściwość określonej wcześniej przestrzeni czynników. Oczywiście należy pamiętać, że istotnym elementem rozpatrywanego podejścia jest zastosowanie metody GMDH jako sposobu tworzenia modeli z wyekstrahowaną z aktualnej rzeczywistości strukturą wielomianową.



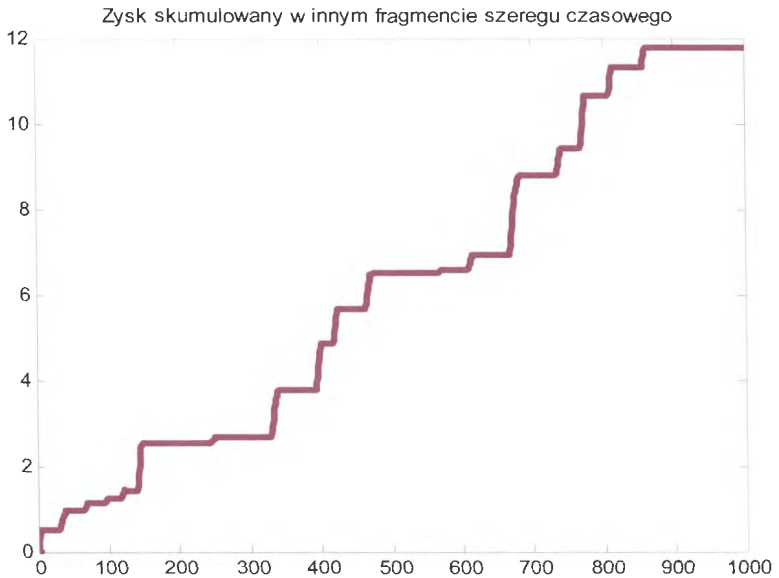
Rys. 68. Wykres narastania kapitału dla algorytmu A^4 wykonanego przy zastosowaniu ograniczenia liczby jednocześnie otwartych pozycji obliczonej jako najlepszej w poprzednim kroku.



Rys. 69. Wyniki uzyskane w czasie badania algorytmu A^4 z wykresu 68. Na wykresie przedstawione są najlepsze obliczone w czasie symulacji liczby jednocześnie otwartych pozycji długich (niebieskie) i krótkich (czerwone).

Dla sprawdzenia, czy zaobserwowana stabilność wyników pomiędzy symulacją wykonaną w warunkach poszukiwania najlepszych liczb otwartych pozycji a symulacją wykonaną w warunkach zastosowania liczby otwartych pozycji najlepszej w chwili poprzedzającej aktualną decyzję inwestycyjną, wykonano badania w zupełnie odmiennej części szeregu czasowego.

W celu próbkowano szereg czasowy nie co 10 świec (jak do tej pory), lecz co 20 a liczbę Ls rozpatrywanych symulacji (co 20 świec) zmieniono z dziesięciu na dwie. To istotne zmiany w parametrach symulacji obejmującej teraz zupełnie inny fragment szeregu czasowego. Podobnie jak poprzednio wykonano dwie niezależne symulacje. Najpierw obliczono, jakie liczby jednocześnie otwartych pozycji byłyby najlepsze w każdym rozpatrywanym kroku symulacji, liczbę te zapamiętano i następnie przeprowadzono eksperyment w warunkach pełnej predykcji stosując ograniczenie liczby otwartych pozycji takie, jakie było najlepsze w poprzednim (a wiec znanym) kroku. Wyniki tej drugiej, predykcyjnej symulacji przedstawiono na rys. 70.



Rys. 70. Narastanie kapitału w modelu predykcyjnym zbudowanym na wynikach symulacji określających najlepsze liczby jednoczenie otwartych pozycji w nierozpatrywanym do tej pory fragmencie szeregu czasowego.

Uzyskano tu wynik całkowity 117764 pipsy, a więc zbliżony do poprzednich. Odnotowano niewielkie obsuniecie maksymalne wynoszące 165 pipsów (niemal niezauważalne na wykresie) i średni zysk rzędu 12 pipsów przy odchyleniu standardowym ok. 40 pipsów. Współczynnik Sharpe'a dla wykresu z rys. 70 wynosił 0.2980.

Wyniki te, pomimo przeniesienia w zupełnie inną rzeczywistość, są zaskakująco zbieżne z dotychczasowymi. Oznaczać to może poprawność (lepszym określeniem może będzie efektywność) wybranej przestrzeni czynnikowej. Zawiera ona szereg oryginalnych zmiennych – wektor kwantyli, zmodyfikowany indykator Lane'a, także czynniki o charakterze technologicznym (np. SL, TP), związane z zarządzaniem ryzykiem (liczba otwartych pozycji, dopuszczalna wielkość asymetrii w inwestowaniu), czy wskaźniki uwzględniające prawdopodobieństwo zdarzeń (dezaprobata dla bezkrytycznego przyjmowania rozkładu normalnego). Przyznać się jednak należy, że rozpatrywana przestrzeń nie powstała przypadkowo, jest rezultatem wielu wcześniejszych niezbyt

udanych prób, w tym także z wykorzystaniem wielu znanych indykatorów uznawanych powszechnie za wartościowe i skuteczne.

Oczywiście wielki jest zdaniem autora wkład metody GMDH w pozyskanie najlepszej struktury modelu opisującego rzeczywistość wyrażoną jako szereg czasowy punktów w tej przetransformowanej przestrzeni. Metody, która pozwala na samoorganizujące się poszukiwanie najlepszej struktury (złożoności modelu) i identyfikację parametrów modelu (współczynników wielomianu). Warto także kolejny raz zwrócić uwagę, że rozpatrywana przestrzeń nie jest niczym stałym raz na zawsze określonym. Największą jej wadą jest być może nadmierna „ostrożność" w rekomendowaniu decyzji inwestycyjnej. Z wykresów na rys. 66 -70 wynika jak niewiele decyzji wystąpiło w dość długim okresie. Na wykresów narastania kapitału widać długie odcinki bezruchu, całkowitego braku sygnałów do otwarcia pozycji. Z drugiej strony, gdy system obserwujący zmiany na rynku dostrzegł sprzyjającą sytuację, to z reguły skutkowało to serią udanych decyzji.

Jak wiele tego typu systemów, związany jest on z charakterem człowieka – jego twórcy, jego emocjami, odpornością psychiczną, stylem percepcji otaczającego świata. Ten, wyżej przedstawiony, odzwierciedla zapewne oczekiwania autora.

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 opisania przyszłości i wymagają ustawicznej adaptacji, zarówno poprzez zmianę parametrów modelu (tu wielomianowego) jak tez 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 zajałby 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, sześcienne 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 osiągniecie wyższego niż niższego zysku. Zawsze lepiej jest osiągnąć założony zysk przy niższym niż wyższym poziomie ryzyka. Jednakże jak zmierzyć te 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 Transacions 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. DeBondt 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 Finacial 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. *Matematiceskije metody postroenija prognozov*. Radio i sviaz. Moskva 1997.

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 forcasting 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 Gruppowogo 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 Clusterization-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., *Samoobujascijesia systemy raspoznavanija i avtomaticeskogo 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., Zaicenko J.P., Dimitrov V.D. *Priniatije resenij na osnove samoorganizacji*. 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., Zadrożny 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., Zadrożny 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., Zadrożny S. *Fuzzy linguistic data summaries as a human consistent, user adaptable solution to data mining.* In: Gabrys B., Leiviskä K., Strackeljan 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 wspomagania 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 Volidation and Active Learning*. NIPS, 7 of 1995.

66. Krutsinger J. *Systemy transakcyjne.Sekrety mistrzów*. Wig-Press, Warszawa 1999.

67. Kulikowski R. *Wspomaganie inwestycji obarczonych ryzykiem*. W: Krawczak M., Jakubowski A., Konieczny P., Kulikowski R., Miklewski A., Szkatuła

G.: *Aktywne zarządzanie inwestycjami finansowymi – wybrane zagadnienia*, Akademicka Oficyna Wydawnicza EXIT, Warszawa 2003, ss. 253-268.

68. Kuś J., Korbicz J. *Statyczne i dynamiczne sieci GMDH* w Sieci neuronowe (eds. Duch, Korbicz, Rutkowski, Tadeusiewicz) w serii pod red. M. 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 Commputational 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 Evaluattion 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 Extarct Knowledge from Data.* Berlin, 1999.

81. Muriel A., *Short-term predictions in forex trading*, Physica A 344 (2004) 190–193

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. 8th 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, Willey, New York 2004.

90. Pedrycz W. *Computational Intelligence: An Introduction*. CRC Press 1997.

91. Penrose R. *Droga do rzeczywistości*. Proszyński i Ska. Warszawa 2006.

92. Piotrowski E. *Dwoistości wartości kapitału*. Politechnika Białostocka 2002.

93. Piotrowski E., Sładkowski 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.

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.* Departament 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 Genectic Program Model.* Springer Berlin/Heidelberg vol.3490/2005, 205-215.

121. Wagner N., Michalewicz Z., Khouja M., McGregor 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 Particiipation 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. Zholnarskij, 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 [dostęp 08/2008]

138. www.oanda.com [dostęp 05/2007 – 09/2008]

139. www.investopedia.com [dostęp 02/2005-10/2008]

140. www.wanadoo.com [dostęp 02/2006-11/2008]

141. www.gmdh.com [dostęp 06/2007-11/2008]

142. 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