

**POLITECHNIKA OPOLSKA**  
**WYDZIAŁ ELEKTROTECHNIKI, AUTOMATYKI I INFORMATYKI**  
**INSTYTUT INFORMATYKI**



**POLITECHNIKA**  
**OPOLSKA**

**PRACA DOKTORSKA**

**BADANIA PORÓWNAWCZE MODELI  
STATYSTYCZNYCH I OBLICZEŃ MIĘKKICH DLA  
IDENTYFIKACJI CIĄGÓW CZASOWYCH I PROGNOZ**

**MGR NATALIA KASHPRUK**

**PROMOTOR:**

**DR HAB. INŻ. RAFAŁ STANISŁAWSKI, PROF. PO**

**PROMOTOR POMOCNICZY:**

**DR HAB. INŻ. MAREK RYDEL, PROF. PO**

**Opole 2020**

**OPOLE UNIVERSITY OF TECHNOLOGY  
DEPARTMENT OF ELECTRICAL, CONTROL AND COMPUTER ENGINEERING**



**POLITECHNIKA  
OPOLSKA**

**PH.D. DISSERTATION**

**COMPARATIVE RESEARCH OF STATISTICAL MODELS  
AND SOFT COMPUTING FOR IDENTIFICATION OF  
TIME SERIES AND FORECASTING**

**NATALIIA KASHPRUK, M.SC.**

**ADVISORS:**

**RAFAŁ STANISŁAWSKI, PH.D.,D.SC.**

**MAREK RYDEL, PH.D.,D.SC.**

Opole 2020

## *Pracę dedykuję mojej rodzinie*

*Autorka składa serdecznie podziękowania  
Pierwszemu Promotorowi pracy doktorskiej  
Pani Profesor Annie Walaszek-Babiszewskiej oraz  
Promotorom niniejszej pracy doktorskiej  
Profesorowi Rafałowi Stanisławskiemu  
i Profesorowi Marku Rydlu  
za cierpliwość, cenne wskazówki, poświęcony mi  
czas oraz wsparcie.*

*Chciałabym również podziękować Tym Wszystkim  
moim Bliskim, którzy pomogli mi w ukończeniu  
mojej pracy naukowej, za pomoc, dobre rady,  
otuchę i nieustanną motywację.*

## STRESZCZENIE

Podjęty temat rozprawy doktorskiej dotyczy badań porównawczych dwóch różnych podejść do budowy i identyfikacji modeli ciągów czasowych: klasycznego podejścia uwzględniającego modele statystyczne AR, ARMA, ARIMA, TAR, STAR, ARX oraz podejścia wykorzystującego tak zwane obliczenia miękkie (*ang. soft computing*), jak modele rozmyte i neuronowo-rozmyte.

Cel pracy ma charakter poznawczy i polega na rozpoznaniu zalet, podobieństw, równoważności i różnic przy analizie i prognozie ciągów czasowych z zastosowaniem modeli statystycznych i obliczeń miękkich.

Badania poprzedzono krótkim przeglądem podejść do modelowania systemów dynamicznych, znanych w literaturze.

Analiza teoretyczna w niniejszej pracy obejmuje podstawy budowy modeli statystycznych ciągów czasowych jako pewnych klas procesów stochastycznych z czasem dyskretnym i ich identyfikację z jednej strony oraz modele rozmyte Mamdaniego, modele Takagi-Sugeno Kanga i model neuronowo-rozmyty Janga wraz z procedurami wnioskowania, z drugiej strony.

W wyniku analizy teoretycznej stwierdzono wiele analogii w procedurach tworzenia modeli w obydwu podejściach.

Dla przeprowadzenia przykładowych obliczeń posłużyły następujące zbiory danych: notowania kursów walut, obserwacje zmienności procesów technologicznych, notowania poboru energii elektrycznej. Identyfikacja modeli statystycznych została przeprowadzona w myśl metody najmniejszych kwadratów a dokładność prognoz weryfikowano nie tylko za pomocą błędów MSE i RMSE ale także za pomocą unormowanej funkcji autokorelacji reszt modelu. Wyznaczanie prognoz za pomocą modeli neuronowo-rozmytych realizowano w systemie obliczeniowym ANFIS MATLAB. Przedstawiono obliczenia prognoz i błędy prognoz dla różnych struktur modeli. Zrobiono prognozę danym poboru energii elektrycznej z liczników za pomocą Facebook Prophet.

Stwierdzono, że chociaż budowa modeli statystycznych i modeli rozmytych przebiega analogicznie, to techniki hybrydowe, np. neuronowo-rozmyte pozwalają na znacznie bardziej komfortowe sposoby uzyskiwanie wyników obliczeń. System neuronowo-rozmyty ANFIS, który jest oparty o technikę adaptacyjnej sieci neuronowej wspomagającej regułowy model rozmyty TSK, jest systemem obliczeniowym wyższej generacji w stosunku do modeli statystycznych i można go zaliczyć do narzędzi inteligencji obliczeniowej. Model Prophet jest najłatwiejszy w użyciu, jest najlepszym rozwiązaniem, gdy priorytetem jest szybkie uzyskanie wyników i nie zależy na dokładności.

Cel pracy został zrealizowany a tezy zostały wykazane.

## ABSTRACT

The subject of the PhD dissertation concerns comparative studies of two different approaches to the construction and identification of time series models: the classical approach taking into account the statistical models AR, ARMA, ARIMA, TAR, STAR, ARX and the approach using so-called soft computing, such as fuzzy and neuro-fuzzy models.

The aim of the work is cognitive and consists of recognizing the advantages, similarities, equivalences and differences in the analysis and prognosis of time series using statistical models and soft calculations.

The research was preceded by a short review of approaches to modeling dynamic systems known in the literature.

The theoretical analysis in this work covers the basics of building statistical models of time sequences as certain classes of discrete time stochastic processes and their identification on one side, and Mamdani's fuzzy models, Takagi-Sugeno-Kang's models and Yang's neuro-fuzzy model along with inference procedures, on the other side.

As a result of the theoretical analysis, many analogies were found in the procedures for creating models in both approaches.

The following data sets were used to carry out sample calculations: quotations of main market indices, exchange rates quotations, observations of variability of technological processes. Identification of statistical models was carried out in accordance with the least squares method and the accuracy of forecasts was verified not only by means of MSE and RMSE errors but also by the normalized function of the autocorrelation of residuals of the model. The determination of forecasts using fuzzy neural models was carried out in the ANFIS MATLAB computational system. Calculations of forecasts and forecast errors for various model structures are presented. A forecast was made of a given electricity consumption from meters using Facebook Prophet.

It was found that although the construction of statistical models and fuzzy models is analogous, hybrid techniques, e.g. neuro-fuzzy techniques, allow for much more expedient ways of obtaining calculation results. The ANFIS neuro-fuzzy system, which is based on the technique of adaptive neural network supporting the TSK fuzzy rule model, is a higher-generation computational system compared to statistical models and can be counted as computational intelligence tools. Prophet is the easiest to use, it is the best solution when getting results quickly is a priority and accuracy is not a concern.

The aim of the work was accomplished and theses were demonstrated.

## SPIS TREŚCI

SPIS OZNACZEŃ I SKRÓTÓW .....	6
1. SFORMUŁOWANIE PROBLEMU BADAWCZEGO .....	7
1.1. Wstęp .....	7
1.2. Cel i teza pracy .....	9
1.3. Zakres pracy .....	10
2. TEMATYKA BADAŃ W LITERATURZE PRZEDMIOTU .....	11
3. MODELE STATYSTYCZNE CIĄGÓW CZASOWYCH .....	16
3.1. Wprowadzenie .....	16
3.2. Procesy stochastyczne - podstawowe definicje i własności .....	16
3.3. Modele autoregresji .....	18
4. WYBRANE METODY IDENTYFIKACJI .....	26
4.1. Metoda najmniejszych kwadratów dla modeli parametrycznych .....	26
4.2. Praktyczne metody doboru modeli .....	30
4.3. Funkcja autokorelacji reszt modelu .....	33
4.4. Składowe trendu .....	34
4.5. Podejście Boxa-Jenkinsa do identyfikacji modeli na przykładzie modelu ARIMA .....	35
4.6. Dobór modeli TAR .....	36
4.7. Facebook Prophet .....	39
4.8. Identyfikacja modeli na przykładzie modelu ARIMA (podejście Boxa-Jenkinsa) .....	39
4.9. Opis danych stanowiących wybrane parametry węgla energetycznego .....	44
4.10. Model ARX i identyfikacja dla ciągu obserwacji procesu technologicznego .....	48
4.11. Prognoza przy stosowaniu modelu TAR .....	51
4.12. Przykład prognozy danych za pomocą pakietu Facebook Prophet .....	53
5. MODELOWANIE MIĘKKIE CIĄGÓW CZASOWYCH I SYSTEMÓW DYNAMICZNYCH DYSKRETNYCH .....	55
5.1. Systemy rozmyte typu Mamdaniego i typu Takagi-Sugeno-Kanga (TSK) .....	55
5.2. Rozmyte modele autoregresji .....	61
5.3. Adaptacyjny system neuronowo-rozmyty (ANFIS) .....	66
5.4. Przykład identyfikacji danych testowych .....	68
5.5. Identyfikacji ciągów czasowych za pomocą systemu neuronowo-rozmytego .....	72
6. PODSUMOWANIA I DALSZE PLANY BADAWCZE .....	86
6.1. Podsumowania metod stosowanych w pracy .....	86
6.2. Wkład własny .....	89
6.3. Kierunki przyszłych badań .....	90
SPIS LITERATURY .....	91

## SPIS OZNACZEŃ I SKRÓTÓW

$D^2$	- wariancja
$k_e(s)$	- funkcja autokorelacji reszt (błędów) modelu
ACF	- funkcja autokorelacji
ANFIS	- adaptacyjny system wnioskowania neuronowo-rozmytego
AR	- model autoregresji
AR(1)	- model autoregresji pierwszego rzędu
AR(2)	- model autoregresji drugiego rzędu
ARIMA	- zintegrowany(scałkowany) model autoregresji ze średnią ruchomą
ARMA	- model autoregresji ze średnią ruchomą
ARX	- model autoregresyjny z zewnętrznym wejściem
I	- całkowanie
MA	- model średniej ruchomej
MAD	- średni błąd bezwzględny
MAPE	- średnia wartość bezwzględnych wartości błędów względnych, w procentach
MSE	- wartość średnia sumy kwadratów błędów prognozy
RMSE	- pierwiastek kwadratowy z wartości średniej sumy kwadratów błędów prognozy
RMSPE	- pierwiastek z sumy kwadratów błędów względnych, wyrażony w procentach
STAR	- autoregresyjny model płynnego przejścia
TAR	- progowy model autoregresji
TSK	- model Takagi-Sugeno-Kanga
$E$	- wartość oczekiwaną
$cov$	- funkcja kowariancji

# 1. SFORMUŁOWANIE PROBLEMU BADAWCZEGO

## 1.1. Wstęp

W codziennych działaniach praktycznych, jak: analiza rynków finansowych, analiza procesów technologicznych, analiza procesów makroekonomicznych spotykamy zbiory danych uwarunkowanych czasowo, reprezentujące wyniki pomiarów lub obserwacji tych procesów. Przedmiotem analizy ciągów czasowych jest budowa modeli, badanie dynamiki oraz przewidywanie przyszłych wartości dla sterowania lub prognoz, czyli dla zadań podejmowania decyzji w różnych obszarach aktywności człowieka. W teorii sterowania procesów oraz w naukach ekonomicznych znane są różne podejścia do budowy modeli ciągów czasowych oraz różne narzędzia programistyczne wspomagające proces budowy modeli.

Modele statystyczne ciągów czasowych, do których zalicza się m. in. modele autoregresji (AR), modele średniej ruchomej (MA), modele ARMA, stanowiące kombinację obydwu typów, zintegrowane modele ARMA, czyli modele ARIMA, mają ugruntowaną pozycję w obszarach nauk ekonomicznych, rynków finansowych a także w teorii sterowania. Nieco później zdefiniowano modele bardziej zaawansowane, nieliniowe, do których należą progowe modele autoregresji (TAR), i modele STAR (Box i Jenkins, 1970), (Hamilton, 1994), (Brockwell i Davis, 1996), (Mantegna i Stanley, 2000), (Tsay, 2002), (Welfe, 2009).

Tak zwane obliczenia miękkie (*ang. soft computing*) (Zadeh, 1999), realizowane za pomocą systemów rozmytych, sieci neuronowych oraz innych narzędzi sztucznej inteligencji, również znalazły zastosowanie w wielu obszarach nauki i techniki. Prognozowanie, w oparciu o dane typy ciągów czasowych jest jednym z zastosowań obliczeń miękkich, które odnoszą rzeczywiste sukcesy. Wśród rozmytych modeli ciągów czasowych, które są stosowane w zadaniach sterowania lub prognoz najczęściej stosowanymi są modele typu Mamdaniego oraz modele Takagi-Sugeno-Kanga (Czogała i Pedrycz, 1985), (Hellendoorn i Driankov, 1997), (Pedrycz i Chen, 2013).

Do narzędzi inteligencji obliczeniowej, wspomagających analizę ciągów czasowych należą sztuczne sieci neuronowe, systemy neuronowo rozmyte, systemy ekspertowe i inne.

Wraz z rozwojem metod i technik sztucznej inteligencji, modele rozmyte i neuronowo rozmyte stały się obiektem zastosowań dla analizy i prognoz ciągów czasowych. W okresie kilku dziesięcioleci istnienia inteligencji obliczeniowej, w literaturze przedmiotu ciągle zbyt mało jest porównań obydwu podejść do budowy modeli, identyfikacji oraz efektów w postaci dogodności stosowania, dokładności modeli, nakładu obliczeń.

Modelowanie szeregów czasowych, zarówno z zastosowaniem klastycznych modeli statystycznych, jak również modeli w postaci rozmytych szeregów czasowych, może być istotnym elementem rozwiązań szeregu problemów naukowych i inżynierskich w dyscyplinach



Automatyka i Robotyka oraz Elektrotechnika, które obecnie stanowią wspólną dyscyplinę Automatyka, Elektronika i Elektrotechnika w dziedzinie nauk inżynieryjno-technicznych.

W obszarze Automatyki i Robotyki modele szeregów czasowych są powszechnie stosowane w układach regulacji automatycznej procesami stochastycznymi. Zastosowanie tych narzędzi możliwa jest predykcję zakłóceń występujących w procesach podlegających regulacji automatycznej, co w konsekwencji prowadzi do poprawy jej efektywności. Dla układów o stosunkowo niewielkim stopniu skomplikowania istnieje możliwość uwzględniania modeli szeregów czasowych *a-priori* w uniwersalnych aproksymatorach heurystycznych, stosowanych powszechnie w zadaniach identyfikacji dla celów sterowania, a w szczególności sterowania predykcyjnego. Do takich modeli można zaliczyć modele: ARMAX (ang. *AutoRegressive Moving Average with eXogenous input*) oraz Boxa-Jenkinsa (Soderstrom and Stoica, 1989). W tych systemach modele szeregów czasowych implementowane są w odrębnych torach „kolorujących” zakłócenia występujące w procesie. Ich zastosowanie może istotnie poprawiać jakość aproksymacji toru głównego, jak również istotnie wpływać na pracę projektowanych układów regulacji, w szczególności w układach sterowania predykcyjnego (Khan, 2020). Należy jednak podkreślić, że w przypadku bardziej złożonych procesów, np. opisywanych z uwzględnieniem złożoności struktury systemu, nie jest możliwe uwzględnienie modeli statystycznych bezpośrednio w zwartych równaniach opisujących proces. Wówczas modele szeregów czasowych są implementowane jako autonomiczne podsystemy wchodzące w skład systemu o złożonej strukturze. Innymi przykładami zastosowania modeli szeregów czasowych w obszarze teorii i praktyki sterowania są np.: modelowanie ewolucji systemów złożonych (Jiang *at al.*, 2016), problematyka zintegrowanych algorytmów sterowania statystycznego i automatycznego (ang. *integrated statistical and automatic process control*) (Tsung and Tsui, 2010), zagadnienia badania odpornej stabilności systemów dynamicznych i stabilności niestacjonarnych systemów stochastycznych (Yang and Wu, 2010), oraz wiele innych.

W obszarze Elektrotechniki modelowanie szeregów czasowych ma wymiar znacznie bardziej praktyczny. Szeregi czasowe są tutaj wykorzystywane w procesie modelowania i prognozowania różnorodnych procesów stochastycznych związanych z szeroko rozumianą elektrotechniką. Do przykładowych zastosowań można zaliczyć np.: prognozowanie zapotrzebowania na energię elektryczną (Dash *at al.*, 1995), (Acakpovi *at al.*, 2020), predykcję mocy generowanej przez odnawialne źródła energii (Silveira-Gontijo and Azevedo Costa, 2020), (Lukhyswara, 2019), przewidywanie przeciążeń transformatorów w sieciach dystrybucyjnych (Chen *at al.*, 2018), i wiele innych aplikacji.

Podsumowując należy podkreślić, że w Automatyce i Robotyce zagadnienia identyfikacji i predykcji szeregów czasowych są istotnym elementem stosowanym do rozwiązania wielu problemów, zarówno w ujęciu teoretycznym, jak i praktycznym. W szczególności są stosowane w takich obszarach jak: identyfikacja i modelowanie systemów dynamicznych (w tym systemów złożonych), analiza układów dynamicznych, oraz synteza układów regulacji/sterowania. Również w Elektrotechnice modelowanie szeregów czasowych jest narzędziem stosowanym do rozwiązania wielu problemów praktycznych. W związku z powyższym, pomimo stosowania m.in. narzędzi inteligencji obliczeniowej do identyfikacji i predykcji różnego rodzaju sygnałów, niniejsza praca istotnie oddziałuje na rozwój dyscypliny Automatyka, Elektronika i Elektrotechnika.

## 1.2. Cel i teza pracy

Podjęty temat pracy doktorskiej dotyczy badań porównawczych dwóch różnych podejść do identyfikacji ciągów czasowych:

- klasycznego podejścia uwzględniającego budowę modeli statystycznych oraz
- podejścia wykorzystującego tak zwane obliczenia miękkie (*ang. soft computing*).

W zadaniach złożonych, takich jak: predykcja, sterowanie czy diagnostyka, w których identyfikacja ciągów czasowych jest jednym z etapów zadania, pierwsze podejście bywa określane jako *model-based* problem. Drugie podejście, znane jest jako *heuristic-based designed*, lub gdy utworzony zostanie system z bazą wiedzy, mówi się *knowledge-based* problem (Hellendoorn i Driankov, 1997).

**Stąd przyjęto, że główny cel pracy doktorskiej ma charakter poznawczy i polega na:**

**rozpoznaniu zalet, podobieństw, równoważności i różnic w metodach analizy i prognozy ciągów czasowych z zastosowaniem modeli statystycznych i obliczeń miękkich.**

**Cel pracy jest zrealizowany poprzez:**

- analizę teoretyczną zasad budowy modeli ciągów czasowych z rodziny autoregresji, czyli modeli AR, MA, ARMA, ARIMA, TAR i STAR z jednej strony oraz modeli rozmytych typu Mamdaniego i Takagi-Sugeno-Kanga oraz neuronowo-rozmytych (ANFIS);
- wykonanie przykładowych obliczeń dla poszczególnych typów modeli;
- analizę porównawczą różnych typów modeli dla tych samych zbiorów danych.

**Przeprowadzone badania w powiązaniu z powyższym celem pracy pozwalają na sformułowanie następującej tezy:**

- 1. obliczenia miękkie i modele statystyczne budowane są w sposób analogiczny;**
- 2. obliczenia miękkie są techniką obliczeniową wyższej generacji w stosunku do modeli statystycznych w zastosowaniu do analizy i prognozy ciągów czasowych.**

### 1.3. Zakres pracy

Praca składa się z sześciu głównych rozdziałów. W rozdziale pierwszym sformułowano problem badawczy, w tym określono cel badań oraz postawiono tezę pracy.

W rozdziale drugim przedstawiono przegląd literatury przedmiotu.

W rozdziale trzecim zostały sformułowane podstawowe modele procesów stochastycznych z czasem dyskretnym, znane z literatury przedmiotu. Wskazano na własności tych modeli.

W rozdziale czwartym określono metodę identyfikacji modeli w oparciu o dane doświadczalne i kryteria identyfikacji. Przedstawiono schemat identyfikacji modeli Boxa-Jenkinsa. Opracowano metodę strojenia modeli TAR. Zwrócono uwagę na problem jakości dopasowania modeli do typu danych doświadczalnych oraz wyznaczono parametry doboru lepszego modelu dla prognozy. Ponadto przedstawiono przykłady identyfikacji dla różnych typów danych: modelowanie zmiany kursu dolara, modelowanie procesu technologicznego poboru próbek frakcji węgla. Przedstawiono modelowanie ekonomii w Facebook Prophet.

W rozdziale piątym przedstawiono główne problemy podejścia obliczeń miękkich do prognozy ciągów czasowych. Omówiono główne typy modeli rozmytych, w tym:

- modele o jednym wejściu i jednym wyjściu Mamdaniego i Assiliana oraz Takagi-Sugeno-Kanga,
- modele o dwóch wejściach i jednym wyjściu powyższych typów,
- model neuronowo-rozmyty o dwóch wejściach i jednym wyjściu.

W szczególności, przedstawiono modele AR(1) i AR(2) jako modele rozmyte typu Mamdaniego i typu Takagi-Sugeno-Kanga.

Wykorzystując system obliczeniowy Matlab ANFIS oraz dane empiryczne jak w rozdziale czwartym wykonano obliczenia prognoz i przedstawiono błędy prognoz dla różnych struktur modeli neuronowo-rozmytych.

W rozdziale szóstym przedstawiono wkład własny, który jest koncepcją pracy doktorskiej, przedstawia analizę porównawczą obydwu podejść do modelowania ciągów czasowych, w sensie metod budowy modeli z zastosowaniem modeli statystycznych i obliczeń miękkich. Ponadto zawiera podsumowanie i wnioski z przeprowadzonych badań oraz kierunki dalszych badań.

## 2. TEMATYKA BADAŃ W LITERATURZE PRZEDMIOTU

Modele ciągów czasowych, w literaturze przedmiotu określane także jako modele szeregów czasowych, zalicza się do statystycznych metod modelowania. Metody statystyczne w badaniach danych empirycznych dla prognoz rozwinęły się najpierw w obszarach ekonomii, finansów, handlu ale także w obszarach sterowania i planowania produkcji. W latach sześćdziesiątych i siedemdziesiątych dwudziestego wieku wielu autorów, wśród nich G.E.P. Box i G.M. Jenkins, opublikowali wiele prac z zakresu statystycznych aspektów układów dynamicznych, sterowania adaptacyjnego, filtracji, prognozy, modeli dyskretnych ciągów czasowych i ich własności. Usystematyzowanie tej tematyki znalazło się w słynnej monografii Boxa i Jenkinsa pt. *Time Series Analysis: Forecasting and Control* (1970). Podejście opisane przez zespół autorów: G.E.P. Box, G.M. Jenkins et al. (2015) jest silnie motywowane ciągami czasowymi w czasie rzeczywistym oraz jest opracowane kompletne podejście do budowania modelu, szacowania, prognozowania i kontroli.

Wśród modeli opracowanych przez Boxa-Jenkinsa szerokie zastosowanie znalazły modele autoregresji skończonego rzędu  $p$ , oznaczane jako AR( $p$ ). Zakłada się, że badany proces stochastyczny jest stacjonarny a prognozowanie jego przyszłych wartości jest kombinacją liniową przeszłych i obecnych zaobserwowanych wartości procesu. Model autoregresji zawiera addytywny proces losowy o nieskorelowanych wyrazach i skończonej wariancji.

Także dla stacjonarnych procesów stochastycznych opracowano model średniej ruchomej, MA( $q$ ). Prognoza procesu jest sumą ważoną skończonej liczby wyrazów nieskorelowanego procesu losowego (szumu białego). Jest to rozwinięcie koncepcji Yule'a, filtru liniowego o skończonej liczbie wyrazów. Kombinacją obydwu modeli jest model ARMA( $p,q$ ), który często jest stosowany w praktyce dla prognozy różnych zjawisk i procesów. Zastosowanie tego modelu również przedstawiono w pracach Minakhi Rout, et al, (2014), Terence C. Mills, (2019) oraz Yong, Lingyun and Jia, (2020).

Na szczególną uwagę zasługuje model ARIMA, zintegrowany (zsumowany) model ARMA, przeznaczony do modelowania niestacjonarnych procesów stochastycznych.

Według niektórych badaczy, mimo swych zalet i elastyczności, model ARIMA jest techniką złożoną; nie jest łatwy w użyciu i wymaga dużego doświadczenia (Bails i Peppers, 1982).

Pomimo szerokiej popularności i przydatności modeli Boxa-Jenkinsa w wielu obszarach zastosowań, kluczowym ograniczeniem tych modeli jest ich liniowy charakter. Jest to powodem nieefektywności tych modeli w przypadku opisywania nieliniowych szeregów czasowych.

W 1978 r. Tong zaproponował progowy model autoregresyjny (TAR). Sukces tego modelu w ekonometrii stworzył nową rodzinę modeli, *autoregresyjne modele przełączania reżimu*. Modele te oparto na idei podziału przestrzeni stanu na kilka podprzestrzeni, z których każda przestrzeń jest modelowana przez model AR. Kluczową własnością modeli TAR jest

nieciągły charakter relacji AR w momencie przekroczenia progu (Tong, 1978), (Chan, 1993), (Hansen, 1997).

Kilka lat później, w 1994 r. Timo Terasvirta zaproponował alternatywny model z *płynnym przejściem autoregresji* (STAR). Modele STAR zapewniają ciągłość funkcji przy przejściu z poszczególnych podprzestrzeni, tzn. od jednego liniowego modelu AR do drugiego (Terasvirta, 1994), (Terasvirta i Kock, 2010), (Terence C. Mills, 2019)

Innym podejściem matematycznym do opisu nieliniowości pewnych szczególnych systemów dynamicznych jest chaos deterministyczny. Analizę dynamiki systemów fizycznych, w których występuje ruch chaotyczny prezentował Henri Poincare już w roku 1892 a wiele lat później, w 1963r. meteorolog M.E. Lorenz odkrył chaos budując model długoterminowych zmian pogody. Stwierdził wrażliwość systemów chaotycznych na zmianę warunków początkowych („efekt motyla”). Geometria fraktalna, której twórcą był B. Mandelbrot (1982), dostarczyła narzędzi do detekcji chaosu w zbiorach ciągów czasowych generowanych przez systemy dynamiczne. Zastosowanie metod matematycznych geometrii fraktalnej w analizie ryzyka (Mandelbrot, 1999), (Mularczyk, 2008), detekcja chaosu dla badania inflacji oraz rynków finansowych (Chappell i Eldridge, 1997), analiza chaotycznych ciągów czasowych w ekonomii (Faggini, 2014) – to tylko niektóre z zastosowań teorii systemów nieliniowych i wykrywania zachowań chaotycznych w zbiorach danych.

Dla porządku należy wspomnieć o jeszcze innym podejściu do modelowania systemów dynamicznych, a mianowicie o modelach układów dynamicznych niecałkowitego rzędu. Teoria systemów bazująca na operatorach różniczkowania niecałkowitego rzędu, której początki zostały położone przez wybitnych matematyków XVIII. i XIX. wieku, obecnie przeżywa duże ożywienie i owocuje wieloma rozwiązaniami teoretycznymi, jak np. w pracy (Stanisławski, Rydel i Latawiec, 2017) oraz zastosowaniami, np. w ekonomii (Fallahgoul i in., 2016), (Liu, Chen i Zhang, 2017).

Zupełnie inne podejście do modelowania systemów powstawało w latach sześćdziesiątych i siedemdziesiątych XX. wieku, stworzono uogólniony aparat matematyczny oraz nieskończenie wymiarową logikę rozmytą. Pionierska praca L.A. Zadeha „Fuzzy Sets” (1965) zawierała definicje zbiorów rozmytych oraz podstawowych operacji logicznych. Kolejne prace tego autora definiowały prawdopodobieństwo zdarzeń rozmytych (1968), podstawy analizy systemów (1973), pojęcie zmiennej lingwistycznej której wartości lingwistyczne precyzują zbiory rozmyte (1975) i wreszcie zasady wnioskowania przybliżonego (1979).

Fundamentalne znaczenie dla budowy teorii systemów rozmytych miała praca E.H. Mamdaniego i S. Assiliana (1975), w której podano przykład sterowania obiektem przemysłowym w oparciu o reguły heurystyczne, wynikające z wiedzy eksperta. Sposób zachowania obiektu został zapisany w postaci zbioru zdań warunkowych *Jeżeli...To*, w których przesłanki i konkluzje są wartościami lingwistycznymi. Zbiór reguł nosi nazwę bazy wiedzy a realizacja sterowania, czyli wnioskowanie odbywa się na podstawie aktualnej przesłanki i bazy reguł za pomocą operatorów *max* i *min*.

Odmienny sposób realizacji systemu wnioskującego z bazą wiedzy zaprezentowany został w pracach zespołu autorskiego: M. Sugeno, G.T. Kang i T. Takagi (1985, 1988). Zbiór reguł tworzących bazę wiedzy ma przesłanki zapisane w postaci lingwistycznej a w

następnikach reguł jest wyrażenie analityczne, funkcja o wartościach numerycznych zmiennych wejściowych i wyjściowej. Wnioskowanie, dla nowych przesłanek zawiera kombinację wyznaczania logicznej wartości przesłanki i wartości funkcji zawartej w następniku.

Systemy modelowania zbudowane i funkcjonujące na tych zasadach noszą nazwę rozmytych systemów ekspertowych Mamdaniego lub Takagi-Sugeno-Kanga. Rozmyty system Takagi-Sugeno-Kanga jest jednym z najczęściej stosowanych modeli w sztucznej inteligencji. Tworzenie systemów TSK składa się z dwóch głównych zadań: identyfikacji konstrukcji i oszacowania parametrów. Identyfikacja struktury wiąże się głównie z określeniem liczby reguł rozmytych. Estymacja parametrów wiąże się z wyznaczeniem parametrów zbiorów rozmytych oraz współczynników funkcji regresji w części następnika. Zadania te można osiągnąć różnymi technikami optymalizacyjnymi, przedstawionymi w pracach: Taieb, A., Soltani, M., Chaari, A., (2018), Tsai, SH, Chen, YW, (2018), Wiktorowicz i Krzeszowski, (2020). W artykule Li, J., Yang, L., Qu, Y. i in., (2018) zaproponowano rozmyty system wnioskowania TSK +, który rozszerzył konwencjonalny system wnioskowania TSK w taki sposób, że można go również zastosować do rzadkich baz reguł i nierównomiernie rozłożonych baz reguł.

Zagadnienia sterowania rozmytego były podejmowane przez wielu autorów, m.in. przez R.M. Tonga (1978, 1979), D. Dubois i H. Prade (1980), a w latach późniejszych przez J. Kacprzyka (1986) oraz R. Yagera D.P. Fileva (1994). R. Yager jest także autorem definicji rozmytego prawdopodobieństwa.

Teoria systemów rozmytych rozwija się już od ponad 50 lat i międzyczasie wzbogacona została o inne podejścia połączone z nią synergicznie, jak np. systemy neuronowo-rozmyte (Rutkowski, 2005), (Łęski, 2008), systemy probabilistyczno-rozmyte (Walaszek-Babiszewska, 2011), (Walaszek-Babiszewska i Rudnik, 2013), systemy rozmyte wspomagane obliczeniami genetycznymi (Kacprzyk i Yager, 2001) lub systemy rozmyte wspomagane przez *clustering* (Hellendorn i Driankov, 1997). Taki rodzaj współistnienia różnych podejść został określony przez L.A. Zadeha jako *soft computing*, a w języku polskim przyjęła się nazwa *obliczenia miękkie*. Powstało także określenie *computational intelligence*.

Zastosowanie systemów rozmytych i hybrydowych do analizy ciągów czasowych jest jednym z ważniejszych tematów wielu konferencji i publikacji zbiorowych, jak np. praca *Time Series Analysis, Modeling and Applications; A Computational Intelligence Perspective*, pod redakcją W. Pedrycza i Shyi-Ming Chena (2013).

Oprócz rozwoju nowych teorii i możliwości ich zastosowań, nie mniej ważne są systemy obliczeniowe umożliwiające zastosowanie wymienionych podejść do modelowania i analizy systemów rzeczywistych. Takim systemem obliczeniowym stał się, m.in. adaptacyjny neuronowo-rozmyty system wnioskujący ANFIS MATLAB, odpowiedni do obliczeń miękkich, czyli do modelowania złożonych, niepewnych i nieprecyzyjnie określonych procesów.

System ANFIS, wykorzystujący możliwości uczenia adaptacyjnych sieci neuronowych, został opracowany na początku lat pięćdziesiątych XX. wieku przez J.S.R. Janga. Funkcjonalnie, system jest równoważny systemowi opartemu o reguły i wnioskowanie typu Takagi-Sugeno-Kanga. W pracy Janga, (1993) znajdujemy dowód równoważności wielowarstwowej sieci neuronowej i systemu rozmytego TSK. Ponieważ system ANFIS był jednym z pierwszych systemów neuronowo rozmytych, stał się standardem w obliczeniach

miękkich oraz doczekał się wielu udanych praktycznych zastosowań. J.S.R. Jang zaproponował specyficzne dobieranie struktury sieci ANFIS: podział przestrzeni zmiennych wejściowych na zbiory rozmyte i wybór singletonów zmiennej wyjściowej oraz dobór wartości współczynnika sprzężenia zwrotnego (Łęski, 2008).

Ogólne podejście do konstrukcji systemów neuronowo rozmytych, których reguły warunkowe zawierały w konkluzjach zmienne rozmyte przedstawił S. Horikawa (Horikawa, Furuhashi, Okuma i Uchikawa, 1990). Metoda Horikawy wymaga jednak uzyskania wstępnej bazy wiedzy od eksperta, co prowadzi do trudności w przypadku systemów o dużej liczbie wejść. K. Cho i B. Wang zaproponowali użycie sieci neuronowej o radialnych funkcjach bazowych do uzyskania systemu neuronowo- rozmytego oraz przedstawili jej zastosowanie do identyfikacji systemów i predykcji ciągów czasowych. W 1998 roku J. Yen i in. zaproponowali metodę uczenia sieci opartą na optymalizacji wskaźników jakości systemu neuronowo-rozmytego (Łęski, 2008).

Szybki rozwój technologii informacyjnych prowadzi do wykorzystania ogromnych ilości danych w praktycznie każdej dziedzinie nauki i przemysłu. Dla prognozowania takiej ilości informacji istnieje potrzeba metod i narzędzi, które zapewnią automatyczne rozwiązania w wielu różnych kontekstach. Jednym z takich narzędzi do analizy i prognozowania szeregów czasowych jest ECOTOOL, nowy zestaw narzędzi MATLAB przedstawiony w artykule autorstwa D.J. Pedregal, (2019).

Analiza współczesnych prac wykazała obecność dużej liczby publikacji związanych z najbardziej aktualnym tematem na świecie - wybuchem pandemii koronawirusa i podejmowaniu niezbędnych środków w celu zareagowania na ten kryzys. Ten temat jest bardzo aktualny, w większości badań naukowych wykorzystano zarówno modele statystyczne ciągów czasowych jak i obliczenia miękkie do przewidywania zachorowalności, chorobowości i śmiertelności spowodowanych przez COVID-19. Warto wziąć pod uwagę że w tych pracach wykorzystano obydwa podejścia do analizy ale bez porównywania ich, co wskazuje na aktualność i nowość tej pracy doktorskiej.

W pracy Li z zespołem, (2020) opracowano funkcję przewidywania trwającego trendu za pomocą analizy opartej na danych i szacowania wielkości ogniska COVID-19 w Chinach. Roosa et al., (2020) wykorzystali sprawdzone modele aby stworzyć i ocenić krótkoterminowe prognozy łącznej liczby potwierdzonych przypadków. W pracy Fanelli i Piazza (2020) zrobiona analiza czasowa dynamiki pandemii COVID-19 w Chinach kontynentalnych, Włoszech i Francji. Roda et al. (2020) porównali standardowe ramy SIR (*ang. Susceptible, Infectious, or Recovered*) i SEIS (*ang. susceptible, exposed, infectious, then susceptible again*) do modelowania COVID-19. Wu et al. (2020) zrobili prognozę rozprzestrzenianie się COVID-19 w skali krajowej i globalnej. Anastassopoulou et al., (2020) badali modelowanie i przewidywanie rozprzestrzeniania się epidemii COVID-19.

W swoich badaniach Wang et al., (2020) opracowali algorytm oparty na informacjach o pacjencie do szacowania śmiertelności z powodu COVID-19 w czasie rzeczywistym z wykorzystaniem publicznie dostępnych danych.

W artykule Zeynep Ceylan, (2020) opracowano zintegrowane modele autoregresyjne ze średniej ruchomej (ARIMA) do przewidywania trendu występowania COVID-19 w krajach najbardziej dotkniętych epidemią. Badania z zastosowaniem modeli ARIMA opisano w następujących pracach naukowych Benvenuto, et al, (2020), Hiteshi Tandon, et al. (2020),

Lazzaro, Silvia, and Massimo, (2020), Soudeep Deb, Manidipa Majumdar, (2020), Teshome, Abebe, (2020), Vinay Kumar Reddy Chimmula, Lei Zhang, (2020), Xingde Duan, Xiaolei Zhang, (2020), Yonar, et al, (2020), Yong, Lingyun and Jia, (2020).

Wraz z rozwojem metod i technik sztucznej inteligencji, modele rozmyte i neuronowo-rozmyte stały się bardzo popularnym obiektem zastosowań dla analizy i prognoz ciągów czasowych Sandya, Hemanth, Himanshi, Susham, (2013), Tsai, Chen, (2018), Dhiman, Nitesh and Sharma, Mukesh,( 2019), Jianwei E, Ye J, Jin H, (2019), Ali Behnood,et al, (2020), Castillo, Melin, (2020), Dhiman, Nitesh & Sharma, Mukesh, (2020). W artykule Rashmi Bhardwaj, Aashima Bangia, (2020) badania oparte na danych do szacowania nowych wypadków zagrożeń związanych z rozprzestrzenianiem COVID-19 za pomocą hybrydowych technik obliczeniowych.

Również bardzo popularnym narzędziem do modelowania i prognozowania są hybrydowe systemy, Al-qaness et al., (2020) zaproponowali metodę optymalizacji prognozowania potwierdzonych przypadków COVID-19 za pomocą ulepszania hybrydowego adaptacyjnego systemu wnioskowania neuro-rozmytego (ANFIS). W pracy Paul A. Adedeji, et al, 2020 stosowali ANFIS do prognozowania zużycia energii na wielu kampusach uniwersyteckich.

Można zauważyć, że po okresie kilku dziesięcioleci istnienia *inteligencji obliczeniowej* w literaturze przedmiotu ciągle zbyt mało jest porównań *obliczeń miękkich* do klasycznych, statystycznych metod budowy modeli, identyfikacji oraz efektów w postaci dogodności stosowania, dokładności, nakładu obliczeń Walaszek-Babiszewska, Rydel i Kashpruk, (2018). Szersze studia na tym polu stanowią prace zespołu autorów, w osobach: J.L. Aznarte i J.M. Benitez, np. praca Aznarte i Benitez, (2013).



## 3. MODELE STATYSTYCZNE CIĄGÓW CZASOWYCH

### 3.1. Wprowadzenie

Modele statystyczne, wypracowane w obszarze statystyki matematycznej, odegrały przełomową rolę w analizie sygnałów dyskretnych i znalazły bardzo szerokie zastosowanie w wielu obszarach nauk. Modele te były rozwijane głównie na potrzeby ekonometrii i teorii sterowania.

Pewne uporządkowanie w teorii ciągów czasowych, głównie ze względu na zastosowanie w teorii sterowania, wprowadziła praca Boxa i Jenkinsa, dotycząca liniowych modeli szeregów czasowych: AR, MA, ARMA, ARIMA (Box i Jenkins, 1970).

Prace w obszarze ekonometrii zaowocowały kolejnymi typami modeli statystycznych, w tym także dotyczącymi modelowania nieliniowych ciągów czasowych, np. progowe modele autoregresyjne TAR (*ang. threshold autoregressive*), autoregresyjne modele płynnego przejścia STAR (*ang. smooth transition autoregressive*), SETAR (*ang. self-exciting threshold autoregressive*), ARCH (*ang. autoregressive conditional heteroscedasticity*), oraz szereg innych modeli (Brockwell i Davis, 1996), (Chatfield, 1975), (Chen i Tsay, 1993), (Welfe, 2009).

Badania statystyczne opierają się na analizie ciągów obserwacji procesów rzeczywistych. Przykładem takich danych są: notowania głównych indeksów rynków (WIG20, NASDAQ, Dow-Jones, S&P), notowania kursów walut, obserwacje zmienności procesów technologicznych, obserwacje parametrów pogody, obserwacje astronomiczne.

### 3.2. Procesy stochastyczne - podstawowe definicje i własności

Uporządkowany pod względem czasu zbiór zmiennych przypadkowych nazywany jest *procesem stochastycznym*, a realizację tych zmiennych w kolejnych okresach lub momentach czasu nazywamy ciągiem czasowym. Ciąg czasowy złożony z kolejnych  $N$  obserwacji  $x' = (x_1, x_2, \dots, x_N)$ , rozumiemy jako próbkę, pobraną z nieskończonej liczby próbek, które mogą być generowane przez ten proces (Fisz, 1967).

Proces stochastyczny  $X_t$  jest zbiorem zmiennych losowych, które są uporządkowane zgodnie ze zmianą czasu  $t$ :

$$\{X(t, \omega); X \in \chi, t \in T, \omega \in \Omega\}, \quad (3.1)$$

gdzie  $\chi \subseteq R$  jest zbiorem wartości procesu,  $T \subset R$  jest zbiorem wartości parametru  $t$  (na ogół czasu),  $\Omega$  jest przestrzenią zdarzeń elementarnych (Doob, 1953), (Fisz, 1967), (Papoulis, 1972).

Dla każdej ustalonej chwili  $t$ ,  $X(t)$  jest zmienną losową z rozkładem prawdopodobieństwa

$$F_t(x) = P\{X(t) < x\}.$$

Dla dowolnego zbioru wartości parametru  $t$ ,  $\{t_1, t_2, \dots, t_n\}$  proces stochastyczny  $X(t)$  ma łączny  $n$ -wymiarowy rozkład prawdopodobieństwa, określony przez dystrybuantę

$$F_{t_1, t_2, \dots, t_n}(x_1, x_2, \dots, x_n) = P\{X(t_1) < x_1, \dots, X(t_n) < x_n\}.$$

Proces stochastyczny jest w pełni opisany przez rodzinę wszystkich skończeniowymiarowych dystrybuant, dla  $n = 1, 2, \dots$  (Fisz, 1967).

Dla ustalonego zdarzenia elementarnego  $\omega' \in \Omega$  funkcja  $X(t) = X(t, \omega')$  jest realizacją tego procesu.

Podstawową definicją w teorii ciągów czasowych jest pojęcie *stacjonarności*. Proces stochastyczny  $X(t)$  jest stacjonarny w ścisłym sensie, jeśli jego  $n$ -wymiarowe dystrybuanty są niezmiennie z przesunięciem czasowym. Oznacza to, że dla każdego  $n$ , dla każdego podzbioru  $\{t_1, t_2, \dots, t_n\} \subset T \subseteq R$  i dla każdego  $\tau$  zachodzi relacja (Fisz, 1967):

$$F_{t_1, t_2, \dots, t_n}(x_1, x_2, \dots, x_n) = F_{t_1 + \tau, t_2 + \tau, \dots, t_n + \tau}(x_1, x_2, \dots, x_n),$$

przy czym

$$\{t_1 + \tau, t_2 + \tau, \dots, t_n + \tau\} \subset T.$$

Jeśli proces  $X(t)$  jest *stacjonarny w ścisłym sensie*, to w szczególności są spełnione charakterystyki:

$$F_t(x) = F_0(x) \text{ oraz } F_{t_1, t_2}(x_1, x_2) = F_{0, t_2 - t_1}(x_1, x_2).$$

Dane równości oznaczają że proces stacjonarny w ścisłym sensie ma jednowymiarowe dystrybuanty takie same dla dowolnych wartości parametru  $t$ , a dystrybuanty dwuwymiarowe zależą tylko od różnicy  $t_2 - t_1$  wartość parametru  $t$  (czasu).

Proces stochastyczny  $X(t)$  jest *stacjonarny w szerszym sensie*, jeśli jego wartość oczekiwana i wariancja nie zależą od czasu, a funkcja kowariancji zależy tylko od różnicy chwil  $t_2 - t_1 = \tau$ , czyli:

$$m_x(t) = m_x = \text{const},$$

$$\sigma_x^2(t) = \sigma_x^2 = \text{const},$$

$$K_X(t_2 - t_1) = K_X(\tau).$$

Funkcja kowariancji procesu stacjonarnego w szerszym sensie spełnia następujące warunki:

1)  $K_X(0) = \sigma_x^2 = \text{const}$ , (funkcja kowariancji procesu dla  $\tau=0$  jest równa wariancji, jest skończona i niezależna od  $t$ ),

2)  $K_X(\tau) = K_X(-\tau)$ , (symetryczna względem  $\tau$ ),

3)  $|K_X(\tau)| \leq K_X(0)$ , (malejąca dla rosnącej wartości  $\tau$ ).

Często w zastosowaniach praktycznych, dla procesów stacjonarnych jest używana unormowana funkcja kowariancji

$$k_X(\tau) = \frac{K_X(\tau)}{K_X(0)}, \quad |k_X(\tau)| \leq 1,$$

której wartości stanowią współczynniki korelacji pomiędzy zmiennymi losowymi  $X(t)$  oraz  $X(t + \tau)$ . Z warunku trzeciego funkcji autokowariancji wynika, że korelacja pomiędzy wyrazami procesu stacjonarnego maleje wraz ze wzrostem przedziału czasowego pomiędzy tymi wyrazami:

$$k_X(\tau) \rightarrow 0, \quad \text{gdy } |\tau| \rightarrow \infty.$$

Jeżeli proces stochastyczny jest stacjonarny w węższym sensie i istnieją momenty drugiego rzędu, to spełnione są warunki stacjonarności w szerszym sensie (Doob, 1953), (Fisz, 1967), (Papoulis, 1972).

Jeżeli przejścia procesu z jednego stanu w inny możliwe są tylko w ściśle ustalonych chwilach  $t_1, t_2, \dots, t_n$ , i w przedziałach czasu pomiędzy tymi momentami (krokami) proces zachowuje swoje poprzednie położenie to nazywamy go procesem z czasem dyskretnym.

Jeśli przejście procesu z jednego położenia w inne jest możliwe w każdym, z góry nieznanym momencie, to proces nazywamy procesem z czasem ciągłym.

Proces deterministyczny jest w każdej chwili czasu jednoznacznie określony. Charakteryzuje się tym, że na podstawie bieżącej próbki jesteśmy w stanie przewidzieć wartości procesu w przyszłości. Przykładem procesu deterministycznego jest trend liniowy.

Procesy *ergodyczne*, są to takie procesy, w których charakterystyki probabilistyczne po zbiorze mogą być wyrażone przez charakterystyki wyznaczone w długim odcinku czasu (Mańczak, 1970).

Aby teoria procesów stochastycznych mogła być stosowana w zagadnieniach praktycznych, potrzebna jest możliwość wyznaczania charakterystyk probabilistycznych procesów na podstawie jednej, dostatecznie długiej realizacji procesu. Wymagany jest warunek ergodyczności procesu (Mańczak, 1970)

### 3.3. Modele autoregresji

Autoregresja – to jest regresja statystyczna w której zmienna objaśniana jest przyszłą wartością ciągu, a zmienna objaśniająca to wartości ciągu czasowego z przeszłości.

Modele, w których bieżące wartości szeregu wyrażają się jako liniowe funkcje poprzednich obserwacji, nazywają się *modelami autoregresji* (Box i Jenkins, 1983), (Niederliński, 1979), (Welfe, 2009).

Jeżeli bieżąca wartość ciągu  $x_n$  zależy liniowo tylko od jednej poprzedniej wartości  $x_{n-1}$ , to taki model jest modelem autoregresji pierwszego rzędu, co oznaczamy jako AR(1).

Jeżeli bieżąca wartość ciągu  $x_n$  zależy liniowo od  $p$  poprzednich wartości  $x_{n-1}, \dots, x_{n-p}$ , to taki model jest modelem autoregresji rzędu  $p$ , AR( $p$ ).

Proces autoregresji skończonego rzędu,  $p > 1$ , przedstawia wzór:

$$x_n = a_1 x_{n-1} + a_2 x_{n-2} + \dots + a_p x_{n-p} + \varepsilon_n, \quad (3.2)$$

gdzie:

$a_1, \dots, a_p$  – parametry modelu procesu,

$\varepsilon_n$  – proces bezwzględnie losowy, tzw. biały szum, o własnościach:

- a)  $E[\varepsilon_n] = 0$ ,
- b)  $D^2[\varepsilon_n] = \sigma_\varepsilon < \infty$
- c)  $cov[\varepsilon_n, \varepsilon_{n-s}] = 0, s \neq 0$ .

### Proces stochastyczny autoregresji pierwszego rzędu AR(1) i jego własności

Procesem autoregresji pierwszego rzędu, AR(1), nazywana jest klasa modeli procesów stochastycznych w dyskretnej przestrzeni czasu o postaci (Welfe, 2009), (Walaszek-Babiszewska, 2015):

$$x_n = \alpha x_{n-1} + \varepsilon_n, n=2, \dots, N, \quad (3.3)$$

gdzie:  $\alpha$  jest nieznanym parametrem procesu,  $\varepsilon_n$  jest nieznanym procesem bezwzględnie losowym, białym szumem.

Aby wyznaczyć wartość oczekiwaną procesu  $E[x_n]$ , przekształćmy wzór (3.3), podstawiając kolejno dla  $x_{n-1}, x_{n-2}, \dots$  postać procesu (3.3):

$$x_n = \alpha x_{n-1} + \varepsilon_n = \alpha(\alpha x_{n-2} + \varepsilon_{n-1}) + \varepsilon_n = \dots = \alpha^m x_{n-m} + \alpha^{m-1} \varepsilon_{n-(m-1)} + \dots + \varepsilon_n \quad (3.4)$$

Dla  $|\alpha| < 1 \quad \lim_{m \rightarrow \infty} \alpha^m = 0$  i z wyrażenia (3.4) wynika, że  $E[x_n] = 0$ .

Można wykazać, że wariancja procesu AR(1), dla  $|\alpha| < 1$  wynosi:

$$D^2[x_n] = \sigma_X^2 = \sigma_\varepsilon^2 \frac{1}{1-\alpha^2} \quad (3.5)$$

Wariancja procesu AR(1) zależy od wariancji szumu i parametru  $\alpha$ .

Kowariancja procesu AR(1),  $cov[x_n, x_{n-m}]$ , dla wartości parametru  $|\alpha| < 1$ , maleje z potęgą  $\alpha^m$ , a mianowicie (Welfe, 2009):

$$cov[x_n, x_{n-m}] = \sigma_\varepsilon^2 \frac{\alpha^m}{1-\alpha^2} = \sigma_X^2 \alpha^m, m=0,1,2, \dots \quad (3.6)$$

Zatem macierz unormowanej funkcji autokorelacji,  $k_x(m) = k_x(x_n, x_{n-m})$ , można zapisać w postaci (Welfe, 2009):

$$k_x(m) = \begin{bmatrix} 1 & \alpha & \alpha^2 & \dots & \alpha^{m-1} \\ \alpha & 1 & \alpha & \dots & \alpha^{m-2} \\ \alpha^2 & \alpha & 1 & \dots & \alpha^{m-3} \\ \dots & \dots & \dots & \dots & \dots \\ \alpha^{m-1} & \alpha^{m-2} & \alpha^{m-3} & \dots & 1 \end{bmatrix}. \quad (3.7)$$

## Proces stochastyczny autoregresji drugiego rzędu AR(2)

Procesem autoregresji drugiego rzędu, AR(2), jest klasa modeli procesów stochastycznych w dyskretnej przestrzeni czasu o postaci:

$$x_n = ax_{n-1} + bx_{n-2} + \varepsilon_n, \quad n=3, \dots, N \quad (3.8)$$

gdzie:  $a, b$  są parametrami modelu,  $\varepsilon_n$  jest procesem stochastycznym białego szumu.

Modele autoregresji są wykorzystywane do opisu procesów stacjonarnych. Charakterystyczną właściwością stacjonarnych ciągów czasowych jest to, że funkcje probabilistyczne stacjonarnych ciągów nie zmieniają się z przesunięciem czasowym.

Proces stochastyczny rzędu drugiego ma zawsze funkcję korelacyjną (Swiesznikow, 1965), (Sobczyk, 1996).

## Modele średniej ruchomej (MA)

Proces ze średnią ruchomą, jest regresją liniową bieżącej wartości szeregów w odniesieniu do poprzednich wartości błędu, związanych z białym szumem lub przypadkowymi zaburzeniami. Zakłada się, że takie przypadkowe zaburzenia w każdym z punktów pochodzą z tego samego rozkładu, zwykle rozkładu normalnego z zerową wartością oczekiwaną i stałą wariancją. Charakterystyczna dla tego modelu jest propagacja losowych zaburzeń do przyszłych wartości szeregów czasowych (Niederliński, Kasprzyk i Figwer, 1997).

Model ze średnią ruchomą rzędu  $q$ , MA( $q$ ) (*ang. moving average*), ma postać (Welfe, 2009):

$$x_n = \varepsilon_n + b_1\varepsilon_{n-1} + b_2\varepsilon_{n-2} + \dots + b_q\varepsilon_{n-q}, \quad (3.9)$$

gdzie:  $b_1, \dots, b_q$  są parametrami modelu,  $\varepsilon_n$  jest procesem białego szumu z czasem dyskretnym.

Rozpatrzmy proces ze średnią ruchomą i jego własności na podstawie najprostszego wzoru modelu MA(1):

$$x_n = \tau_n + b\tau_{n-1}, \quad (3.10)$$

gdzie:  $\tau_n$  jest procesem białego szumu o własnościach:

- a)  $E[\tau_n] = 0$ ,
- b)  $D^2[\tau_n] = \theta_\tau^2$ ,
- c)  $cov(\tau_n, \tau_{n-s}) = 0, \quad s \neq 0$ .

Wartość oczekiwaną, wariancję i funkcję kowariancji procesu MA(1) można wyznaczyć z odpowiednich definicji, w następujący sposób (Welfe, 2009), (Walaszek-Babiszewska, 2015):

- 1)  $E(x_n) = E(\tau_n) + bE(\tau_{n-1}) = 0$ ;
- 2)  $D^2(x_n) = D^2(\tau_n) + b^2D^2(\tau_{n-1}) = \theta_\tau^2(1 + b^2)$

$$\begin{aligned}
3) \operatorname{cov}(x_n, x_{n-1}) &= E\{[x_n - E(x_n)][x_{n-1} - E(x_{n-1})]\} = E[x_n x_{n-1}] = \\
&= E[(\tau_n + b\tau_{n-1})(\tau_{n-1} + b\tau_{n-2})] = \\
&= E[(\tau_n \tau_{n-1} + b\tau_{n-1}^2 + b\tau_n \tau_{n-2} + b^2 \tau_{n-1} b\tau_{n-2})] = E[b\tau_{n-1}^2] = b\theta_\tau^2 \\
\operatorname{cov}(x_n, x_{n-s}) &= 0 \text{ dla } s > 1
\end{aligned}$$

Proces średniej ruchomej pierwszego rzędu ma wartość średnią równą zero, skończoną wariancję zależną od wariancji zakłóceń i parametru modelu  $b$ . Proces MA(1) jest procesem o wyrazach nieskorelowanych, czyli tak jak biały szum.

Model średniej ruchomej ma swoje wady i zalety:

- do zalet należą łatwość i szybkość w obliczeniach,
- do wad można zaliczyć małą dokładność prognozy, utratę danych w ciągu, brak uwzględnienia trendów i efektów sezonowych.

### Model autoregresji ze średnią ruchomą (ARMA)

Model autoregresji ze średnią ruchomą ARMA (*ang. autoregressive moving average*), jest modelem matematycznym służącym do analizy i przewidywania stacjonarnych szeregów czasowych w statystyce (Zhang, Moore, 2015), (James Chen, 2019).

Model ARMA obejmuje dwa proste modele ciągów czasowych - model autoregresji (AR) i model średniej ruchomej (MA). Proces autoregresji skończonego rzędu  $s$ ,  $s > 1$  przedstawia wzór:

$$x_n = a_1 x_{n-1} + a_2 x_{n-2} + \dots + a_s x_{n-s} + \varepsilon_n, \quad (3.11)$$

gdzie:

$a_1, \dots, a_s$  – parametry modelu procesu,

$\varepsilon_n$  – składnik losowy, który jest procesem stochastycznym białego szumu.

Proces ze średnią ruchomą rzędu  $q$ , MA( $q$ ) ma postać:

$$x_n = \varepsilon_n + b_1 \varepsilon_{n-1} + b_2 \varepsilon_{n-2} + \dots + b_q \varepsilon_{n-q} \quad (3.11a)$$

gdzie:

$b_1, \dots, b_q$  – parametry modelu,

$\varepsilon_n$  – składnik losowy procesy białego szumu.

Proces autoregresji za średnią ruchomą rzędu  $s$ ,  $q$ , ARMA( $s$ ,  $q$ ) jest kombinacją obydwu procesów (Welfe, 2009):

$$x_n = a_1 x_{n-1} + a_2 x_{n-2} + \dots + a_s x_{n-s} + \varepsilon_n + b_1 \varepsilon_{n-1} + b_2 \varepsilon_{n-2} + \dots + b_q \varepsilon_{n-q} \quad (3.12)$$

Modele ARMA są używane do opisu stacjonarnych ciągów czasowych. Także niektóre rodzaje niestacjonarności można sprowadzać do procesów stacjonarnych przez tworzenie procesów różnic i dalej stosować np. modele ARMA. Ponad to poniżej przedstawiono klasę autoregresyjnych zintegrowanych modeli średniej ruchomej (ARIMA).

Użyteczność modeli ARMA polega na ich oszczędnej reprezentacji. Podobnie jak w przypadku AR i MA, właściwości modeli ARMA można zwykle scharakteryzować za pomocą ich funkcji autokorelacji (Chan, Ngai, 2011), (Zhang, Moore, 2015).

## Model ARIMA

Model ARIMA (*ang. autoregressive integrated moving average*) jest to zintegrowany proces ARMA. Na ogół modele zintegrowane służą do opisów procesów stochastycznych niestacjonarnych (Stellwagen i Tashman, 2013).

ARIMA jest modelem do analizy statystycznej, który wykorzystuje dane ciągów czasowych do przewidywania przyszłych tendencji.

Zintegrowany model autoregresji ze średnią ruchomą jest formą analizy regresji, która mierzy siłę jednej zmiennej zależnej w stosunku do innych zmiennych. Model ten jest szeroko stosowany do prognozowania przyszłych wartości danych, badając różnice między wartościami w ciągach zamiast wartości oryginalnych.

Model ARIMA można uważać jako złożony z następujących elementów: autoregresja (AR), całkowanie (I) oraz średnia ruchoma (MA).

Przez autoregresję rozumiemy model, który pokazuje regresję zmiennej zgodnie z własnymi poprzednimi wartościami.

Całkowanie to proces różnic nieprzetworzonych obserwacji, aby utworzyć ciągi czasowe, których wyrazy są różnicami między wartościami danych w chwilach sąsiednich (James Chen, 2019).

W przypadku modeli ARIMA domyślną notacją będzie ARIMA z parametrami  $p$ ,  $d$  i  $q$ , gdzie:

$p$  - liczba opóźnień obserwacji w modelu.

$d$  - stopień różnicy.

$q$  - zakres średniej ruchomej.

Model ARIMA można zbudować na wiele sposobów, aby pełnił funkcję modelu ARMA, lub prostych modeli AR, I lub MA. Modele ARIMA są wykorzystywane przy opisie szumu w zakłóceniach (James Chen, 2019).

Podstawą budowy modeli ARIMA są procesy różnic, tworzone w oparciu o obserwacje podstawowego procesu  $\{x_n, x_{n-1}, x_{n-2}, \dots, x_{n-m}, \dots\}$ , w sposób następujący:

- proces pierwszych różnic:

$$\begin{aligned}\Delta^1 X_n &= x_n - x_{n-1}, \\ \Delta^1 X_{n-1} &= x_{n-1} - x_{n-2} \dots \\ \Delta^1 X_{n-m} &= x_{n-m} - x_{n-m}.\end{aligned}$$

- proces drugich różnic:

$$\Delta^2 X_n = \Delta^1 X_n - \Delta^1 X_{n-1} = x_n - x_{n-1} - (x_{n-1} - x_{n-2}) = x_n - 2x_{n-1} + x_{n-2} \dots$$

- proces trzecich różnic:

$$\begin{aligned}\Delta^3 X_n &= \Delta^2 X_n - \Delta^2 X_{n-1} = x_n - 2x_{n-1} + x_{n-2} - (\Delta^1 X_{n-1} - \Delta^1 X_{n-2}) = x_n - 2x_{n-1} + \\ & x_{n-2} - [x_{n-1} - x_{n-2} - (x_{n-2} - x_{n-3})] = x_n - 3x_{n-1} + 3x_{n-2} - x_{n-3}.\end{aligned}$$

Jeśli, dla przykładu, proces różnic rzędu  $l$ ,  $\Delta^l X_n$ , zostanie opisany modelem AR( $p$ ), to po estymacji parametrów  $a_1, \dots, a_p$  należy rozwiązać powyższe układy równań i wyznaczyć proces podstawowy.

## ARX

Jednym z modeli obiektów dynamicznych, jest model autoregresji z zewnętrznym (dodatkowym) wejściem – ARX (*ang. autoregressive with exogenous input*). Model wejściowo-wyjściowy ARX stosuje się dla obiektów stochastycznych. Wyjście obiektu w chwili  $n$ -tej jest liniowo zależne od poprzednich swoich wartości oraz opóźnionego wejścia. Model ARX ma postać (Niederliński, 1979):

$$x_n + a_1 x_{n-1} + \dots + a_p x_{n-p} = b_1 u_{n-s} + \dots + b_k u_{n-s-k} + e_n \quad (3.13)$$

Modele ARX stosuje się do opisu systemów, gdzie  $x$  oznacza zmienną wyjściową systemu,  $u$  oznacza zmienną wejściową systemu,  $e$  – jest procesem bezwzględnie losowym. Rząd modelu ARX jest określony przez liczby kroków wejścia i wyjścia branych pod uwagę oraz przez opóźnienie  $s$ .

ARX jest najprostszym modelem procesu dynamicznego przedstawionego przez dane wejściowe w obecności niepewności. Modele te opisują badany wynik procesu jako sumę regresji poprzednich obserwacji funkcji wejścia-wyjścia oraz białego szumu, który opisuje błąd równania (Soderstrom, Stoica, 1997). Ten kontekst stochastyczny, podobnie jak wszystkie inne modele błędów równań, nie daje wyraźnych założeń dotyczących pochodzenia ich zgodności między obserwacjami a wynikiem procesu (Diversi, et al, 2010).

## TAR

Prace w obszarze ekonometrii zaowocowały kolejnymi typami modeli statystycznych, w tym także dotyczącymi modelowania nieliniowych ciągów czasowych, np. progowe modele autoregresyjne TAR (*ang. threshold autoregressive*) oraz autoregresyjne modele płynnego przejścia STAR (*ang. smooth transition autoregressive*) (Welfe, 2009).

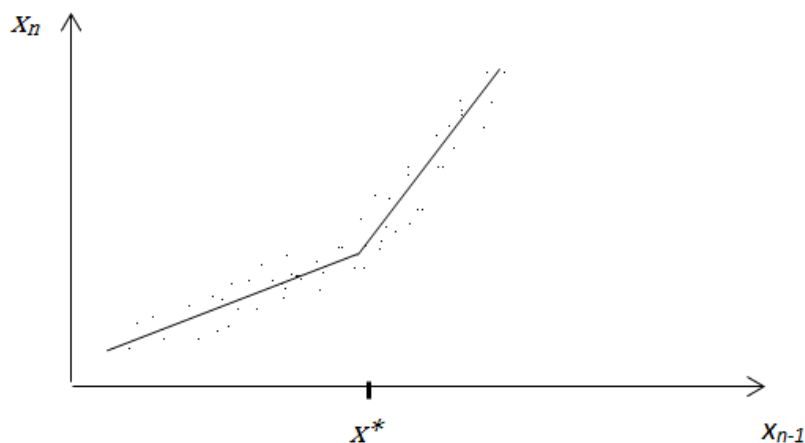
W praktyce prognoz, warunek stacjonarności procesu rzadko jest spełniony, zwłaszcza co do stałej wartości średniej, dlatego też liniowe modele autoregresji wyznaczane dla całego zbioru obserwacji przestają być dość dobrym odwzorowaniem.

Jednym z rozwiązań jest odcinkowo liniowy model, np. progowy model autoregresji (TAR). Jeśli istnieją podstawy aby przypuszczać, że w pewnej części przestrzeni stanu procesu,  $\chi$ , badany proces opisują różne modele autoregresji, wówczas model TAR daje możliwość takiego opisu, zarazem łącząc modele lokalne w model całościowy. Dla wybranej zmiennej, np. zmiennej niezależnej  $x_{n-1}$  dzieli się przestrzeń  $\chi$  na rozłączne przedziały, w których wyznaczone zostają modele autoregresji  $x_n^r = f(x_{n-s})$ ,  $s=1,2$  dla każdego  $r$ -tego przedziału. Przejście z jednej linii regresji na kolejną określa tzw. funkcja przejścia,  $\lambda_r$ , która przyjmuje wartość 1 dla  $r$ -tego przedziału, a w pozostałych jest równa zero. Model globalny można zapisać w postaci

$$x_n = \sum_r \lambda_r x_n^r, \quad r = 1, 2, \dots, R \quad (3.14)$$

gdzie  $x_n^r$  jest modelem autoregresji AR(1) lub AR(2) wyznaczonym w  $r$ -tym przedziale wybranej zmiennej. Granice przedziałów wyznaczają progi modelu,  $x^*$ .



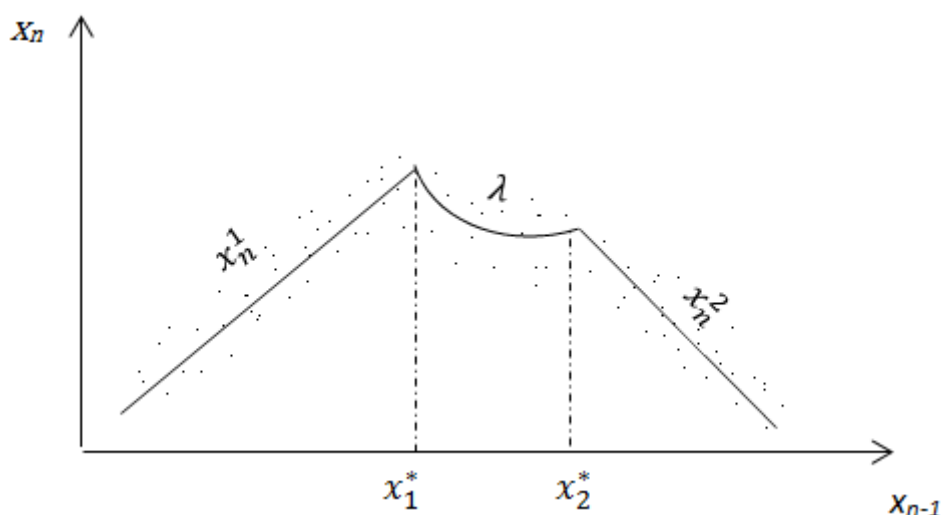


Rys. 3.1. Przykład modelu TAR na tle zbioru obserwacji w układzie współrzędnych  $(x_n, x_{n-1})$

Ogólnie, model TAR jest zbiorem modeli AR, wykorzystywanych w przypadku nieliniowości w obliczeniach (Aznarte i Benitez, 2013). Praktyczne zastosowanie modelu TAR zostało przedstawione w rozdziale 6.

### STAR

Płynny, przejściowy model autoregresji, STAR, (*ang. smooth transition autoregressive*) zawiera ciągłe funkcje przejścia  $\lambda_r(b^r, x_{n-1}^r)$ , które określają położenie i kształt przejścia między lokalnymi modelami autoregresyjnymi  $x_n^r$ . Funkcja przejścia  $\lambda_r(b^r, x_{n-1}^r)$  przyjmuje swoje wartości z przedziału jednostkowego  $[0,1]$  i zwykle jest nieliniową funkcją zmiennych niezależnych (Walaszek-Babiszewska, Rydel i Kashpruk, 2018).



Rys. 3.2. Przykład modelu STAR na tle zbioru obserwacji w układzie współrzędnych  $(x_n, x_{n-1})$

Modele STAR, charakteryzujące się płynnym a nie dyskretnym przełączaniem reżimów, są stosowane w zaawansowanej analizie finansowej dla wypadkowej dynamiki ciągów czasowych, gdy decyzje inwestorów nie są podejmowane jednocześnie (Aznarte i Benitez, 2013).

Tab. 3.1 Tabela porównawcza modeli parametrycznych

	Modele				Ilość zmiennych		Liczba parametrów		Modele procesów	
	Linowe	Nieliniowe	Sygnalów	Obiektów	jedna	wiele	jeden	wiele	stacjonarne	niestacjonarne
AR	+		+		+		+		+	
MA	+		+		+		+		+	
ARMA	+		+		+			+	+	
ARIMA	+		+		+			+		+
ARX	+			+		+		+	+	
TAR		+	+		+			+	+	
STAR		+	+		+			+	+	

W tabeli 3.1 porównano omówione w tym rozdziale modele parametryczne procesów stochastycznych.

## 4. WYBRANE METODY IDENTYFIKACJI

Identyfikacją procesu lub systemu, nazywa się w automatyce proces ustalania modelu matematycznego systemu na podstawie badań eksperymentalnych. Celem identyfikacji może być ustalenie zależności pomiędzy zmiennymi obiektu w stanie ustalonym lub czasie trwania procesów przejściowych. Mówimy wtedy o identyfikacji charakterystyki (modelu) statycznej lub dynamicznej. Na ogół zakłada się pewną klasę modeli hipotetycznych. Uporządkowanie w klasie modeli wprowadza kryterium identyfikacji (Mańczak, 1971), (Bubnicki, 1974), (Mańczak i Nahorski, 1983).

W bieżącym rozdziale będziemy się zajmować identyfikacją modeli statystycznych. Przedstawimy metodę najmniejszej sumy kwadratów błędów modelu, która obok metody największej wiarygodności, jest jednym z najczęściej stosowanych kryteriów wyboru modelu parametrycznego. Wybór modelu dokonuje się przez wyznaczenie parametrów, w oparciu o pewien zbiór danych.

Wprowadzono wzory do obliczenia estymatorów modeli AR(1) i AR(2), przyjęto miary oceny dopasowania modeli do danych doświadczalnych, w tym także funkcję autokorelacji reszt modelu, wyprowadzono wzory na składowe trendu modeli. Przedstawiono schemat procesu identyfikacji modeli AR(1) i AR(2), który stał się podstawą do algorytmu obliczeń.

W rozdziale zostały przedstawione przykłady obliczeń identyfikacyjnych modeli AR(1) i AR(2) dla różnych typów empirycznych ciągów czasowych. Zbiory danych stanowią realizacje procesów stochastycznych z czasem dyskretnym o różnych właściwościach.

Ponadto, opracowano algorytm wyznaczania parametrów modelu systemu stochastycznego ARX. W przykładzie identyfikacji wykorzystano ciągi obserwacji dwóch zmiennych dotyczące tego samego procesu technologicznego.

### 4.1. Metoda najmniejszych kwadratów dla modeli parametrycznych

Metoda najmniejszej sumy kwadratów błędów modelu, jest jednym z kryteriów wyboru modelu, z przyjętej klasy modeli parametrycznych, i pozwala określić parametry modelu odpowiednie dla danych eksperymentalnych.

Ta metoda została zdefiniowana jako element klasycznej metody analizy regresji, w której przyjmuje się tzw. założenia Gaussa-Markowa dotyczące probabilistycznych własności zmiennych modelu. Takie założenia pozwalają na wnioskowanie statystyczne dotyczące estymatorów modeli (Mańczak, 1971), (Welfe, 2009). Sama metoda najmniejszych kwadratów nie wymaga takich założeń i pozwala wyznaczyć parametry modelu, którego postać jest liniowa względem parametrów.

Analiza regresji, przy znanych rozkładach warunkach Gaussa-Markowa, pozwala na wyznaczenie istotności równania regresji i istotności i współczynników  $\hat{a}$  i  $\hat{b}$  (warunki Gaussa-Markowa) (Mańczak, 1971), (Welfe, 2009).

W niniejszej pracy będzie stosowana wyłącznie metoda najmniejszych kwadratów, ponieważ analiza regresji wymaga rozkładów normalnych badanych zmiennych co w badaniach doświadczalnych rzadko jest spełnione.

Założmy, że dany jest ciąg obserwacji  $\{(x_i, y_i)\}_{i=1,2,\dots,N}$ , zmiennej wejściowej (niezależnej)  $x$  i zmiennej losowej wyjściowej (zależnej)  $y$ . Poszukuje się warunkowej wartości oczekiwanej:

$$\hat{y} = E[y|x], \quad (4.1)$$

dla której zakładamy postać liniową (prostą regresji pierwszego rzędu)

$$E\{y|x\} = \hat{a}x + \hat{b}. \quad (4.2)$$

spełniająca warunek:

$$E[y_i - (\hat{a}x_i + \hat{b})]_{a,b}^2 \min.$$

Estymatory parametrów modelu  $\hat{a}$  i  $\hat{b}$  wyznacza się minimalizując sumę kwadratów błędów modelu:

$$E[e_i]^2 \cong \frac{1}{N} \sum [y_i - (\hat{a}x_i + \hat{b})]_{a,b}^2 \min \quad (4.3)$$

Warunek konieczny istnienia ekstremum prowadzi do układu równań algebraicznych.

$$\begin{aligned} \frac{d}{d\hat{a}} \left\{ \frac{1}{N} \sum_{i=1}^N [y_i - (\hat{a}x_i + \hat{b})]^2 \right\} &= 0 \\ \frac{d}{d\hat{b}} \left\{ \frac{1}{N} \sum_{i=1}^N [y_i - (\hat{a}x_i + \hat{b})]^2 \right\} &= 0 \end{aligned}$$

Po wykonaniu działań otrzymujemy:

$$\begin{cases} \hat{a} \sum_{i=1}^N x_i^2 + \hat{b} \sum_{i=1}^N x_i = \sum_{i=1}^N x_i y_i \\ \hat{a} \sum_{i=1}^N x_i + \hat{b} N = \sum_{i=1}^N y_i \end{cases}$$

Dla uproszczenia zapisów, podstawiamy odpowiednie litery A,B,C,D, do układu równań algebraicznych:

$$\begin{cases} \hat{a}C + \hat{b}A = D \\ \hat{a}A + \hat{b}N = B \end{cases}$$

Układ równań algebraicznych rozwiążemy wykorzystując wyznaczniki:

$$W = \begin{vmatrix} C & A \\ A & N \end{vmatrix} = CN - A^2, W \neq 0.$$

$$W_{\hat{a}} = \begin{vmatrix} D & A \\ B & N \end{vmatrix} = DN - BA,$$

$$W_{\hat{b}} = \begin{vmatrix} C & D \\ A & B \end{vmatrix} = CB - AD.$$

Ostatecznie poszukiwane estymatory  $\hat{a}$ ,  $\hat{b}$  wynoszą:

$$\hat{a} = \frac{W_{\hat{a}}}{W}, \quad \hat{b} = \frac{W_{\hat{b}}}{W}.$$

Poszukiwane równanie regresji (4.2) jest opisane za pomocą wyznaczonych estymatorów:

$$\hat{y} = \hat{a}x + \hat{b} \quad (4.4)$$

Zasada najmniejszych kwadratów może także być wykorzystana dla identyfikacji modeli dynamicznych. Analogicznie jak dla modelu statycznego (4.2) przeprowadza się obliczenia parametrów dla modeli dynamicznych, np. AR(1), AR(2).

### Obliczenia estymatora modelu AR(1)

Jeśli badany jest rzeczywisty proces stochastyczny, to przyjęcie modelu AR(1) jest pewną hipotezą. Zakładając stacjonarność i ergodyczność badanego procesu możliwe jest, w oparciu o zbiór obserwacji  $\{x_n\}$  procesu, wyznaczenie estymatora modelu AR(1). Do obliczenia estymatora  $\hat{a}$  w modelu AR(1) można wykorzystać metodę najmniejszych kwadratów, poszukując minimum następującej funkcji kryterialnej:

$$Q = \sum_{n=2}^N (x_n - \hat{a}x_{n-1})^2 \xrightarrow{\hat{a}} \min. \quad (4.5)$$

Warunek konieczny istnienia ekstremum  $\frac{dQ}{d\hat{a}} = 0$ , ma postać:

$$\frac{dQ}{d\hat{a}} = \sum_{n=2}^N 2(x_n - \hat{a}x_{n-1})(-x_{n-1}) = 0. \quad (4.6)$$

Po przekształceniach otrzymujemy:

$$\sum_{n=2}^N (x_n - \hat{a}x_{n-1})x_{n-1} = 0,$$

$$\sum_{n=2}^N x_n x_{n-1} = \hat{a} \sum_{n=2}^N x_{n-1}^2.$$

Estymator  $\hat{a}$  metody najmniejszych kwadratów parametru  $\alpha$  wyznacza się z danych empirycznych  $\{x_n\}$ , według wzoru:

$$\hat{a} = \frac{\sum_{n=2}^N x_n x_{n-1}}{\sum_{n=2}^N x_{n-1}^2} \quad (4.7)$$

Ponadto, dla uzupełnienia wiedzy o jakości modelu, wyznacza się następujące wskaźniki badanego procesu:

$\bar{x}$  - wartość średnia w zbiorze danych,

$D_x^2$  - wariancja w zbiorze danych,

$e^2$  - suma kwadratów reszt modelu (suma kwadratów błędów modelu),

$$e^2 = \sum_n (x_n - \hat{a}x_{n-1})^2, \quad (4.8)$$

oraz jej wartość średnia

$$\bar{E} = \frac{1}{N-1} e^2. \quad (4.9)$$

### Obliczenia estymatorów parametrów $a, b$ modelu AR(2)

Analogicznie jak dla modelu AR(1), przy założeniu stacjonarności i ergodyczności badanego, rzeczywistego procesu w oparciu o zbiór danych obserwacji procesu można wyznaczyć estymatory hipotetycznego modelu AR(2). Obliczenia estymatorów parametrów  $a, b$  modelu AR(2), metodą najmniejszych kwadratów przebiega analogicznie jak dla modelu AR(1):

$$Q = \sum_{n=3}^N [x_n - (a x_{n-1} + b x_{n-2})]^2 \xrightarrow{a,b} \min \quad (4.10)$$

Wychodząc z warunku koniecznego istnienia ekstremum funkcji (4.10) doprowadzimy do układu dwóch równań algebraicznych ze względu na  $a$  i  $b$ .

$$\begin{cases} \frac{dQ}{da} = 0 \\ \frac{dQ}{db} = 0 \end{cases}, \quad (4.11)$$

$$\begin{cases} a \sum_{n=3}^N x_{n-1}^2 + b \sum_{n=3}^N x_{n-1} x_{n-2} = \sum_{n=3}^N x_n x_{n-1} \\ a \sum_{n=3}^N x_{n-1} x_{n-2} + b \sum_{n=3}^N x_{n-2}^2 = \sum_{n=3}^N x_n x_{n-2} \end{cases}$$

Dla uproszczenia zapisu, oznaczmy jako  $A, B, C, D, E$  wartości odpowiednich sum w układzie równań (4.11).

$$\begin{cases} aA + bB = C \\ aB + bE = D \end{cases}.$$

Powyższy układ równań algebraicznych rozwiążemy za pomocą wyznaczników:

$$W = AE - B^2, \quad W \neq 0$$

$$W_a = \begin{vmatrix} C & B \\ D & E \end{vmatrix} = CE - DB,$$

$$W_b = \begin{vmatrix} A & C \\ B & D \end{vmatrix} = AD - BC,$$

$$\hat{a} = \frac{W_a}{W},$$

$$\hat{b} = \frac{W_b}{W}.$$

Wyznaczony model AR(2) ma więc postać:

$$x_n = \hat{a} x_{n-1} + \hat{b} x_{n-2}. \quad (4.12)$$

## 4.2 Praktyczne metody doboru modeli

Istnieje wiele sposobów szacowania współczynników, takich jak:

- zwykła procedura najmniejszych kwadratów,
- metoda równań Yule-Walkera (metoda momentów),
- metoda Burga,
- algorytm innowacji,
- algorytm Hannana-Rissanena.

Wstępne oszacowanie dla modeli AR(p) polega na zastosowaniu metody Yule-Walkera lub Burga, podczas gdy dla procesów MA (q) używamy algorytmu innowacji lub algorytmu Hannana-Rissanena. Po dopasowaniu modelu wstępnego wyszukujemy estymatory maksymalnego prawdopodobieństwa Gaussa, korzystając z algorytmu innowacji. Wówczas można łatwo uzyskać przedziały ufności dla oszacowanych współczynników (Sahu, 2005).

**Równania Yule-Walkera**, to metoda momentów polegająca na porównaniu wartości kowariancji badanych danych. Jeśli zastąpimy funkcję autokorelacji jej oszacowaniem z próbki i rozwiążemy równania, otrzymamy oszacowanie Yule-Walkera dla parametrów. Przykładowe oszacowanie dla funkcji autokorelacji przedstawia wzór:

$$\gamma_m = E[X_{t-m}X_t] = \sum_{j=1}^p \varphi_j E[X_{t-m}X_{t-j}] + E[X_{t-m}Z_t] = \sum_{j=1}^p \varphi_j \gamma_{|m-j|} + \sigma^2 \delta_m \quad (4.13)$$

dla  $m > 0$ , mamy układ równań:

$$\begin{bmatrix} \gamma_1 \\ \gamma_2 \\ \dots \\ \gamma_p \end{bmatrix} = \begin{bmatrix} \gamma_0 & \gamma_1 & \gamma_2 & \dots \\ \gamma_1 & \gamma_0 & \gamma_1 & \dots \\ \gamma_2 & \gamma_1 & \gamma_0 & \dots \\ \dots & \dots & \dots & \dots \end{bmatrix} \begin{bmatrix} \varphi_1 \\ \varphi_2 \\ \dots \\ \varphi_p \end{bmatrix}$$

Z uwagi na estymację  $\gamma_i$ , odwracamy ten system, aby uzyskać wynik  $\varphi_j$  dla  $j > 1$ . Macierz jest odwracalna, ponieważ jest dodatnia. Po czym podstawiamy do równania, aby znaleźć  $\sigma^2$ . To podejście nazywa się równaniami Yule-Walkera. Dana metoda pokazuje, że zawsze można znaleźć model AR dopasowany do dowolnej liczby momentów w ACF ([en.wikipedia.org/wiki/Autoregressive\\_model#Yule-Walker\\_equations](http://en.wikipedia.org/wiki/Autoregressive_model#Yule-Walker_equations)).

Kolejnym algorytmem estymacji parametrów jest **algorytm Burga**. Ten algorytm minimalizuje błędy predykcji liniowej za pomocą średniej harmoniczej współczynników korelacji częściowej. Procesy liniowe są niezmiennie w odniesieniu do odwrócenia czasu. Główną zaletą tego algorytmu jest to, że zawsze zapewnia stabilne modele. Przewyższa ona metodę Yule-Walkera dla krótkich zapisów danych, a także usuwa kompromis między używaniem obciążonych i nieobciążonych oszacowań autokorelacji podczas dzielenia wag ([opus-codec.org/docs/vos\\_fastburg](http://opus-codec.org/docs/vos_fastburg)).

**W algorytmie Hannana-Rissanena** wyznaczenie parametrów jest podzielone na dwa etapy:

1. Dopasowujemy model AR (m) wyższego rzędu do danych, gdzie  $m > \max(p, q)$ . Model ten jest używany do wyznaczania warunków szumu.

$$\epsilon_n = x_n - \sum_{j=1}^m \hat{a}_j x_{n-j} \quad (4.14)$$

2. W drugim kroku parametry modelu ARMA (p,q) są estymowane przez liniową regresję  $x_n$  metodą najmniejszych kwadratów na  $(x_{n-1}, \dots, x_{n-j}, \epsilon_{n-1}, \dots, \epsilon_{n-j})$  (Eckehard Olbrich, 2008).

Dla nieliniowych modeli stosujemy **metodę największej wiarygodności, MLE** (*ang. maximum likelihood estimation*). Różnica między metodą MLE a metodą najmniejszych kwadratów polega na tym, że metoda MLE uwzględnia rozkład pierwszych składników  $p$ .

Estymatorem największej wiarygodności parametru  $\theta$  nazywamy wartość parametru, dla której funkcja wiarygodności przyjmuje wartość największą. Metoda MLE może być dość wolna w przypadku dużych zbiorów danych. Zazwyczaj używa się jednej z powyższych metod aby znaleźć dobre wstępne przybliżenie, a następnie wykorzystuje się optymalizację numeryczną do znalezienia MLE (Kochanek, (2016).

Podstawą metody MLE jest funkcja wiarygodności zbudowana na podstawie wartości funkcji gęstości prawdopodobieństwa dla poszczególnych elementów próby losowej. Estymacja parametrów polega na takim dobraniu zestawu parametrów rozkładu, aby funkcja wiarygodności osiągnęła maksimum (Mańczak Dzendzelyuk, Kostiv, i Rabyk, 2013), (Brockwell i Davis, 2016).

### Kryteria oceny

W literaturze przedmiotu stosuje się następujące miary błędów (reszt) wyznaczonych modeli (prognoz): MSE (*ang. mean square error*), RMSE (*ang. root mean square error*), MAD (*ang. mean absolute deviation*), RMSPE (*ang. root mean square percentage error*), MAPE (*ang. mean absolute percentage error*) (Niederliński, 1979).

Niech  $\hat{x}_t$  oznacza prognozowaną wartość ciągu czasowego w chwili  $t$ , a  $x_t$  będzie zaobserwowaną wartością procesu chwili  $t$ . Formuły poszczególnych błędów są następujące:

- wartość średnia sumy kwadratów błędów prognozy dla N-observacji

$$MSE = \frac{1}{N} \sum_t (x_t - \hat{x}_t)^2 \quad (4.15)$$

- pierwiastek kwadratowy z wartości średniej sumy kwadratów błędów prognozy dla N-observacji :

$$RMSE = \sqrt{\frac{1}{N} \sum_t (x_t - \hat{x}_t)^2} = \sqrt{MSE} \quad (4.16)$$

- średni błąd bezwzględny dla N observacji:

$$MAD = \frac{1}{N} \sum_t |x_t - \hat{x}_t| \quad (4.17)$$

- pierwiastek z sumy kwadratów błędów względnych, wyrażony w procentach

$$RMSPE = \sqrt{\sum_t \left( \frac{x_t - \hat{x}_t}{x_t} \right)^2} \cdot 100\% \quad (4.18)$$

- średnia wartość bezwzględnych wartości błędów względnych dla N-observacji, wyrażona w procentach



$$MAPE = \frac{1}{N} \sum_T \left| \frac{x_t - \hat{x}_t}{x_t} \right| \cdot 100\% \quad (4.19)$$

Pierwsze trzy kryteria wyrażają błąd w jednostkach miary procesu i ich wartość zależy od specyfiki szeregów czasowych. Dwa ostatnie kryteria są mierzone w jednostkach względnych, więc możemy porównywać modele dla różnych procesów.

Dokładność prognozy dla błędów względnych bywa określana lingwistycznie (słownie) na przykład:

- a) jeśli stanowi mniej niż 10% - dokładność wysoka,
- b) od 10% do 20% prognoza jest dobra,
- c) między 20% i 40% średnia albo zadowalająca,
- d) wszystko co powyżej 40% - dokładność zła, niezadowalająca.

Podane wartości nie są stałe i zmieniają się w zależności od celu prognozowania. Wybieramy granice miary dokładności, które zadowolają oczekiwania.

### **Kryteria informacyjne dopasowania modeli (AIC i BIC)**

Dla wyboru najlepszego modelu regresji do opisu odpowiednich szeregów czasowych, korzystamy z kryteriów informacyjnych, AIC i BIC .

**AIC- Kryterium informacyjne Akaikego** (*ang. Akaike Information Criterion*) – kryterium wyboru pomiędzy modelami statystycznymi o różnej liczbie predyktorów. Jest to jeden ze wskaźników dopasowania modelu, który stosuje się do wyznaczenia miary jakości modeli statystycznych dla badanego zbioru danych. Na ogół model o większej liczbie predyktorów daje dokładniejsze przewidywania, ponadto ma większą skłonność do przeuczenia. AIC przedstawia wzór poniżej (Akaike, 1973):

$$AIC = \frac{2*k}{n} + \ln \left( \frac{1}{n} \sum_{i=1}^n (x_i - \hat{x}_i)^2 \right) , \quad (4.20)$$

gdzie:

- $x_i$  rzeczywista wartość ciągu,
- $\hat{x}_i$  wyjściowa wartość modelu,
- $n$  długość ciągu czasowego,
- $k$  liczba parametrów.

**BIC - Bayesowskie kryterium informacyjne Schwartz** (*ang. Bayesian Information Criterion*), w kategoriach sumy kwadratów reszt oblicza się za pomocą wzoru (Schwarz, 1978):

$$BIC = \frac{k*\ln n}{n} + \ln \left( \frac{1}{n} \sum_{i=1}^n (x_i - \hat{x}_i)^2 \right) , \quad (4.21)$$

gdzie:

- $x_i$  rzeczywista wartość ciągu,
- $\hat{x}_i$  wyjściowa wartość modelu,
- $n$  długość ciągu czasowego,
- $k$  liczba parametrów.

Te kryteria informacyjne są nazywane metodami porównywania modeli dla zmiennej zależnej, i stosowane są do wyboru najlepszego modelu. Zgodnie do przyjętych norm, najlepszym jest model dla którego wartości kryterium AIC oraz BIC są najniższe.

W statystyce kryterium informacyjne Bayesa (BIC) lub kryterium informacyjne Schwarza stanowią kryteria wyboru modelu spośród skończonego zbioru modeli; preferowany jest model z najniższym BIC. Opiera się on częściowo na funkcji prawdopodobieństwa i jest ściśle powiązany z kryterium informacyjnym Akaike (AIC) (Chen, J. i Chen, Z., 2008).

### 4.3. Funkcja autokorelacji reszt modelu

Podczas procesu modelowania i identyfikacji systemów dynamicznych, szczególną uwagę należy zwrócić na badanie reszt modeli. W przypadku obiektów statycznych, suma kwadratów błędów lub błąd średniokwadratowy stanowią wystarczające kryteria. Natomiast gdy jest testowany model ciągów czasowych wówczas funkcje autokorelacji reszt powinny być badane i porównywane z funkcjami autokorelacji typowych przykładów procesów stochastycznych, (Mańczak i Nahorski, 1983), (Kashpruk, 2016b).

Autokorelacja występuje wtedy, gdy skutki działania zmienności losowej nie wygasają w danym okresie  $t$ , lecz są przenoszone na okresy przyszłe.

Autokorelacja w modelu może być:

- a) dodatnia, gdy obok siebie występują składniki losowe takich samych znaków,
- b) ujemna, gdy obok siebie występują składniki losowe o różnych znakach.

Funkcję autokorelacji reszt modelu, obliczymy zgodnie z definicją unormowanej funkcji autokorelacji, (Mańczak i Nahorski, 1983):

$$k_e(s) = \frac{\sum_{n=m+1}^{N-s} (e_n - \bar{e})(e_{n+s} - \bar{e})}{\sqrt{\sum_{n=m+1}^{N-s} (e_n - \bar{e})^2} \sqrt{\sum_{n=m+1}^{N-s} (e_{n+s} - \bar{e})^2}}, \quad s=0,1,\dots,S \quad (4.22)$$

gdzie:

$e_n$  – reszta modelu,

$m$  – rząd modelu, np.  $m=2$ , dla AR(2),  $e_n = x_n - (ax_{n-1} + bx_{n-2})$ ,

$s$  – argument funkcji autokorelacji (krok),  $s = 0,1, \dots, S \leq N/3$ ,

$N$  – liczba obserwacji badanego ciągu  $x_n$ .

### Funkcja autokorelacji częściowej

Omówiliśmy wcześniej funkcję autokorelacji, która podaje korelacje między obserwacjami w bieżącym momencie a obserwacjami we wszystkich poprzednich punktach w czasie.

Funkcja częściowej autokorelacji (PACF) jak sama nazwa wskazuje, PACF jest podzbiorem funkcji autokorelacji (ACF). ACF i PACF stosowane są do wyboru rzędu modeli (Cory Maklin, 2019). PACF pokazuje zależność między  $x_n$  a poprzednimi wartościami, pomijając wpływ pośrednich opóźnień, jest przedstawiona wzorem:

$$x_n = \mu + \rho_1 x_{n-1} + \dots + \rho_s x_{n-s} + e_n, \quad (4.23)$$

PACF wyraża korelację między obserwacjami poczynionymi w dwóch punktach w czasie, uwzględniając wpływ innych punktów danych. Ta funkcja jest konstruowana przez obliczenie częściowej korelacji statystycznie korygując wpływ opóźnień pośrednich.

#### 4.4. Składowe trendu

Każdy ciąg czasowy często rozpatruje się jako sumę składowych: trendu, składowej sezonowej i szumu. Dla ujawnienia i analizy trendu wykorzystuje się aparat analizy regresji i średniej ruchomej. Dla analizy sezonowego efektu stosuje się modele sezonowego wygładzania i sezonowej autoregresji (Box i Jenkins, 1970), (Niederliński, Kasprzyk i Figwer, 1979).

Wahania trendu ujawniają się podczas analizy spektralnej, a dla opisu i prognozowania takich procesów wykorzystuje się modele harmoniczne albo modele autoregresji. W zadaniach prognozowania takich ciągów czasowych ważne jest założenie że tendencja, która zaznaczyła się w przeszłości, jest wyraźna i stosunkowo stabilna.

Wybór linii trendu nie zawsze jest prostym zadaniem. Ponadto, badany ciąg może jeszcze zawierać składowe sezonowe i cykliczne. Cykliczne składowe różnią się od sezonowych większą trwałością i nieciągłością amplitudy. Okresy składowej sezonowej mierzy się w dniach, tygodniach albo miesiącach, a cyklicznej – w latach albo dziesiątkach lat (np., cykle ekonomiczne) (Welfe, 2009). Do głównych metod analizy takich złożonych ciągów czasowych można zaliczyć metodę średniej ruchomej, wyrównywania wykładniczego i projektowania tendencji.

Dowolny ciąg czasowy można zapisać jako sumę składników: deterministycznego i przypadkowego:

$$x_n = f(u) + e_n, \quad (4.24)$$

gdzie:

$f(u)$  – linia trendu,

$e_n$  – składnik przypadkowy.

Dowolną obserwację ciągu czasowego możemy rozłożyć na trzy czynniki:

- a) trend ( $T$ ),
- b) sezonowość ( $S$ ),
- c) składnik losowy, czyli biały szum.

Budowa tych modeli sprowadza się do wyznaczenia składowych  $T$ ,  $S$  i  $e$  dla każdego szeregu. Dla modeli addytywnych składowe  $T$  się sumują:

$$X = T + S + e. \quad (4.25)$$

Dla modelu multiplikatywnego ciągu czasowego mamy zależność:

$$X = T \cdot S \cdot e. \quad (4.26)$$

Model multiplikatywny oznacza, że obserwowane wartości zmiennej prognozowanej są iloczynem składowych.

Metody służące do wyodrębniania tendencji rozwojowej, wahań okresowych i wahań przypadkowych z szeregów czasowych są określane mianem metod eliminacji w czasie lub metod dekompozycji (rozłożenia) szeregów czasowych (Welfe, 2009).

#### **4.5. Podejście Boxa-Jenkinsa do identyfikacji modeli na przykładzie modelu ARIMA**

Istnieje wiele metod i technik do przewidywania wartości zmiennej zależnej od czasu. Jednym z modeli branych pod uwagę jest ARIMA (ang. Autoregressive Integrated Moving Average). Przy badaniu takich modeli ciąg czasowy można podzielić na 3 składniki: trend (dane rosną lub maleją w długim okresie notowań), zmienność sezonowa (wzrost lub spadek w pewnych okresach lub momentach) oraz szum (zakłócenia w przypadkowych odstępach czasu).

Na ogół struktura modelowania szeregów czasowych ma następujące etapy (Box i Jenkins, 1983):

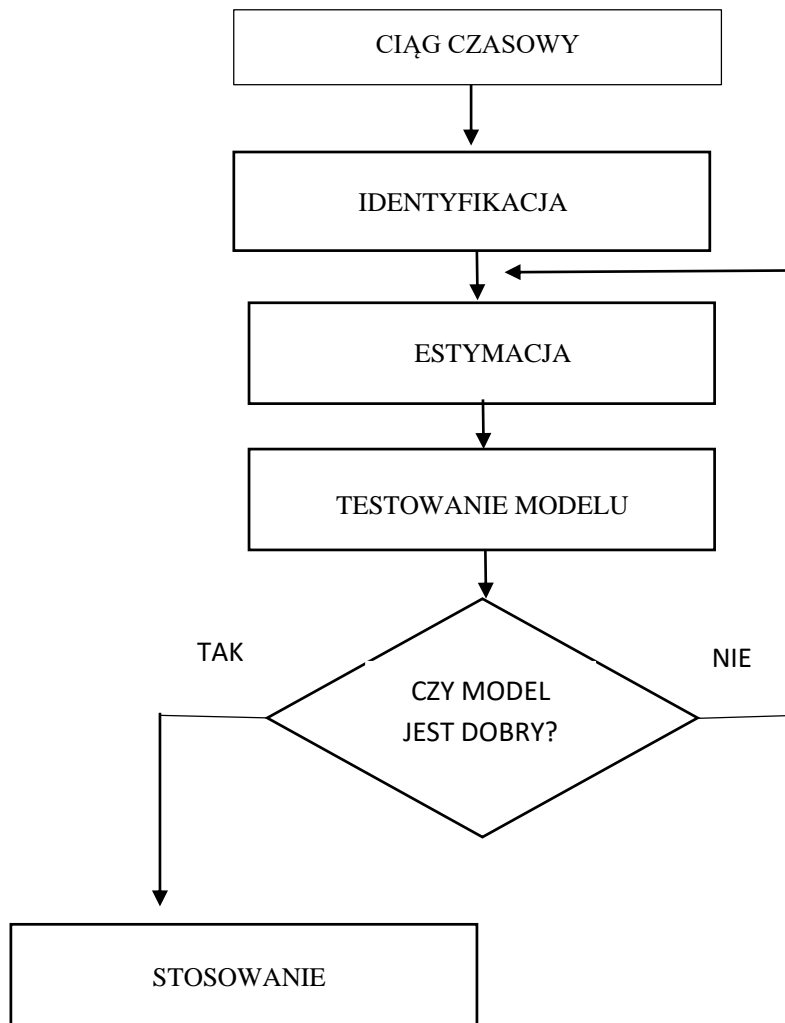
- testowanie stacjonarności i sezonowości,
- znajdowanie obiecujących propozycji modeli ,
- dopasowywanie modeli i zawężanie ich wyboru za pomocą AIC / BIC,
- wykonywanie testów diagnostycznych modeli,
- tworzenie prognoz.

Analiza Boxa-Jenkinsa odnosi się do systematycznej metody identyfikacji, dopasowywania, sprawdzania i wykorzystywania zintegrowanych modeli szeregów czasowych z autoregresją i średnią ruchomą (ARIMA). Metoda jest odpowiednia dla długich i średnich ciągów czasowych (co najmniej 50 obserwacji).

Skuteczną procedurę Boxa-Jenkinsa budowania empirycznych modeli ciągów czasowych można zrealizować w trzech etapach:

- specyfikacja lub identyfikacja modelu; wykorzystując wykresy danych, autokorelacje, częściowe autokorelacje i inne informacje, wybiera się klasę prostych modeli ARIMA, co sprowadza się do oszacowania odpowiednich wartości parametrów  $p$ ,  $d$  i  $q$ ;
- estymacja modeli;
- diagnostyka lub testowanie modeli (dopasowany model jest sprawdzany pod kątem niedoskonałości, biorąc pod uwagę autokorelację reszt modelu).

Powyższe kroki są stosowane iteracyjnie, dopóki krok trzeci nie przyniesie żadnej poprawy w modelu (Jerry Dr., Hintze L., 2007).



Rys. 4.1. Schemat Boxa-Jenkinsa do identyfikacji modeli

Przeanalizujemy krok po kroku schemat Boxa-Jenkinsa który pokazany na rys. 4.1.:

1) Najpierw sprawdzamy czy dane są stacjonarne za pomocą wykresu i testu Dickey-Fullera. Jeżeli dane nie są stacjonarne to wyliczamy zmiany procentowe kursu i ponownie robimy test.

2) Stosując funkcje ACF i PACF, wyznaczmy najmniejsze wartości  $p$  i  $q$ , które pokazują najniższe rzędy modeli AR i MA.

3) Następnym krokiem jest oszacowanie rzędu  $p$ ,  $q$ .

4) Dalej prowadzimy badania na resztach modeli i sprawdzamy jawność białego szumu. Jeżeli nie mamy oznak białego szumu, to dodajemy opóźnienia  $p$  lub  $q$  lub obydwie.

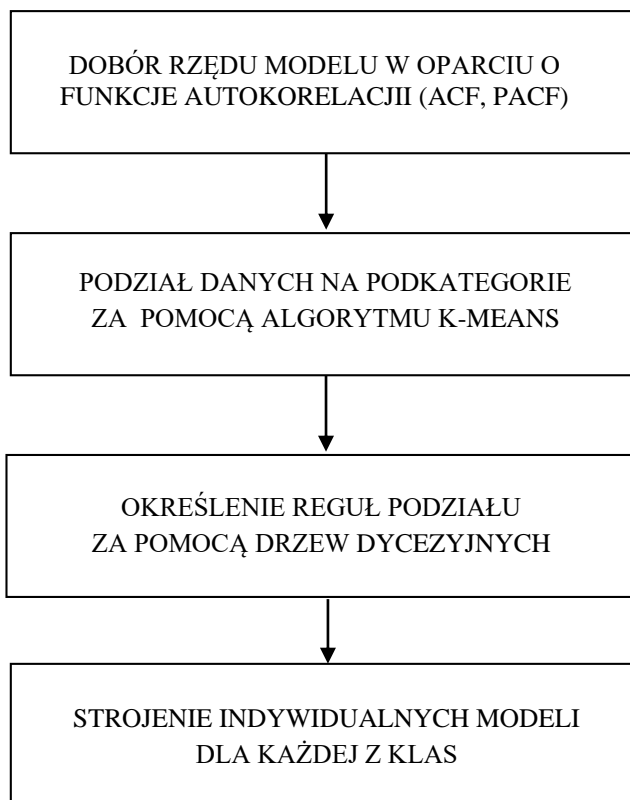
5) Dla uproszczenia wyboru możemy użyć kryteriów informacyjnych AIC, BIC (Hannan i Quinn, 1979).

## 4.6. Dobór modeli TAR

Metodę budowy modeli statystycznych dla ciągu jednowymiarowego  $\{x_n\}, n = 1, 2, \dots, N, x \in \chi$ , zaczyna się od przyjęcia klasy modeli, np. modeli autoregresji  $AR(k, A_k)$ ,

gdzie  $k$  oznacza rząd modelu liniowego z wektorem parametrów  $A_k = [a_0, a_1, \dots, a_k]$ . Estymatory  $\hat{A}_k$  parametrów modelu uzyskuje się na drodze obliczeń statystycznych z wykorzystaniem zbioru danych  $\{x_n\}$  i metody najmniejszych kwadratów.

Zbiór danych wejściowych  $\{x_n\}$  i przyjęta arbitralnie klasa modeli determinują dokładność modeli. Dla przykładu, ciąg danych o charakterze procesu białego szumu nie daje możliwości modelowania, ponieważ nie zawiera w sobie korelacji pomiędzy poszczególnymi wyrazami ciągu.



Rys. 4.2. Schemat doboru modeli TAR

W przypadku modeli progowych TAR, w których każdy model AR(k) jest obliczany dla pewnej części przestrzeni  $\chi$ , o dokładności modelu TAR decydują takie czynniki jak: zbiór danych wejściowych, trafność podziału przestrzeni  $\chi$  na podzbiory, w których liczone są poszczególne modele oraz wybór klas modeli.

Zarówno w modelach AR, jak i TAR decydujące znaczenie może mieć wiedza eksperta, w związku z wyborem klasy AR(k) oraz podziałem przestrzeni dla modeli TAR.

Dopasowanie modeli do danych doświadczalnych, a więc także uporządkowanie w klasach modeli wprowadzają kryteria identyfikacji reszt.

### Klasteryzacja $k$ -means

Klasteryzacja  $k$ -means ( $k$ -średnich) jest algorytmem grupowania obiektów. Grupowanie w algorytmie jest realizowane według kryterium minimum sumy kwadratów odległości między danymi a odpowiednimi środkami grupy (klastra).

Podstawowy algorytm  $k$ -średnich można przedstawić następująco:

1. Zakładamy liczbę grup równą  $k$  i wybieramy  $k$  początkowych środków tych grup, może być losowo spośród wszystkich punktów poszczególnych klastrów.
2. Wyznaczenie przynależności obiektów (punktów) do poszczególnych grup według przyjętej miary podobieństwa; każdy punkt do najbliższego klastra.
3. Wyznaczenie nowych środków grup jako średniego położenia danych przynależnych do danej grupy.
4. Sprawdzenie kryterium zatrzymania algorytmu. Jeśli kryterium podobieństwa nie zostanie spełnione, wtedy przechodzimy ponownie do drugiego kroku.

Ogólnym kryterium zbieżności jest minimalna różnica przy ponownym przypisaniu punktów do wyznaczonych centrów klastrów.

Głównymi zaletami algorytmu  $k$ -średnich są jego prostota i szybkość w obliczeniach, co pozwala do obliczeń na dużych zbiorach danych. W rezultacie grupowania  $k$ -średnich każdy punkt należy do tylko jednego zbioru. W klastrach rozmytych  $k$ -średnich, tak jak w logice rozmytej, każdy punkt może być przypisany do wielu klastrów jednocześnie, ma pewien stopień przynależności do pewnych zbiorów (Rutkowski, L., 2020), (Talar, J., 2007).

## Drzewa decyzyjne

Drzewa decyzyjne (*ang. decision trees*) są ważnym narzędziem do klasyfikacji danych. Struktura jest złożona z gałęzi oraz węzłów, spośród których można wyróżnić korzeń i liście. Drzewa decyzyjne to acykliczne grafy skierowane, w których wierzchołkami są węzły, a powiązania między węzłami to gałęzie. Węzły odpowiadają testom, które przeprowadzane są na wartościach atrybutów warunkowych, krawędzie odpowiadają wynikom testów, a liście - etykietom kategorii. Wierzchołki bez potomków to liście, a wierzchołek bez rodzica to korzeń. W celu zaklasyfikowania atrybutu w drzewie, jest on przesuwany odgórnie od węzła głównego (korzenia) w kierunku węzłów liści poprzez węzły wybierane zgodnie z wynikami testów reprezentowanymi przez węzły wewnętrzne, dopóki nie zostanie osiągnięty ostatni węzeł zwany liściem. W tym momencie, etykieta związana z węzłem liścia jest etykietą klasy przewidzianą dla atrybutu (Probiez Barbara, 2017).

Metoda drzew decyzyjnych należy do metod klasyfikacji, i pozwala na przedstawienie wiedzy w prostej i zrozumiałej formie. Celem stosowania drzewa decyzyjnego jest znalezienie zbioru reguł logicznych. Podstawowe kryterium, to maksymalizacja zdobywania informacji (lub minimalizacja heterogeniczności w węzle). Istnieje wiele różnych kryteriów stosowanych do oceny zdobywania informacji dla różnych algorytmów drzew decyzyjnych.

Podejście do automatycznego konstruowania drzew decyzyjnych opiera się na metodzie „dziel i zwyciężaj”, która polega na iteracyjnej procedurze odgórnego wyboru najlepszego atrybutu do znakowania wewnętrznych węzłów drzewa. W pierwszym kroku wybierany jest atrybut do reprezentowania korzenia drzewa. Po wybraniu pierwszego atrybutu tworzona jest gałąź dla każdej możliwej wartości atrybutu. Zestaw danych dzielony jest na podzbiory według wartości wybranego atrybutu. Procedura selekcji jest następnie stosowana rekurencyjnie do każdego kanału za pomocą węzła odpowiedniego podzbiory atrybutów, które mają podobne

wartości. Obiekty z podzbioru mają tę samą etykietę klasy lub gdy inne kryterium zatrzymania jest spełnione, tworzą węzeł liści do reprezentowania etykiety klasy, którą należy przewidzieć (Probiez Barbara, 2017).

Algorytm CART (*ang. Classification and Regression Trees*) to algorytm stosowany do konstruowania drzew decyzyjnych i jest bardzo popularną i przydatną metodą prezentowania wiedzy. CART to metoda, która generuje podzbiór zbioru danych na dwa podzbiory według kryterium minimalizacji heterogeniczności. Każdy podział oparty jest na jednej zmiennej. Niektóre zmienne mogą być użyte kilka razy, a inne w ogóle (Nowak-Brzezińska, 2012).

## 4.7. Facebook Prophet

Prognozowanie ciągów czasowych jest bardzo popularne, i najczęściej stosowane są modele autoregresji. Otrzymanie wiarygodnych prognoz nie jest takie proste, ponieważ wymaga specjalisty oraz dostępności odpowiedniego oprogramowania.

Powołując się na artykuł ekspertów Facebooka, przedstawimy pakiet prognozowania ciągów czasowych – Prophet, który pozwala na tworzenie dokładnych modeli predykcyjnych w trybie półautomatycznym. Bardziej szczegółowy opis pakietu Prophet i jego zastosowanie został przedstawiony w artykule Taylor i Letham (2017). Podstawą w tym pakiecie jest procedura dopasowania modeli regresji addytywnej (*Generalized Additive Models*, GAM):

$$y(t) = g(t) + s(t) + h(t) + \varepsilon_t, \quad (4.27)$$

gdzie

$g(t)$  – funkcja, która przedstawia trend ciągu,

$s(t)$  - wahania sezonowe (roczne, tygodniowe itp.),

$h(t)$  - funkcja, która przedstawia skutki świąt i innych wydarzeń,

$\varepsilon_t$  - błąd modeli, zawierający różne informacje i zaburzenia które nie są przedstawione w modelu.

Estymację parametrów dopasowanego modelu przeprowadza się z wykorzystaniem zasad statystyki bayesowskiej (metodą znajdowania późniejszego maksimum (MAP)).

W tym celu wykorzystuje się bibliotekę dla Pythona – Facebook Prophet ([facebook.github.io/prophet/](https://facebook.github.io/prophet/)).

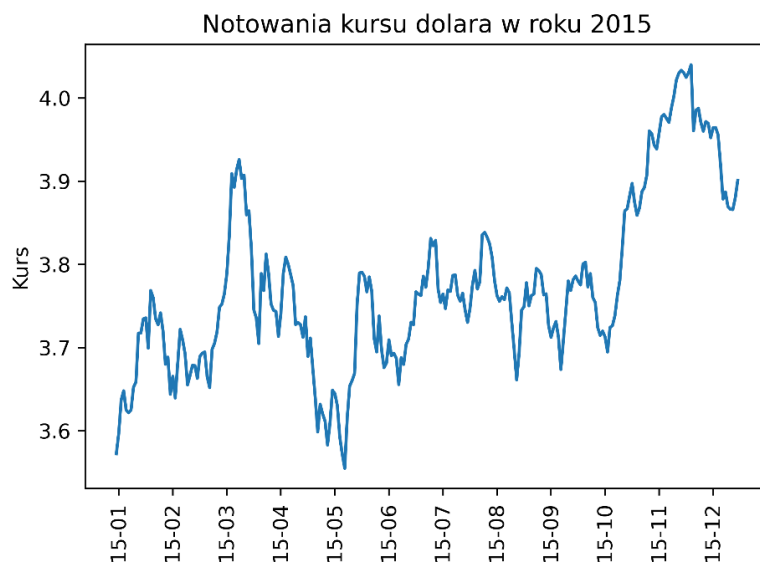
Model Prophet ciągów czasowych jest zaprojektowany do prognozowania biznesowych ciągów czasowych. Model ma intuicyjne parametry, co pozwala na dostosowanie bez znajomości szczegółów podstawowego modelu (Taylor i Letham, 2018).

## 4.8. Identyfikacja modeli na przykładzie modelu ARIMA (podejście Boxa-Jenkinsa )

Dla przykłady przedstawione zostaną badania dziennych notowań walut USD/PLN. Dane w liczbie 250 dotyczą roku 2015 ([money.pl/pieniadze/nbparch/srednie/?symbol=USD](https://money.pl/pieniadze/nbparch/srednie/?symbol=USD)).

Przebieg notowań kursu dolara do złotego jest przedstawiony na rysunku 4.3.





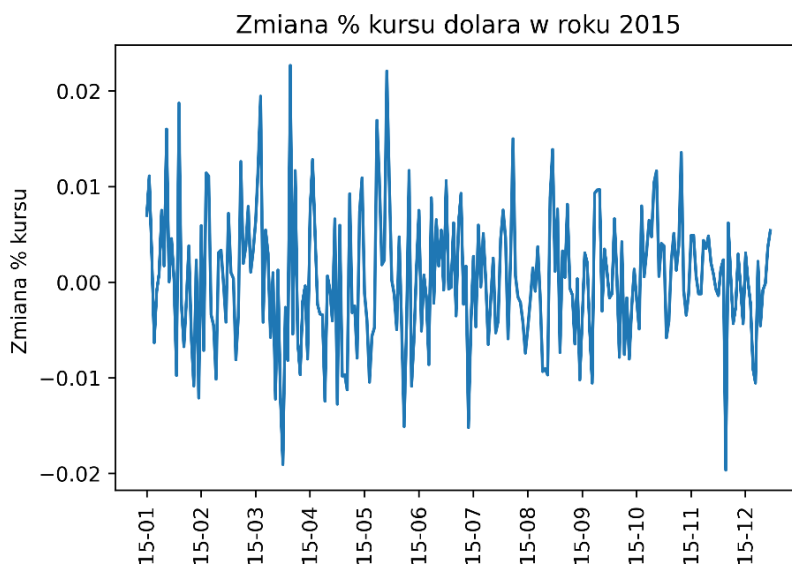
Rys. 4.3. Przebieg notowań kursu dolara w 2015 roku

Dla sprawdzenia stacjonarności ciągu robimy analizę za pomocą rozszerzonego testu Dickeya Fullera (ADF). Przy analizie testu ADF warto zwrócić uwagę, że hipoteza zerowa wskazuje na istnienie niestacjonarności. Jeżeli obliczona wartość  $p$  jest mniejsza od zadanej alfa, to ta hipoteza zerowa zostaje odrzucona, a ciąg jest stacjonarny. W przeciwnym razie dane są niestacjonarne.

Tab. 4.1. Wyniki testu ADF dla ciągu danych notowań zmiany kursu USD

ADF Statistic:	-2.357800
p-value:	0.154003

Wynika z tego że model jest białym szumem i jest stacjonarny. Po czym za pomocą funkcji `pct_change()` obliczamy procentową zmianę między bieżącym a poprzednim elementem notowania kursu USD. Otrzymany ciąg danych jest przedstawiony na rysunku 4.4.



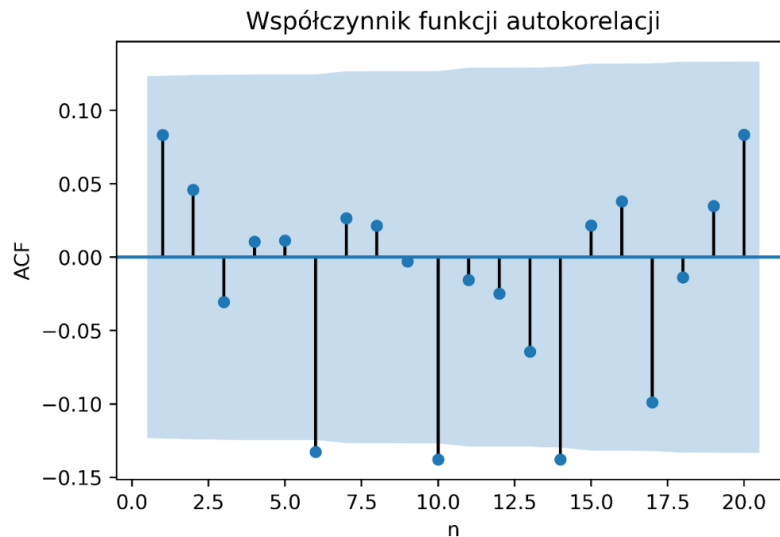
Rys. 4.4. Zmiana procentowa notowań kursu USD w 2015 roku.

Odpowiednie jak dla ciągu notowań kursu USD obliczamy test ADF dla danych zmiany procentowej kursu USD. Wyniki testu dla % zmiany kursu USD przedstawiono w tabeli 4.2.

Tab. 4.2. Wyniki testu ADF dla ciągu % zmiany kursu USD

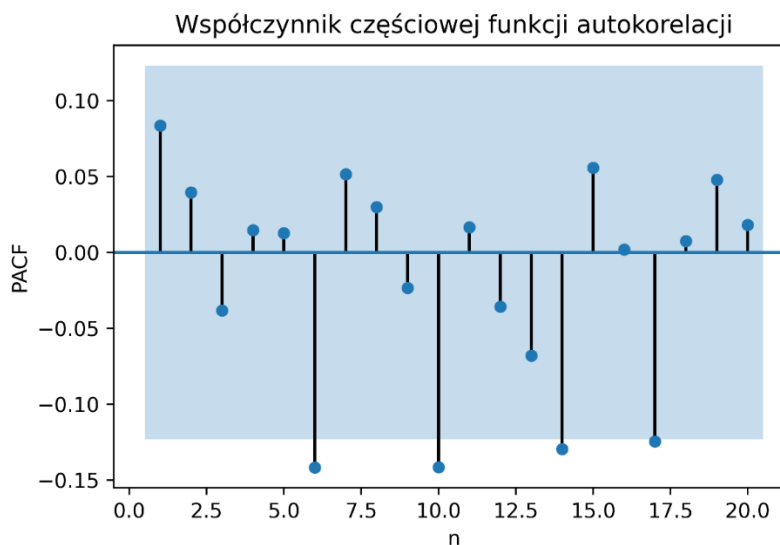
ADF Statistic:	-14.555140
p-value:	0.000000

Jak widać z wyników testu ADF nasz model nie jest stacjonarny, możemy kontynuować badania i wyliczamy funkcje autokorelacji oraz częściowej autokorelacji.



Rys. 4.5. Wykres funkcji autokorelacji dla danych % zmiany kursu USD

Funkcja autokorelacji (ACF), podaje korelacje między obserwacjami w bieżącym momencie a obserwacjami we wszystkich poprzednich punktach w czasie (rysunek 4.5.).



Rys. 4.6. Wykres częściowej funkcji autokorelacji dla danych % zmiany kursu USD

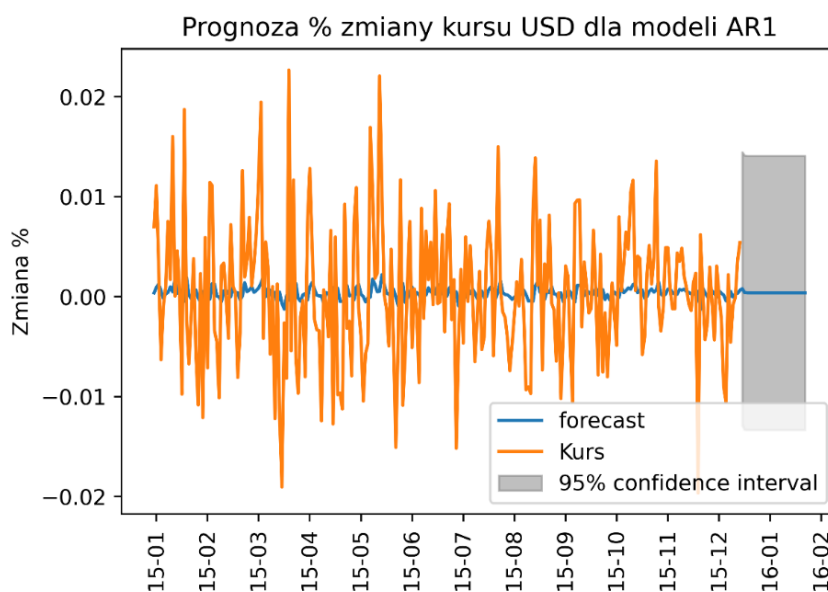
Funkcja częściowej autokorelacji (PACF), wyraża korelację między obserwacjami poczynionymi w dwóch punktach w czasie, uwzględniając wpływ innych punktów danych. Ta funkcja jest obliczana przez obliczenie częściowej korelacji, statystycznie korygując wpływ opóźnień pośrednich. Funkcję częściowej autokorelacji dla danych % zmiany kursu USD przedstawiono na rysunku 4.6.

W kolejnym kroku wyznaczamy za pomocą funkcji *auto.arima* najlepsze parametry modelu dla prognozowania. Wyliczono kryteria informacyjne AIC, BIC, HQIC odpowiednio, do których został wybrany najlepszy model, ARIMA(3,1,1), oraz wyznaczono parametry modeli (tab. 4.3.)

Tab. 4.3. Parametry modelu ARIMA(3,1,1)

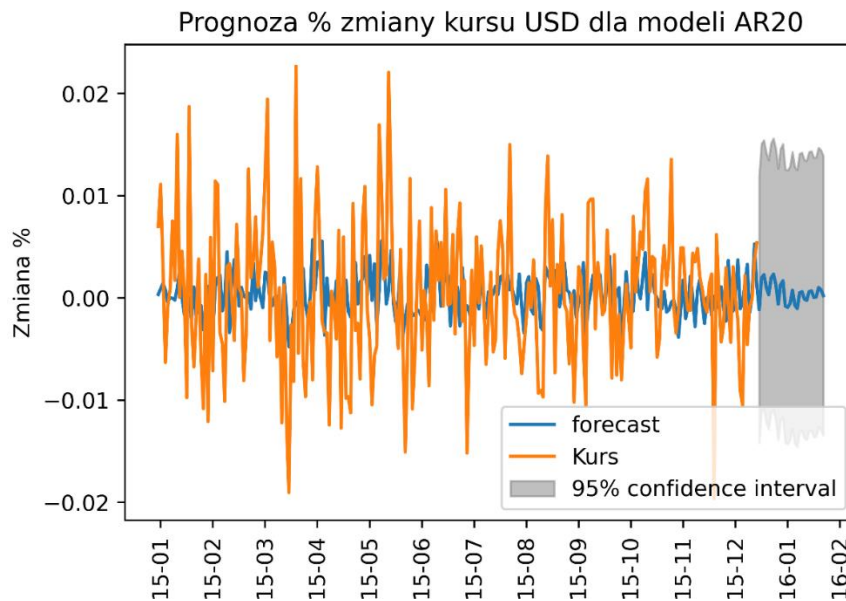
Model:	ARIMA(3,1,1)
Log Likelihood	883.735
AIC	-1755.470
BIC	-1734.293
HQIC	-1746.949
intercept	$-8.522e - 06 \pm 4.33e - 05$
AR.L1	$0.1356 \pm 0.072$
AR.L2	$-0.0995 \pm 0.068$
AR.L3	$-0.0622 \pm 0.069$
MA.L1	$-0.9147 \pm 0.042$
sigma2	$0.0025 \pm 0.000$

Funkcja autokorelacji wskazuje, że może to być proces stochastyczny o długiej pamięci. Jest to proces stacjonarny w szerszym sensie, gdyż funkcja autokorelacji maleje do zera (rysunek 4.5.). Współczynniki ACF i PACF wskazują na liniową statystyczną zależność co potwierdza celowość wprowadzenia liniowego modelu AR(1) oraz dla porównania weźmiemy model AR(20).



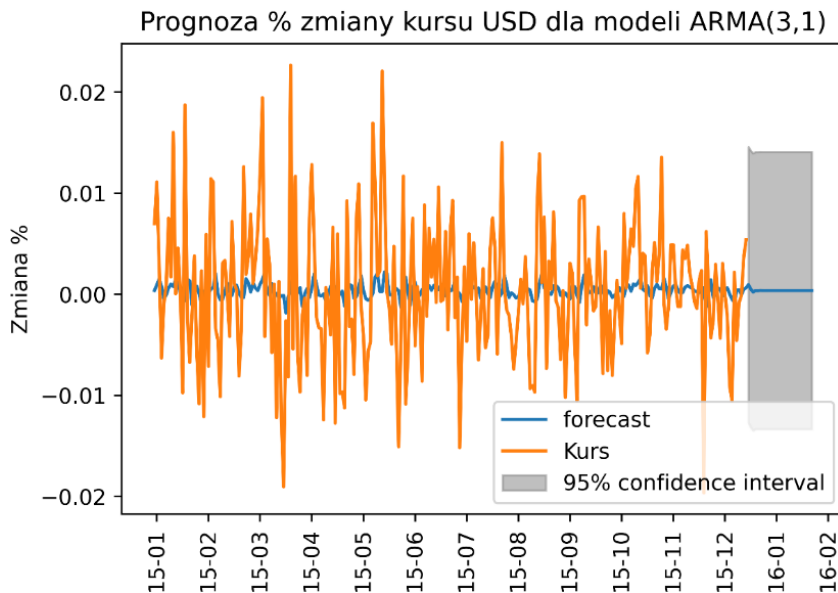
Rys. 4.7. Wykres % zmian kursu USD i prognozy wyznaczonej za pomocą modeli AR(1)

Prognoza, jak pokazano na rysunkach 4.7 i 4.8, jest mocno uśredniona. Ponadto prognozowane dane znajdą się w zakresach  $[-0.014; 0.014]$  dla modeli AR(1) oraz  $[-0.0125; 0.0125]$  dla modelu AR(20), z wiarygodnością 95%.



Rys. 4.8. Wykres % zmian kursu USD i prognozy wyznaczonej za pomocą modeli AR(20)

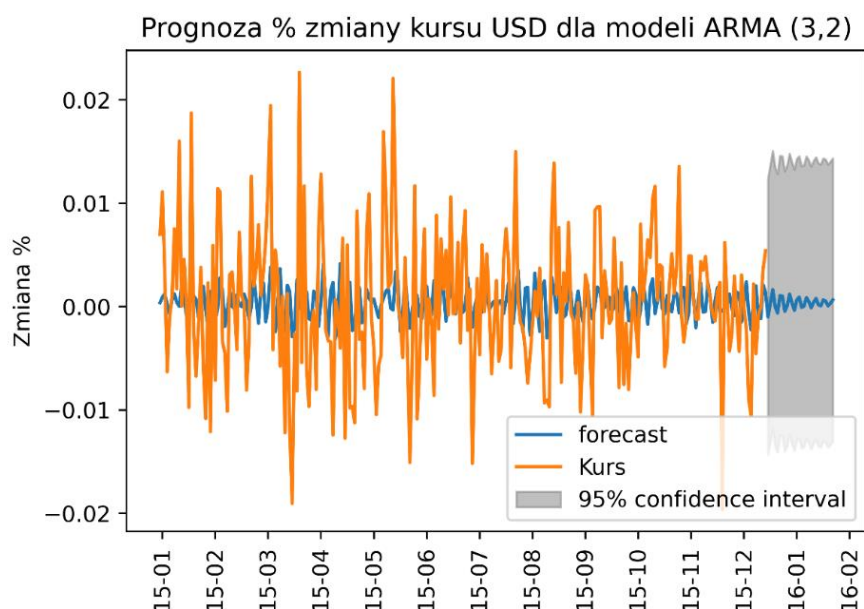
Dla porównania pokażemy modele ARMA(3,1) i ARMA(3,2) dla danych procentowej zmiany kursu USD. Przebieg danych wraz z prognozami pokazano na rysunkach 4.9. i 4.10. Ponadto obliczono ich parametry (tab4.4.).



Rys. 4.9. Wykres % zmian kursu USD i prognozy dla modeli ARMA(3,1)

Tab. 4.4. Porównanie parametry modeli ARMA(3,1) i ARMA(3,2)

Model:	ARMA(3,1)	ARMA(3,2)
Log Likelihood	897.709	902.696
AIC	-1783.417	-1791.393
BIC	-1762.217	-1766.659
HQIC	-1774.888	-1781.442
const	0.0002 ± 0.000	0.0004 ± 0.000
AR.L1	-0.361 ± 0.781	-0.3346 ± 0.088
AR.L2	0.0524 ± 0.088	0.8833 ± 0.046
AR.L3	-0.0335 ± 0.072	0.0087 ± 0.073
MA.L1	0.1180 ± 0.780	0.4275 ± 0.065
MA.L2	-	0.9662 ± 0.052



Rys. 4.10. Wykres % zmian kursu USD i prognozy wyznaczonej za pomocą modeli ARMA(3,2)

#### 4.9. Opis danych stanowiących wybrane parametry węgla energetycznego

Do analizy posłużą dane opisujące własności węgla surowego w próbkach pobieranych w zakładzie przeróbki mechanicznej węgla. Węgiel jest materiałem uziarnionym i niejednorodnym, w tym sensie, że jego ziarna mają różne rozmiary i różny skład chemiczny. Dla celów energetycznych ważna jest zawartość substancji organicznej, palnej oraz zawartość skały, jako substancji nieużytecznej. Jako paliwo węgiel musi być oczyszczony ze skały.

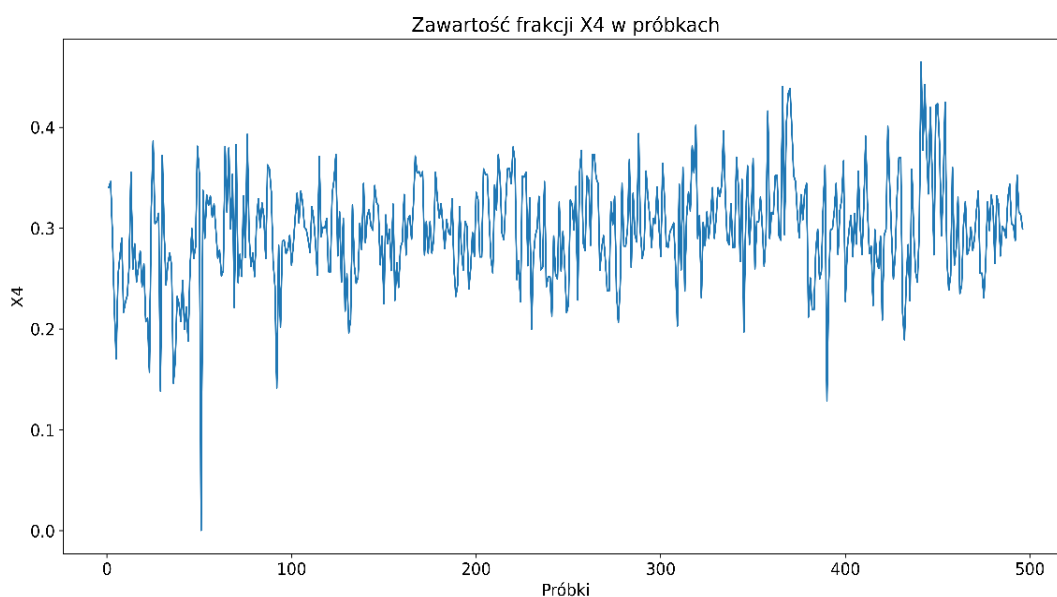
Dane do modelowania pochodzą z próbek pobieranych w zakładzie przeróbki węgla. Pobierano próbki z materiału, który płynie na przenośnikach, w stałych odstępach czasu co 4 minuty. Ziarna węgla mają średnice poniżej 20 mm, co ogólnie jest określane jako węgiel drobny. Na ogół pomiar parametrów węgla nie jest sprawą łatwą, ze względu na niejednorodność materiału oraz czynnik losowy wpływający na wybór próbki poddawanej pomiarowi. Próbkowanie jest czynnikiem, który generuje największe błędy pomiarów (Walaszek-Babiszewska A. (2004)).



Rys.4.11. Przykładowy podajnik węgla z kopalni Dobropolska, obwód Doniecki, Ukraina.  
(Źródło: <https://metallurgprom.org/uk>)

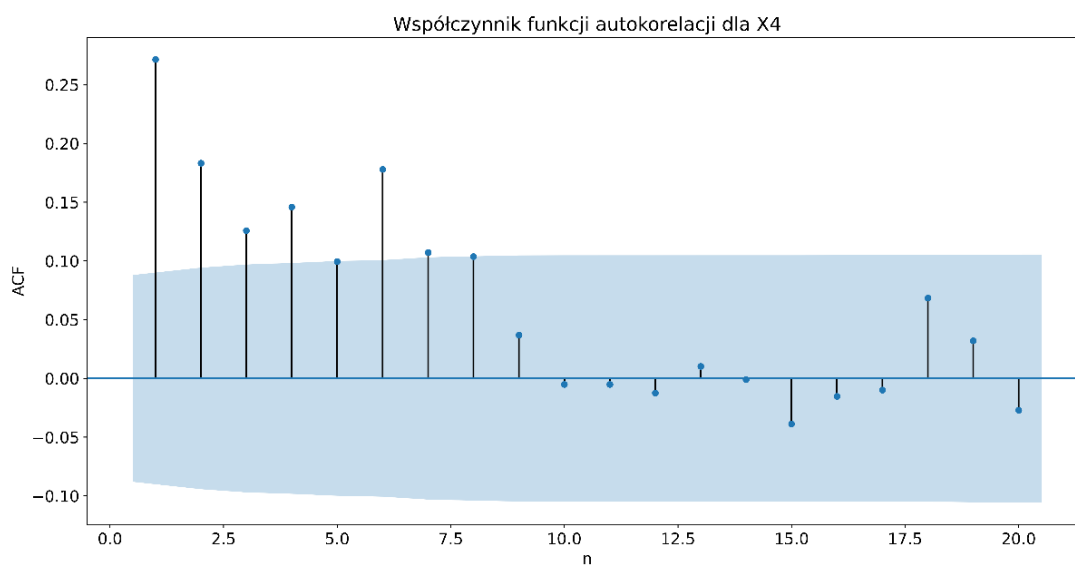
W laboratorium, przy spaleniu próbki węgla uwalnia się energia, która zależy od zawartości części palnej w węglu. Ta frakcja próbki ma najmniejszą gęstość, jest najlżejsza. Jako produkt przeróbki jest to koncentrat. Masa popiołu pozostała po spaleniu próbki wskazuje na udział substancji niepalnej w węglu, tym samym na obciążenie środowiska. Dane dotyczące jakości węgla opisuje 8 zmiennych. Udziały czterech frakcji gęstościowych są wpisane w kolumnach X1-X4 (dane sumują się do jedności), gdzie X1 oznacza zawartość najlepszej frakcji w próbce węgla oraz X4 - zawartość frakcji węgla z największą ilością substancji niepalnej. Dla każdej frakcji ziaren węgla, po spaleniu otrzymujemy substancję niepalną (popiół), której zawartość w danej frakcji wyrażono w procentach. Zawartości substancji niepalnej umieszczono w kolumnach X5-X8. Przy spaleniu X1 otrzymujemy X5, itd. Liczba danych w każdej kolumnie jest równa 496 (Walaszek-Babiszewska A., Czabak A., Błaszczak K., 2008).

Wartości udziału frakcji ciężkiej, X4, w próbkach węgla, czyli ziaren o najniższej zawartości substancji palnej, przedstawia rysunek 4.12. Przedstawiony ciąg wartości cechuje się dużymi i szybkimi zmianami w czasie.



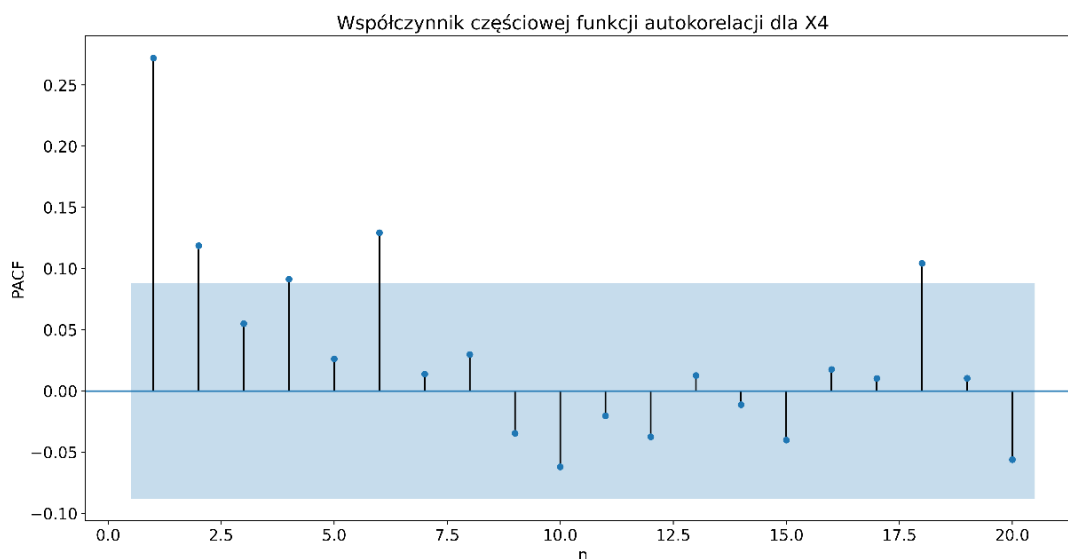
Rys. 4.12. Wykres zmian w czasie zawartości frakcji ciężkiej X4 w próbkach węgla

Funkcję autokorelacji przebiegu X4 przedstawiono na rysunku 4.13. Kształt funkcji autokorelacji wskazuje na brak znaczących zależności statystycznych.



Rys .4.13. Wykres funkcji autokorelacji dla danych X4

Funkcję autokorelacji częściowej przebiegu X4 przedstawiono na rysunku 4.14. Kształt funkcji autokorelacji potwierdza również brak znaczących zależności statystycznych.



Rys. 4.14. Wykres częściowej funkcji autokorelacji dla danych X4

Tak silna losowość zmian wynika z istoty badań próbkowych niejednorodnego materiału ziarnionego, co wykazano w pracy [Walaszek-Babiszewska, 1992].

Następnie, dla sprawdzenia ciągu na stacjonarność przeprowadzono rozszerzony test Dickey-Fullera, przedstawione w tabeli 4.5.

Tab. 4.5. Wyniki testu ADF dla danych X4

ADF Statistic:	-6.256060;
p-value:	0.000000

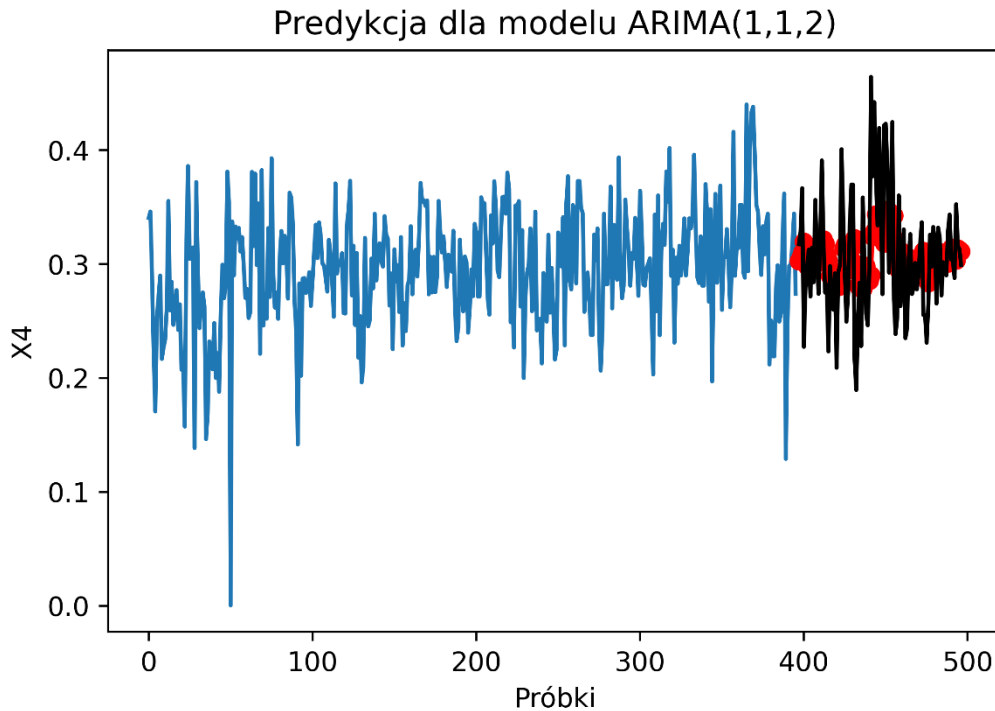
Odpowiednio jak dla danych dolara, wyznaczono za pomocą funkcji *auto.arima* najlepsze parametry modelu dla prognozowania. Wyliczono kryteria informacyjne AIC, BIC, HQIC odpowiednio do których został wybrany najlepszy model, ARIMA(3,1,1), oraz wyznaczono parametry modeli (tab. 4.6.).

Tab. 4.6. Kryteria doboru lepszego modelu dla prognozowania danych X4

Best model:	ARIMA(1,1,2)
Log Likelihood	780.753
AIC	-1553.506
BIS	-1536.687
HQIC	-1546.903
AR.L1	0.7550 ± 0.088
MA.L1	-1.5568 ± 0.104
MA.L2	0.5618 ± 0.101
SIGMQ	0.0025 ± 0.000
Mean squared error:	0.002582014240362711
SMAPE:	12.73645478699387



Wykres propozycji najlepszego modelu dla danych X4 odpowiednio do wyznaczonych parametrów przedstawiono na rysunku 4.15.



Rys. 4.15. Prognoza danych X4 na podstawie ciągu uczącego dla modelu ARIMA

#### 4.10. Model ARX i identyfikacja dla ciągu obserwacji procesu technologicznego

Proces ARX jest modelem liniowym obiektu dynamicznego z czasem dyskretnym w którym  $x$  oznacza zmienną wejściową,  $y$  oznacza zmienną wyjściową.

Model może mieć następującą postać:

$$y_n = ay_{n-1} + bx_{n-2} + \varepsilon_n, \quad (4.28)$$

gdzie:

- zmienna  $y$  jest modelowana jako AR(1),
- zmienna wejściowa jest opóźniona o 2 kroki,
- $\varepsilon$  reprezentuje proces bezzwzględnie losowy,
- parametry modelu  $a$ ,  $b$  można wyznaczyć metodą najmniejszych kwadratów, w oparciu o zbiór danych  $\{(x_n, y_n)\}$ ,  $n=1,2,\dots,N$ , w oparciu o zbiór danych ciągu obserwacji procesu technologicznego.

Jako dane wejściowe  $x_n$  przyjęto frakcje X8, zawartość popiołu dla frakcji X4 w każdej próbkce. Jako dane wyjściowe przyjęto udział frakcji X4 w każdej próbkce. Dane wejściowe i wyjściowe przedstawiono na rysunku 4.16. i 4.12. Przyjmuje się kryterium identyfikacji (Niederliński, 1979) ,(Mańczak i Nahorski, 1983):

$$Q = \sum_{n=3}^N [y_n - (ay_{n-1} + bx_{n-2})]^2 \rightarrow \min. \quad (4.29)$$

Rozwiązania, w postaci estymatorów  $\hat{a}$   $\hat{b}$  poszukuje się z warunku koniecznego istnienia ekstremum funkcji (4.28.).

$$\begin{cases} \frac{\partial Q}{\partial a} = 0, & \frac{\partial Q}{\partial a} = \sum_{n=3}^N \{2[y_n - (ay_{n-1} + bx_{n-2})](-y_{n-1})\} \\ \frac{\partial Q}{\partial b} = 0, & \frac{\partial Q}{\partial b} = \sum_{n=3}^N \{2[y_n - (ay_{n-1} + bx_{n-2})](-x_{n-2})\} \end{cases}$$

$$\sum_{n=3}^N [-y_n y_{n-1} + ay_{n-1}^2 + bx_{n-2} y_{n-1}] = 0$$

$$\sum_{n=3}^N [-y_n x_{n-2} + ay_{n-1} x_{n-2} + bx_{n-2}^2] = 0$$

$$\begin{cases} a \sum_n y_{n-1}^2 + b \sum_n x_{n-2} y_{n-1} = \sum_n y_n y_{n-1} \\ a \sum_n y_{n-1} x_{n-2} + b \sum_n x_{n-2}^2 = \sum_n y_n x_{n-2} \end{cases}$$

Powyższy układ równań rozwiązujemy korzystając z arkusza kalkulacyjnego. Podstawiając A,B,C,D,E, zamiast wyrażeń jak poniżej.

$$A = \sum_n y_{n-1}^2$$

$$B = \sum_n y_{n-1} x_{n-2}$$

$$C = \sum_n y_n y_{n-1}$$

$$D = \sum_n y_n x_{n-2}$$

$$E = \sum_n x_{n-2}^2$$

Otrzymujemy układ równań algebraicznych ze względu na  $a$  i  $b$

$$\begin{cases} \hat{a}A + \hat{b}B = C \\ \hat{a}B + \hat{b}E = D \end{cases}$$

Wyznacznik główny układu równań:

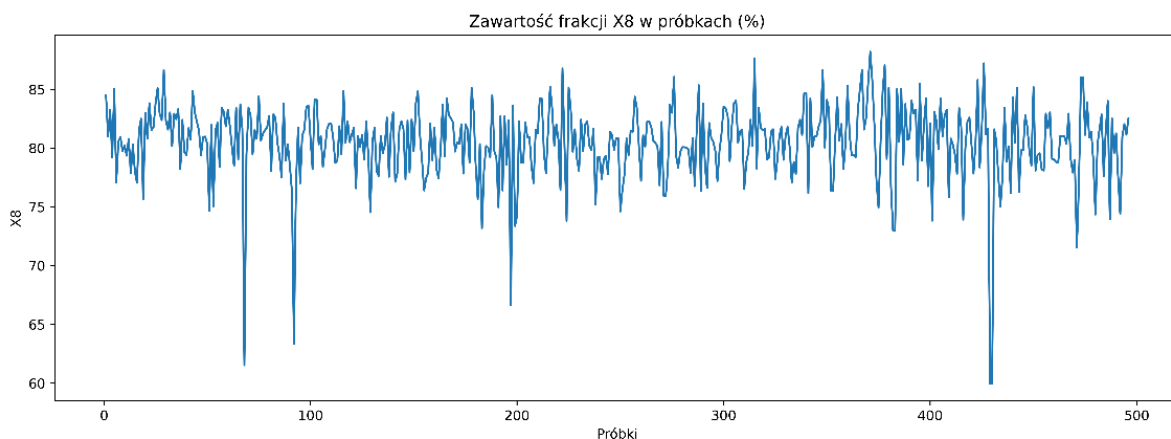
$$W = \begin{vmatrix} C & A \\ A & B \end{vmatrix} = AE - B^2, W \neq 0,$$

$$W_{\hat{a}} = \begin{vmatrix} C & B \\ D & E \end{vmatrix} = CE - DB,$$

$$W_{\hat{b}} = \begin{vmatrix} A & C \\ B & D \end{vmatrix} = AD - CB,$$

$$\hat{a} = \frac{W_{\hat{a}}}{W}, \quad \hat{b} = \frac{W_{\hat{b}}}{W}.$$

ARX to model autoregresji z zewnętrznym wejściem. Modele ARX stosuje się do opisu systemów, gdzie  $x$  jest zmienną wejściową systemu,  $y$  oznacza zmienną wyjściową systemu. Rząd modelu ARX jest określony przez liczby kroków wejścia i wyjścia branych pod uwagę oraz przez opóźnienie.



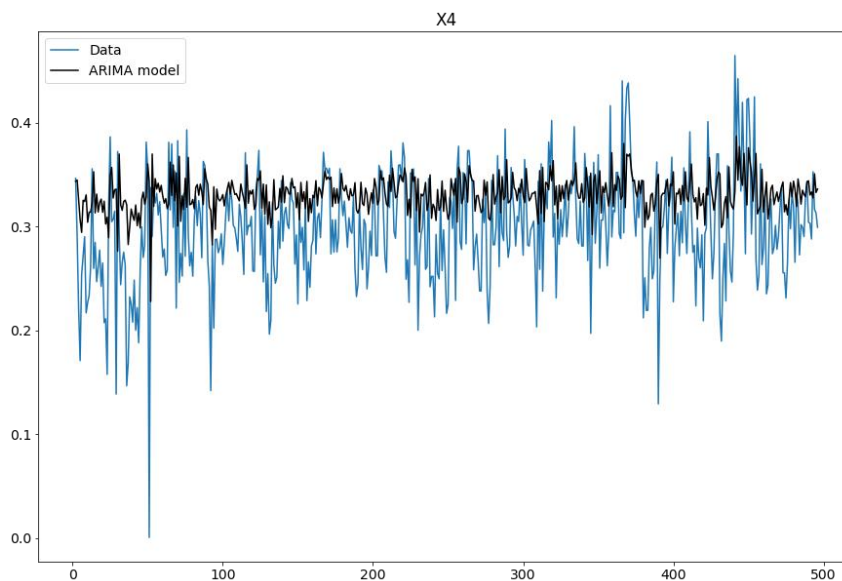
Rys. 4.16. Wykres zawartości popiołu w poszczególnych próbkach dla X4

Odpowiednio jak w poprzednich przykładach, za pomocą analizy kryteriów informacyjnych przedstawionych w tabeli 4.7., dobrano najlepszy model dla przewidywania zawartości popiołu w frakcji X4, ARIMAX (1,0,1).

Tab. 4.7. Kryteria informacyjne dla modelu ARIMAX (1,0,1)

<b>Model:</b>	<b>ARIMAX (1,0,1)</b>
Method	MLE
Log Likelihood	675.0903
AIC	-1338.180618076955
BIC	-1312.9532715015428

Wykres prognozy przedstawiono na rysunku 4.17.



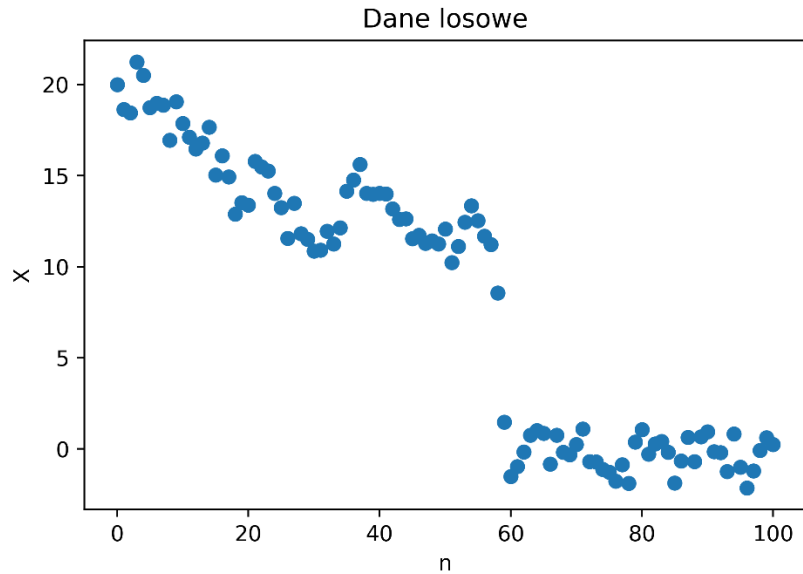
Rys. 4.17. Prognoza X4 dla modelu ARIMAX (1,0,1).

## 4.11. Prognoza przy stosowaniu modelu TAR

Dla przykładu wygenerowano 100 obserwacji z następującego modelu:

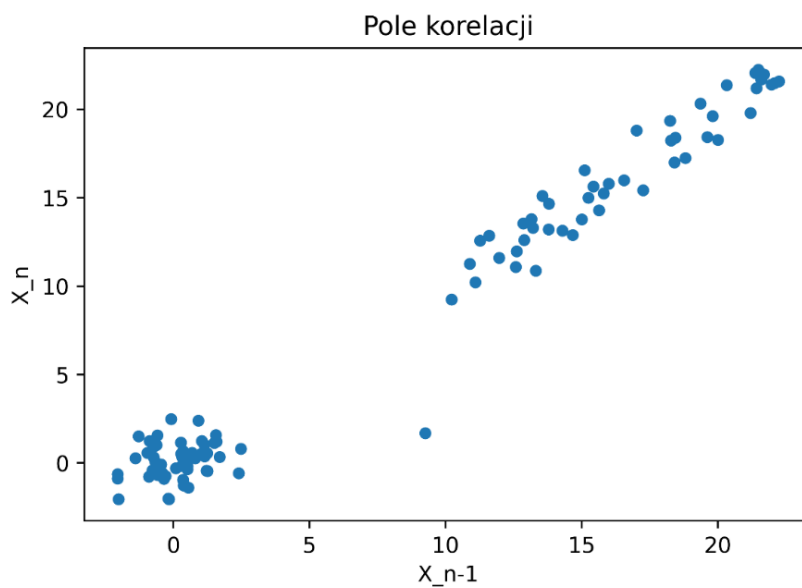
$$x[k + 1] = \begin{cases} 0.3 x[k] + \varepsilon[k] & \text{dla } x < 10 \\ 0.99 x[k] + \varepsilon[k] & \text{dla } x > 10 \end{cases} \quad (4.30)$$

Wykres wygenerowanych danych przedstawiono na rys.4.18.



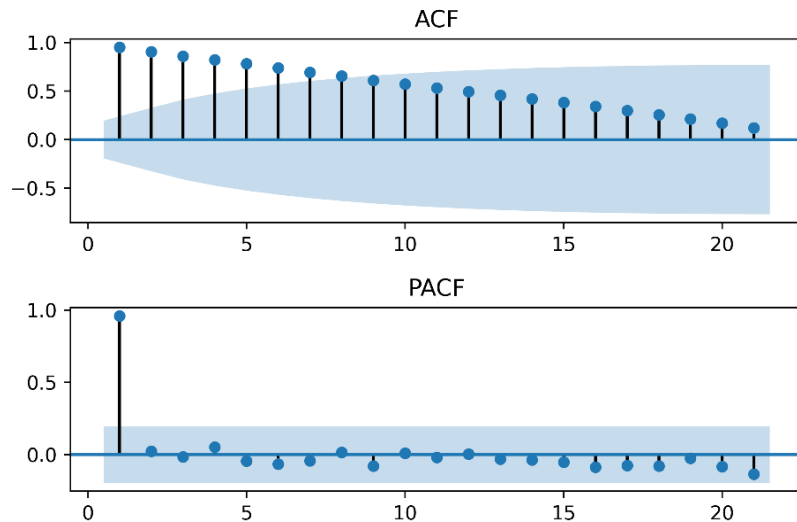
Rys. 4.18. Wykres losowo-wygenerowanych danych wejściowych

Przy analizie pola korelacji przedstawionego na rys.4.19. widoczna jest korelacja pomiędzy danymi.



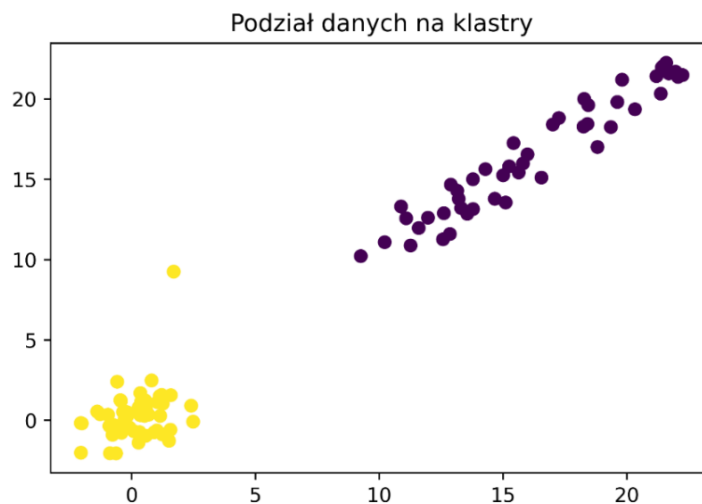
Rys. 4.19. Pole korelacji danych wejściowych

Stacjonarność procesu w szerszym sensie oznacza stałą w czasie wartość średnią, stałą wariancję i funkcję autokorelacji zależną tylko od przesunięcia czasowego (kroku). Jeśli szereg nie spełnia tych wymagań, to zazwyczaj stosuje się różnicowanie aż do osiągnięcia stacjonarności. Przebieg badanego ciągu na rys.4.18 i funkcje ACF i PACF pokazane na rys.4.20 wskazują, że proces nie jest stacjonarny. Dla tego zbioru danych zostanie zastosowany algorytm klasteryzacji k-means.



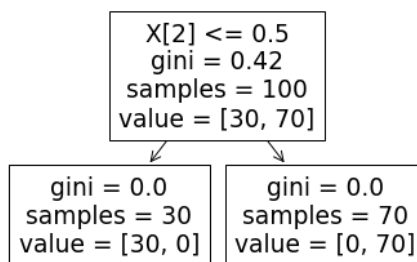
Rys. 4.20. Wykresy funkcji autokorelacji oraz autokorelacji częściowej

Grupowania danych w algorytmie jest realizowane według kryterium minimum sumy kwadratów odległości między danymi a odpowiednimi środkami grupy (klastra). Jak widać z rysunku 4.21. mamy dane podzielone na dwa klastry.



Rys. 4.21. Wykres podziału danych wejściowych na klastry

Kolejnym etapem jest stosowanie metody CART, która generuje podzbiór zbioru danych na dwa podzbiory według kryterium minimalizacji heterogeniczności. Każdy podział oparty jest na jednej zmiennej. Podział danych na podzbiory przedstawiono na rysunku 4.22.



Rys. 4.22. Drzewo decyzyjne

Za pomocą drzewa decyzyjnego, otrzymujemy dwa modele, których kryteria informacyjne są porównane w tabeli 4.8.

Tab. 4.8. Kryteria informacyjne dla modeli progowych

Model	TAR	
	0	1
Liczba obserwacji	30	70
Log Likelihood	-41.210	-113.212
AIC	88.421	232.424
BIC	92.624	239.169
HQIC	89.765	235.103
const	13.6723 ± 2.740	-0.3310 ± 0.216
AR.L1	0.9630 ± 0.040	0.3295 ± 0.117

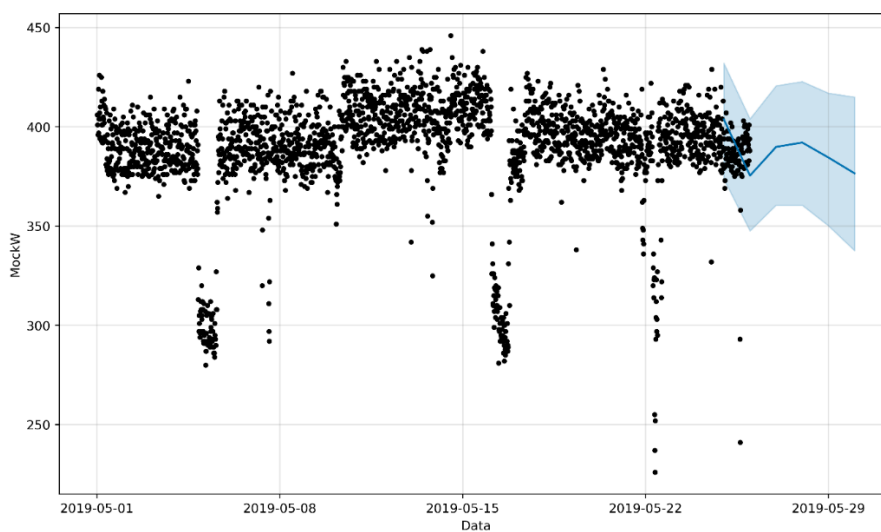
## 4.12. Przykład prognozy danych za pomocą pakietu Facebook Prophet

Dla przykładu przeprowadzimy prognozę danych z liczników elektrycznych z zastosowaniem pakietu Prophet. Przykładowy licznik pokazano na rysunku 4.23. Dane były pobierane co 15 min z liczników elektrycznych w okresie od północy 1.05.2019 do północy 1.06.2019.



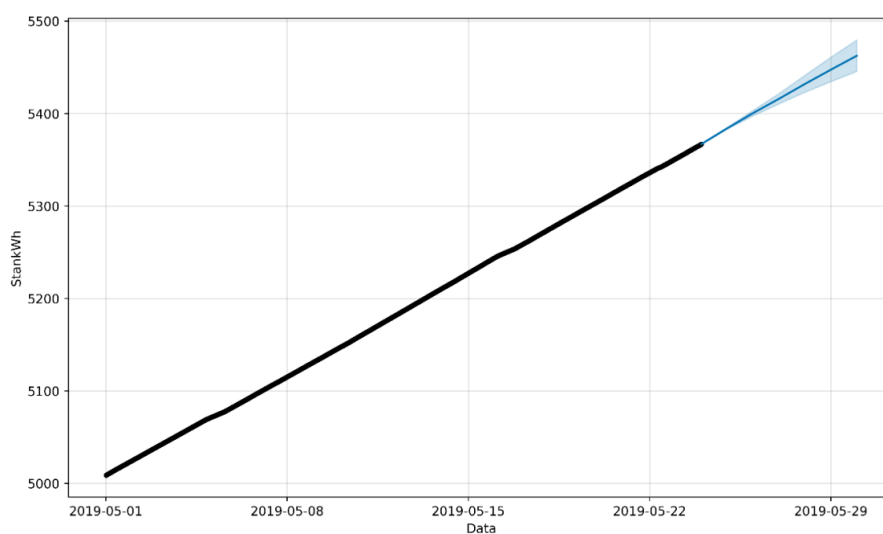
Rys. 4.23. Przykładowy licznik pomiaru energii elektrycznej (źródło: shop.lumel.com.pl)

Zbiór danych reprezentujących moc w kW przedstawiono na rysunku 4.24. Na rysunku przedstawiono niebieską linią uśrednioną prognozę wyznaczoną automatycznie w pakiecie Prophet oraz przedziały ufności prognozy.



Rys. 4.24. Prognoza mocy w kW za pomocą Facebook Prophet

Na rysunku 4.25 przedstawiono stan zużycia energii (kWh) wskazywany przez dany licznik oraz prognozę zużycia, co jest podstawą do wystawiania rachunków.



Rys. 4.25. Prognoza Stan kWh za pomocą Facebook Prophet

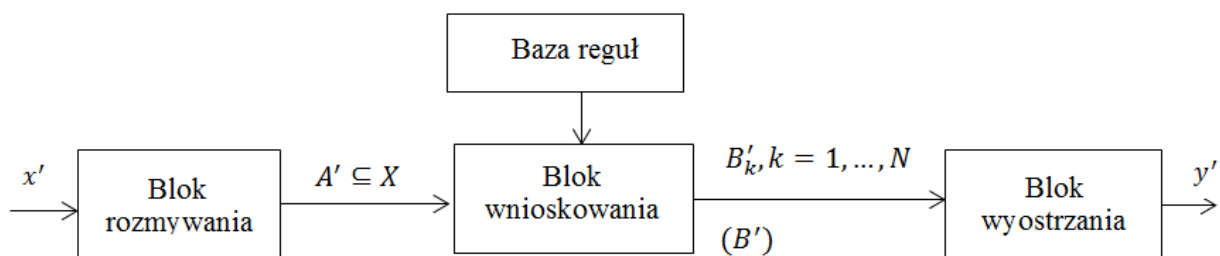
Wizualnie prognoza modelu wydaje się wystarczająco dobrą, z wykresów można wywnioskować, że większość punktów mieści się w przedziale ufności.

## 5. MODELOWANIE MIĘKKIE CIĄGÓW CZASOWYCH I SYSTEMÓW DYNAMICZNYCH DYSKRETNYCH

Obliczenia miękkie (*ang. soft computing*) zapoczątkowane przez L.A. Zadeha są rozumiane jako podejście do rozwiązywania zadań naśladujące możliwości ludzkiego umysłu do percepcji i uczenia się w niepewnym i nieprecyzyjnym otoczeniu. Znana jest tak zwana zasada *niekompatabilności Zadeha*: „jeżeli skomplikowanie systemu wzrasta, nasza zdolność formułowania precyzyjnych i wciąż wiarygodnych stwierdzeń o jego zachowaniu zmniejsza się, a po przekroczeniu pewnego skomplikowania precyzja i wiarygodność wzajemnie się wykluczają” (Łęski, 2008). Do obliczeń miękkich zalicza się takie metody i techniki, jak: zbiory rozmyte, zbiory przybliżone, sztuczne sieci neuronowe, algorytmy genetyczne, teoria prawdopodobieństwa i szereg innych. W praktyce tworzenia informatycznych inteligentnych systemów modelujących panuje tendencja do tworzenia systemów hybrydowych, na przykład neuronowo-rozmytych, rozmytych z zastosowaniem algorytmów genetycznych optymalizujących model rozmyty i t. p. (Zadeh, 1996).

### 5.1. Systemy rozmyte typu Mamdaniego i typu Takagi-Sugeno-Kanga (TSK)

Idea rozmytego systemu wnioskującego została przedstawiona na rysunku 5.1.



Rys. 5.1. Rozmyty system wnioskujący (Rutkowski, 2005, s. 105)

Prezentowane systemy należą do systemów ekspertowych opartych o logikę rozmytą (Zadeh, 1965). Rozmyty system wnioskujący składa się z następujących elementów:

- bazy reguł,
- bloku rozmywania,
- bloku wnioskowania,
- bloku wyostrzania

(Yager i Filev, 1995), (Helendorn i Drianków, 1997), (Rutkowski, 2005), (Łęski, 2008), (Piegat, 2009).



W bloku rozmywania numeryczne wartości wejściowe  $x'$  są przetwarzane na ich odpowiedniki rozmyte,  $A'$ . Zmienna wejściowa  $x$  jest zdeterminowana przez zbiór wartości numerycznych z pewnej przestrzeni  $X$ . W przestrzeni  $X$  określa się tzw. zbiory rozmyte odniesienia, np.  $A_1, A_2, \dots, A_k$  za pomocą funkcji przynależności (Zadeh, 1965). Zbiory te mogą reprezentować pewne lingwistyczne pojęcia opisujące poziom zmiennej wejściowej w poszczególnych przedziałach, jak np. *niski* ( $N$ ), *średni* ( $S$ ), *wysoki* ( $W$ ) lub odpowiednio  $A_1, A_2, A_3$ .

Baza reguł (baza wiedzy) stanowi model wiedzy dotyczącej danego zagadnienia, któremu dedykowany jest system, np. wiedza dotycząca sterowania pewnym obiektem. Wiedza ta jest zapisana w postaci zbioru reguł warunkowych, wiążących zmienne wejściowe (jako przyczyny, przesłanki) ze zmiennymi wyjściowymi (skutki, konkluzje):

*Jeżeli (Przesłanka) To (Konkluzja).*

Przesłanka reguły może mieć postać, np.  $x$  jest  $A_i$ , co rozumiemy jako fakt, że zmienna wejściowa  $x$  przyjmuje wartości ze zbioru rozmytego  $A_i$ , określonego w przestrzeni wejścia  $X$ . Konkluzja reguły może mieć postać, np.  $y$  jest  $B_i$ , gdzie  $B_i$  jest zbiorem rozmytym określonym w przestrzeni wyjścia  $Y$ . Każda reguła stanowi relację rozmytą  $A_i \times B_i$  w pewnej części przestrzeni wejścia-wyjścia,  $X \times Y$ . Podział przestrzeni na poszczególne obszary jest osobnym zagadnieniem, istotnym dla systemów rozmytych. Wiąże się z tym problem pozyskiwania wiedzy oraz optymalnej liczby reguł w bazie wiedzy.

Blok wnioskowania zawiera procedurę wyprowadzania konkluzji rozmytej  $B'$  na podstawie bazy reguł dla zadanej na wejściu wartości numerycznej  $x'$  lub rozmytej  $A'$ . Zasady wnioskowania w logice rozmytej zostaną szerzej omówione w kolejnych podrozdziałach. Jeśli w procesie wnioskowania bierze udział kilka reguł aktywnych i z każdej reguły wyprowadzana jest konkluzja rozmyta  $B_k'$ ,  $k=1,2,\dots,N$  to łączna konkluzja rozmyta  $B'$ , jest wynikiem operacji agregowania poszczególnych konkluzji  $B_k'$ .

Ze względu na postać bazy wiedzy i zasadę wnioskowania rozróżnia się dwa główne typy systemów rozmytych: lingwistyczne systemy typu Mamdaniego i rozmyte systemy Takagi-Sugeno-Kanga, którym poświęcone są następane podrozdziały.

W bloku wyostrzania systemu wnioskującego rozmyta konkluzja  $B'$  jest przetwarzana na wartość numeryczną  $y'$ , która w szczególnym przypadku może być centroidą zbioru rozmytego  $B'$ . Istnieje wiele metod operacji wyostrzania (defuzyfikacji) wartości rozmytych. Wybór właściwej metody jest istotny dla uzyskanej wartości konkluzji nierozmytej  $y'$ .

### **Ogólne zasady wnioskowania w logice rozmytej**

Logika rozmyta sformułowana przez L.A. Zadeha jest logiką nieskończenie wartościową. Wartość prawdy zdania „ $x$  jest  $A$ ” jest równa wartości funkcji przynależności  $\mu_A(x)$ , elementu  $x$  w podzbiorze  $A$  przestrzeni  $X$

Wnioskowanie w logice rozmytej odbywa się głównie na podstawie dwóch schematów (Zadeh, 1973, 1979):

- uogólniona reguła *modus ponens* (ang. *generalized modus ponens*)

Implikacja:  $p \Rightarrow q$

$$\begin{aligned} & \text{Przesłanka: } p' \\ & \text{Wniosek: } q' = p' \circ (p \Rightarrow q) \end{aligned} \quad (5.1)$$

- uogólniona reguła *modus tollens*

$$\begin{aligned} & \text{Implikacja: } p \Rightarrow q \\ & \text{Przesłanka: } q' \\ & \text{Wniosek: } p' = (p \Rightarrow q) \circ q' \end{aligned} \quad (5.2)$$

Pierwszy ze schematów, (5.1), określa tzw. wnioskowanie „w przód”, z nowej przesłanki  $p'$  pozwala wyprowadzić nowy wniosek  $q'$ . Schemat (5.2) pozwala wyprowadzić przesłankę  $p'$ , jeżeli znana jest nowa konkluzja  $q'$ .

W powyższych schematach symbole oznaczają:

- $p, q$  zdania, np.  $x$  jest  $A$ ,  $y$  jest  $B$ ;
- $p \Rightarrow q$  implikacja rozmyta, *Jeżeli  $x$  jest  $A$  To  $y$  jest  $B$* ;
- $p'$  zdanie  $x$  jest  $A'$  (nowa przesłanka)
- $q'$  zdanie  $y$  jest  $B'$  (nowa konkluzja)
- $A, A'$  zbiory rozmyte o funkcjach przynależności  $\mu_A(x), \mu_{A'}(x), x \in X$ ;
- $B, B'$  zbiory rozmyte o funkcjach przynależności  $\mu_B(y), \mu_{B'}(y), y \in Y$ ;
- o operacja złożenia *max – min*.

Wniosek reguły (5.1) można zapisać za pomocą funkcji przynależności

$$\mu_{B'}(y) = \sup_{x \in X} \min[\mu_{A'}(x), \mu_{A \Rightarrow B}(x, y)] \quad (5.3)$$

Lub bardziej ogólnie, używając  $t$ -normy

$$\mu_{B'}(y) = \sup_{x \in X} t[\mu_{A'}(x), \mu_{A \Rightarrow B}(x, y)]. \quad (5.4)$$

Wniosek reguły (5.2) można zapisać w następujących postaciach:

$$\mu_{A'}(x) = \sup_{y \in Y} \min[\mu_{A \Rightarrow B}(x, y), \mu_{B'}(y)] \quad (5.5)$$

Lub bardziej ogólnie, używając  $t$ -normy

$$\mu_{A'}(x) = \sup_{y \in Y} t[\mu_{A \Rightarrow B}(x, y), \mu_{B'}(y)] \quad (5.6)$$

Wzory (5.3) – (5.6) pozwalają określić stopnie prawdy  $\mu_{A'}(x)$ ,  $\mu_{B'}(y)$  otrzymanych wniosków a w algorytmach praktycznych pozwalają wyznaczyć zbiory rozmyte konkluzji  $A'$ ,  $B'$ .

Pokazana we wzorach implikacja rozmyta o symbolu  $\Rightarrow$ , posiada wiele interpretacji (np. w (Łęski, 2008)). Jedną z wielu, to implikacja rozmyta Kleene-Dienes:

$$\mu_{A \Rightarrow B}(x, y) = \max[1 - \mu_A(x), \mu_B(y)] \quad (5.7)$$

$$\text{dla } \forall x \in X, \forall y \in Y.$$

Operacje  $\max[\mu_A(x), \mu_B(y)]$  oraz  $\min[\mu_A(x), \mu_B(y)]$  są operacjami sumy  $A \cup B$  i przecięcia (iloczynu)  $A \cap B$  zbiorów rozmytych. Bardziej ogólnie, operacje te wyrażają się poprzez  $s$ -normy i  $t$ -normy.

Funkcję dwóch zmiennych  $a, b$ ,  $T: [0,1] \times [0,1] \rightarrow [0,1]$  nazywamy  $t$ -normą, jeżeli spełnia następujące warunki (Rutkowski, 2005), (Łęski, 2008):

- 1) funkcja  $T$  jest niemalejąca względem obu argumentów,
- 2) funkcja  $T$  spełnia warunek przemienności,
- 3) funkcja  $T$  spełnia warunek łączności,
- 4) funkcja  $T$  spełnia warunek brzegowy  $T(a, 1) = a$ .

Ponadto,  $T(a, 0) = 0$ .

Podobnie,  $s$ -norma jest funkcją dwóch zmiennych  $a, b$   $S: [0,1] \times [0,1] \rightarrow [0,1]$ , spełnia warunki 1-3 oraz warunek brzegowy

$$S(a, 1) = 1$$

$$\text{oraz } S(a, 0) = a.$$

### Lingwistyczne modele typu Mamdaniego

Rozmyty system wnioskujący typu Mamdaniego został wprowadzony przez Mamdaniego i Assiliana w roku 1975 (Mamdani i Assilian, 1975).

Baza wiedzy typu Mamdaniego, stanowiąca model rozmyty systemu o jednym wejściu i jednym wyjściu (SISO), ma postać zbioru reguł warunkowych, typu (Yager i Filev, 1995), (Helendorn i Drianków, 1997), (Rutkowski, 2005), (Łęski, 2008):

$$R_i: \text{Jeżeli } x \text{ jest } A_i \text{ To } y \text{ jest } B_j \quad (5.8)$$

$$i = 1, \dots, I, j = 1, \dots, J$$

Zbiór reguł (5.8) opisuje pewien system (zjawisko, proces), w którym  $x$  jest zmienną wejściową,  $y$  jest zmienną wyjściową,  $A_i$ ,  $i = 1, \dots, I$  są zbiorami rozmytymi w przestrzeni zmiennej wejściowej,  $B_j$ ,  $j = 1, \dots, J$  są zbiorami rozmytymi określonymi w przestrzeni zmiennej wyjściowej.

Część reguły stanowiącą warunek nazywa się *poprzednikiem*, a część po słowie *To* nazywa się *następnikiem*. Reguły są powiązane pomiędzy sobą spójnikiem *Lub (Także)*. Każda reguła reprezentuje zależność między zmiennymi w pewnym elementarnym obszarze przestrzeni  $X \times Y$ , wyznaczonym przez iloczyn kartezjański zbiorów rozmytych  $A_i \times B_j$ .

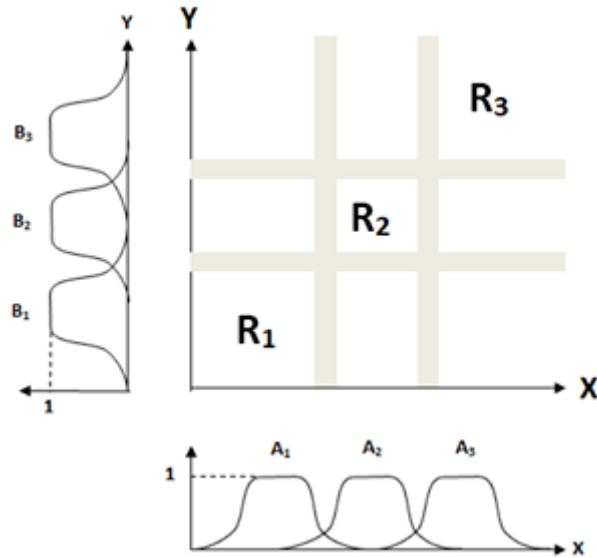
Na rysunku 5.2. przedstawiono podział przestrzeni  $X \times Y$  na obszary rozmyte:  $A_i \times B_j$ ,  $i=1,2,3, j=1,2,3$ . Przykładowa baza wiedzy systemu Mamdaniego może mieć postać:

$$R1: \text{Jeżeli } x \text{ jest } A_1 \text{ To } y \text{ jest } B_1$$

$$R2: \text{Jeżeli } x \text{ jest } A_2 \text{ To } y \text{ jest } B_2$$

$$R3: \text{Jeżeli } x \text{ jest } A_3 \text{ To } y \text{ jest } B_3.$$

Procedura wnioskowania rozpoczyna się wtedy, gdy na wejście systemu zostanie podany nowy fakt (nowa przesłanka, wartość  $x$ ) czyli zdanie w logice rozmytej  $x \text{ jest } A'$ .



Rys. 5.2. Podział przestrzeni  $X \times Y$  za pomocą zbiorów rozmytych

W myśl uogólnionej rozmytej reguły *modus ponens*, "potwierdza przez potwierdzenie", odpowiedzią będzie sygnał  $B'_j$  konkluzji rozmytej wynikający z  $j$ -tego następnika (Zadeh, 1979).

$$B'_j = A' \circ (A_i \rightarrow B_j), \quad (5.9)$$

gdzie symbol „ $\circ$ ” oznacza złożenie typu *max-min*, symbol  $\rightarrow$  oznacza implikację rozmytą.

W modelu Mamdaniego funkcja przynależności relacji rozmytej reprezentującej zdanie warunkowe (implikację rozmytą) jest *t-normą minimum*

$$\mu_{A \Rightarrow B}(x, y) = \min(\mu_{A_i}(x), \mu_{B_j}(y)), \forall (x, y) \in X \times Y, \quad (5.10)$$

więc funkcja przynależności zbioru  $B'_j$  ma postać:

$$\mu_{B'_j}(y) = \sup_{x \in X} \{\min[\mu_{A'}(x), \min(\mu_{A_i}(x), \mu_{B_j}(y))]\}. \quad (5.11)$$

Konkluzja rozmyta  $B'$  wyznaczona z całej bazy reguł, czyli zagregowana, jest sumą rozmytą zbiorów  $B'_j$  wyznaczonych w poszczególnych regułach

$$B' = \cup_j B'_j \quad (5.12)$$

Funkcję przynależności zagregowanej konkluzji wyznacza się za pomocą operatora *max* (ogólnie: *supremum*)

$$\mu_{B'}(y) = \sup_{y \in Y} [\mu_{B'_1}(y), \dots, \mu_{B'_j}(y)]. \quad (5.13)$$

Procedura wnioskowania w modelu Mamdaniego opiera się tylko na operatorach *supremum* (*maximum*) oraz *minimum* (Mamdani i Assilian, 1975).

Baza wiedzy typu Mamdaniego, stanowiąca model rozmyty systemu o dwóch wejściach i jednym wyjściu, MIMO, jest zbiorem reguł warunkowych o postaci (Yager i Filev, 1995),

(Helendorn i Drianków, 1997), (Rutkowski, 2005), (Łęski, 2008), (Walaszek-Babiszewska, Bryniarska, 2018):

$$R_{ijk}: \text{Jeżeli } x \text{ jest } A_i \text{ I } y \text{ jest } B_j \text{ To } z \text{ jest } C_k \quad (5.14)$$

$$i=1, \dots, I, j=1, \dots, J, k=1, \dots, K.$$

Poprzednik reguły zawiera koniunkcję dwóch zdań dla dwóch zmiennych w przestrzeniach  $X$  i  $Y$ . Regułę warunkową można więc zapisać (Walaszek-Babiszewska i Bryniarska, 2018):

$$R_{ijk} = [(A_i \times B_j) \rightarrow C_k]. \quad (5.15)$$

Implikacja rozmyta w modelu Mamdaniego jest operacją *minimum*, podobnie jak iloczyn kartezyjański zbiorów poprzednika.

$$\mu_{(A_i \times B_j) \Rightarrow C_k}(x, y, z) = \min[\min(\mu_{A_i}(x), \mu_{B_j}(y)), \mu_{C_k}(z)], \quad (5.16)$$

$$\forall x \in X, \forall y \in Y, \forall z \in Z.$$

Modele wnioskowania Mamdaniego charakteryzują się tym, że występujące koniunkcje, implikacje i alternatywy rozmyte zastępowane są tylko operatorami *minimum* i *maximum*.

### Modele Takagi-Sugeno-Kanga (TSK)

System TSK pozwala na utworzenie rozmytego modelu systemu w postaci warunkowych reguł opisujących działanie systemu w poszczególnych częściach przestrzeni wejścia-wyjścia. Każda z reguł w modelu TSK łączy w sobie lingwistyczny sposób formułowania warunku z funkcyjnym następnikiem (Takagi i Sugeno, 1985). Dla systemu o dwóch wejściach i jednym wyjściu pojedynczą regułę możemy zapisać w postaci:

$$R_k: \text{Jeżeli } x \text{ jest } A_i \text{ I } y \text{ jest } B_j \text{ To } z = f_k(x, y) \quad (5.17)$$

$$i = 1, \dots, I, j = 1, \dots, J, k=1, \dots, K.$$

Następnik  $f_k(x, y)$  jest nierozmytą funkcją zmiennych numerycznych wejść. Często funkcja ta jest modelem liniowym.

Algorytm wnioskowania z modelu TSK o dwóch wejściach możemy zapisać następująco:

- niech aktualne wartości wejść wynoszą  $x'$  i  $y'$ , odpowiednio;
- poziom zapłonu (aktywizacji)  $k$ -tej reguły,  $\tau_k$ , wyznacza się jako t-normę (np. minimum) ze stopni przynależności obydwu wejść:

$$\tau_k = \min[\mu_{A_i}(x'), \mu_{B_j}(y')]; \quad (5.18)$$

- poziomy zapłonu wyznacza się dla każdej reguły, w której stopnie przynależności  $\mu_{A_i}(x')$ ,  $\mu_{B_j}(y')$  są niezerowe;
- wyznacza się wartości funkcji  $f_k(x', y') = z_k$  w poszczególnych regułach aktywnych;
- wyjście modelu globalnego  $z^*$  jest sumą ważoną wyjść wszystkich reguł aktywnych

$$z^* = \sum_k \frac{\tau_k}{\sum_k \tau_k} z_k . \quad (5.19)$$

Tak więc, w całej przestrzeni wejścia-wyjście model TSK można uważać za zbiór modeli w podprzestrzeniach połączonych operatorem wnioskowania.

Procedura wnioskowania dla modeli rozmytych TSK zostanie szczegółowo omówiona na przykładach w rozdziale 5.2.

## 5.2. Rozmyte modele autoregresji

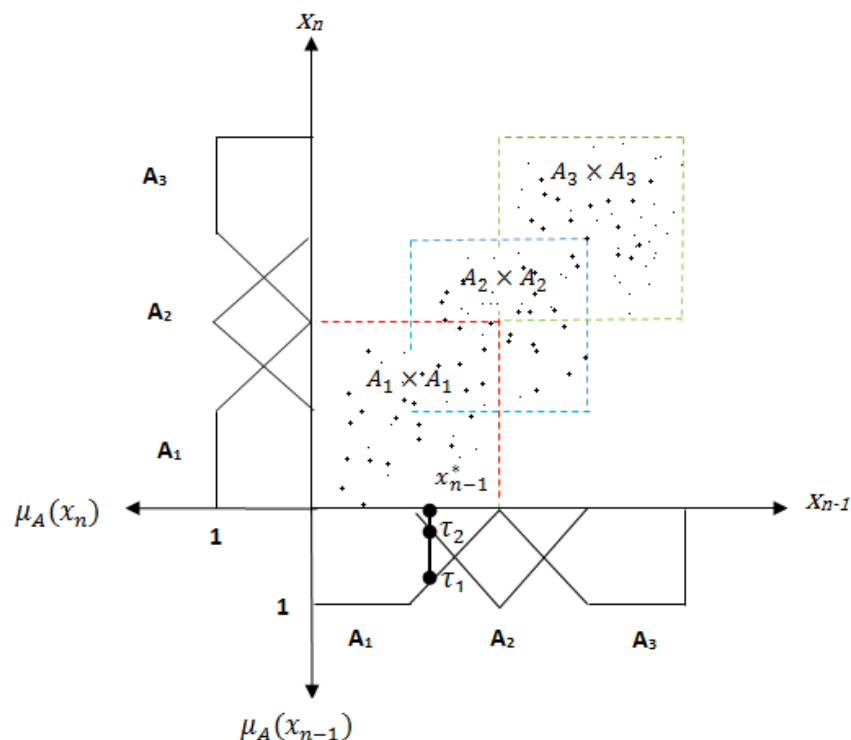
### Rozmyte modele AR(1) i AR(2) oparte o reguły Mamdaniego

Baza wiedzy typu Mamdaniego dla systemu dynamicznego, który opisywany jest za pomocą modelu AR(1) ma postać reguł:

$$R_{i,j}: \text{Jeżeli } X_{n-1} \text{ jest } A_i \text{ To } X_n \text{ jest } A_j \quad (5.20)$$

$$i,j=1,2,\dots,I.$$

Rysunek 5.3. przedstawia podział przestrzeni  $\chi \times \chi$ , gdzie  $x_{n-1}, x_n \in \chi$ , za pomocą zbiorów rozmytych. Zbiory rozmyte są opisywane za pomocą funkcji przynależności  $\mu_{A_i}(x)$ . Funkcje te są odcinkami liniowymi.



Rys. 5.3. Ilustracja podziału przestrzeni  $\chi \times \chi$  i wyznaczania zapłonu reguł

Baza wiedzy uwzględnia wyniki obserwacji, przedstawione na rysunku 5.3. w postaci pola korelacji. Zmienność obiektu może być ujęta w postaci trzech reguł:

$R_1$ : Jeżeli  $x_{n-1}$  jest  $A_1$  To  $x_n$  jest  $A_1$

$R_2$ : Jeżeli  $x_{n-1}$  jest  $A_2$  To  $x_n$  jest  $A_2$

$R_3$ : Jeżeli  $x_{n-1}$  jest  $A_3$  To  $x_n$  jest  $A_3$ .

W pierwszej kolejności realizowany jest proces obliczania poziomu aktywacji reguł. W tym celu dla każdej zmiennej w przesłankach reguły wyznaczane są stopnie przynależności do odpowiedniego zbioru rozmytego. Ze względu na operację logiczną AND, poziomu aktywacji reguły jest obliczany jako wartość minimalna ze wszystkich stopni przynależności. Jeśli wartość poziomu aktywacji reguły jest zerowa, wówczas uznaje się, że aktywacja reguły nie nastąpiła. Jednocześnie wyznaczany jest zbiór rozmyty będący rezultatem uaktywnienia reguły. Ten zbiór zależy od kształtu funkcji przynależności oraz obliczonego poziomu aktywacji reguły. W następnym kroku następuje agregacja aktywnych reguł. Ten proces polega na sumowaniu rozmytych zbiorów wynikowych ze wszystkich reguł. Otrzymany w ten sposób zbiór rozmyty jest zbiorem końcowym wnioskowania rozmytego.

Wnioskowanie według paradygmatu Mamdaniego (Yager i Filev, 1995), (Rutkowski, 2005) można sformułować w kilku krokach. Dla ustalonego wejścia  $x_{n-1}^*$  poziomu aktywacji (zapłonu,  $\tau_1$  i  $\tau_2$ ) reguł wynoszą (rysunek 5.3.):

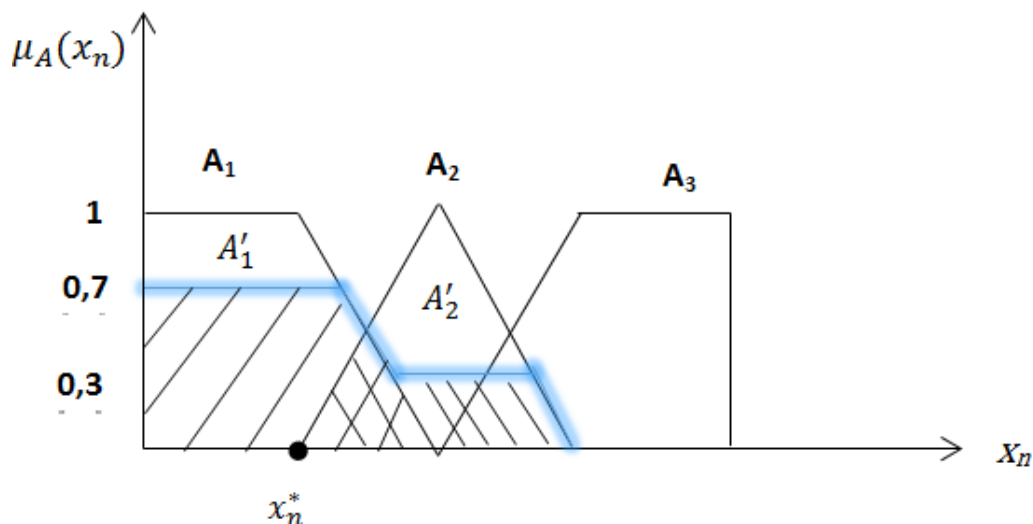
$$\mu_{A_1}(x_{n-1}^*) = \tau_1 = 0,7$$

$$\mu_{A_2}(x_{n-1}^*) = \tau_2 = 0,3$$

Konkluzje rozmyte  $A'_1, A'_2$  z poszczególnych reguł wyznacza się korzystając z operatora  $\min$  (rysunek 5.4.):

$$\mu_{A'_1}(x_n) = \min[\tau_1, \mu_{A_1}(x_n)] \quad (5.21)$$

$$\mu_{A'_2}(x_n) = \min[\tau_2, \mu_{A_2}(x_n)] \quad (5.22)$$



Rys. 5.4. Ilustracja wyznaczania konkluzji rozmytej

Agregację konkluzji rozmytych  $A'_1, A'_2$  wyznacza się za pomocą operatora  $\max$ , jak na rysunku 5.4.:

$$\mu_{A'}(x_n) = \max_{x_n}(\mu_{A'_1}(x_n), \mu_{A'_2}(x_n)). \quad (5.23)$$

Wyostżanie, to jest wyznaczenie wartości numerycznych prognozy  $x_n^*$  realizuje się stosując metodę środka ciężkości. Wyznaczona tą metodą wartość  $x_n^*$  została zaznaczona na rysunku 5.4.

Dla dowolnej wartości wejścia  $x_{n-1}$  w sposób analogiczny otrzymujemy odpowiednie wartości wyjść  $x_n^*$ .

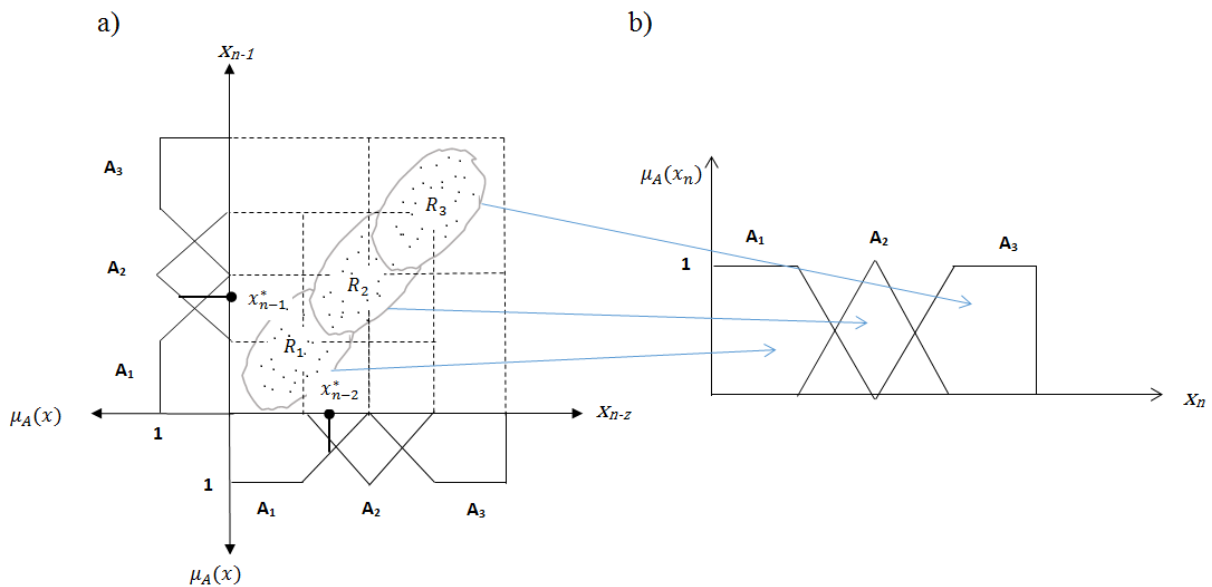
Dla systemu dynamicznego AR(2) baza wiedzy modelu lingwistycznego ma postać następujących reguł:

$$R_{ijk}: \text{Jeżeli } x_{n-1} \text{ jest } A_i \text{ I } x_{n-2} \text{ jest } A_j \text{ To } x_n \text{ jest } A_k, \quad (5.24)$$

$$i, j, k = 1, 2, \dots, I.$$

Zbiory rozmyte  $A_i, A_j, A_k$  są określone w przestrzeni  $\chi$ . Załóżmy dla przykładu, że dany jest zbiór obserwacji  $\{x_1, x_2, \dots, x_m\}$ , który jest przedstawiony w postaci pola korelacji  $x_n = f(x_{n-1}, x_{n-2})$ , oraz  $x_{n-1} = f(x_{n-2})$  jak dla danych notowań USD/PLN przedstawionych na rysunku 5.4.

Przestrzeń zmiennych wejściowych  $\chi \times \chi$  została podzielona za pomocą zbiorów rozmytych i tworzą regularną siatkę 3x3. Na tle siatki, na podstawie pola korelacji można określić trzy nierozłączne części przestrzeni:  $A_1 \times A_1, A_2 \times A_2, A_3 \times A_3$  (rysunek 5.5a.).



Rys. 5.5. Podział przestrzeni wejścia na zbiory rozmyte (a), zbiory rozmyte w przestrzeni wyjścia (b)

Baza wiedzy, wynikająca z pola korelacji, jak na rysunku 5.5., ma postać trzech reguł:

$R_1$ : Jeżeli  $x_{n-1}$  jest  $A_1$  I  $x_{n-2}$  jest  $A_1$  To  $x_n$  jest  $A_1$

$R_2$ : Jeżeli  $x_{n-1}$  jest  $A_2$  I  $x_{n-2}$  jest  $A_2$  To  $x_n$  jest  $A_2$

$R_3$ : Jeżeli  $x_{n-1}$  jest  $A_3$  I  $x_{n-2}$  jest  $A_3$  To  $x_n$  jest  $A_3$ .

Wnioskowanie (prognoza) oparte o zasady Mamdaniego i Assiliana może przebiegać następująco.



Niech obserwacje  $x_{n-1}, x_{n-2}$  wynoszą  $x_{n-1}^*, x_{n-2}^*$  (jak na rysunku 5.5a.). Wartości  $x_{n-1}^*, x_{n-2}^*$  te powodują że aktywne stają się dwie reguły:  $R_1$  i  $R_2$ .

Wartości zapłonu reguł wyznaczymy z rysunku 4.5a. Ten etap procedury wnioskowania dotyczy tzw. rozmywania.

$$\mu_{A_1}(x_{n-2}^*) = \tau_{21} = 0,6$$

$$\mu_{A_2}(x_{n-2}^*) = \tau_{22} = 0,4$$

$$\mu_{A_1}(x_{n-1}^*) = \tau_{11} = 0,3$$

$$\mu_{A_2}(x_{n-1}^*) = \tau_{12} = 0,7$$

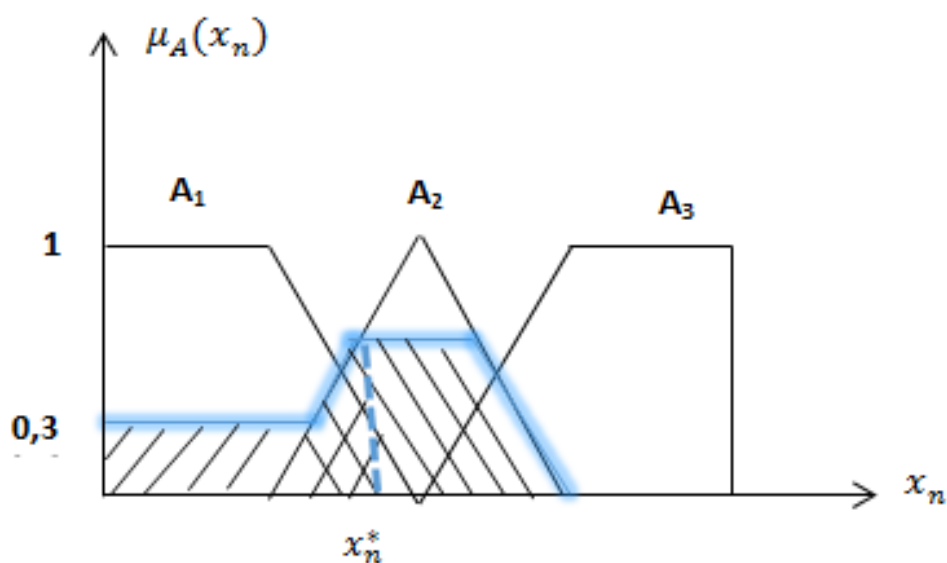
Poziomy zapłonu każdej reguły uwzględniają zbiory rozmyte występujące w danej regule. Operator *min* odpowiada spójnikowi *I* pomiędzy zmiennymi wejściowymi w regule (5.24):

$$\tau_1 = \min(\tau_{11}, \tau_{21}) = \min(0.3, 0.6) = 0.3 \quad (5.25)$$

$$\tau_2 = \min(\tau_{12}, \tau_{22}) = \min(0.7, 0.4) = 0.4 \quad (5.26)$$

Rozmyta konkluzja (zbiór rozmyty)  $A'$  ma więc postać:

$$\mu_{A'}(x_n) = \max \left[ \min(\tau_1, \mu_{A_1}(x_n)), \min(\tau_2, \mu_{A_2}(x_n)) \right] \quad (5.27)$$



Rys. 5.6. Konkluzja rozmyta  $A'$  z modelu AR(2)

Zbiór  $A'$  przedstawiono na rysunku 5.6. jako zakreskowany obszar. Wartość nierozmytą konkluzji, odpowiadającą prognozie  $x_n^*$  wyznacza się np. metodą środka ciężkości zbioru  $A'$ .

Wszystkie przedstawione operacje, dotyczące modelowania rozmytego można zrealizować w systemie MATLAB Fuzzy Logic Toolbox.

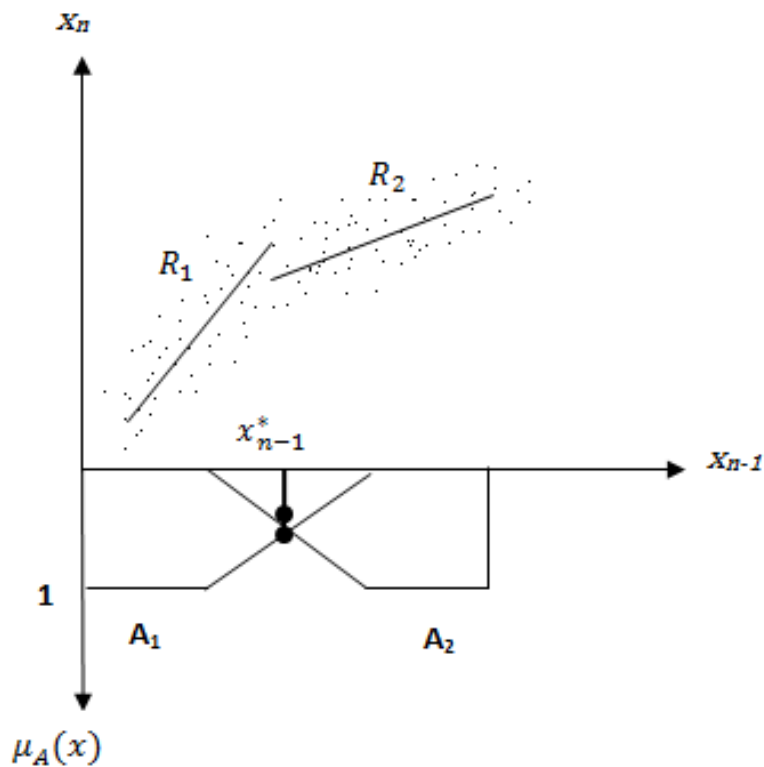
## Rozmyte modele AR(1) i AR(2) oparte o reguły TSK

System dynamiczny AR(1) można przedstawić w postaci modelu rozmytego typu TSK w sposób taki, jak pokazuje poniższy przykład. Podziału przestrzeni  $\chi$ ,  $x_{n-1} \in \chi$  na zbiory rozmyte dokonują zazwyczaj eksperci na podstawie znajomości procesu lub też w oparciu o pomiary i pole korelacji (Marszałek i Burczyński, 2013), na przykład jak na rysunku 5.7.

Baza wiedzy składa się z dwóch reguł:

$$R_1: \text{Jeżeli } x_{n-1} \text{ jest } A_1 \text{ To } x'_n = a_1 x_{n-1} + b_1 \quad (5.28)$$

$$R_2: \text{Jeżeli } x_{n-1} \text{ jest } A_2 \text{ To } x''_n = a_2 x_{n-1} + b_2 \quad (5.29)$$



Rys. 5.7. Zbiory rozmyte poprzednika i funkcje liniowe następnika modelu TSK, AR(1)

Następnik każdej reguły jest funkcją liniową numerycznych wartości zmiennej poprzednika,  $x_{n-1}$ . Parametry funkcji liniowych  $a_i, b_i$  zostają obliczone jako estymatory metody najmniejszych kwadratów, w oparciu o znane obserwacje.

Wnioskowanie przeprowadza się następująco: poziomy zapłonu reguł wyznacza się tak samo jak dla modeli Mamdaniego

$$\tau_1 = \mu_{A_1}(x_{n-1}^*) = 0.6$$

$$\tau_2 = \mu_{A_2}(x_{n-1}^*) = 0.4$$

Konkluzję nierozmytą (prognozę)  $x_n^*$  otrzymuje się jako następującą średnią ważoną obliczonych wcześniej wartości funkcji liniowych  $x_n'$  i  $x_n''$ .

$$x_n^* = \frac{\tau_1 x_n'(x_{n-1}^*) + \tau_2 x_n''(x_{n-1}^*)}{\tau_1 + \tau_2} \quad (5.30)$$

Konkluzja (5.30) z modelu TSK jest także stosowana w systemach neuronowo-rozmytych.

W modelach rozmytych TSK bardzo istotny dla wyniku modelu jest właściwy podział przestrzeni zmiennych poprzednika, czyli zdefiniowanie bazy reguł.

System dynamiczny AR(2) może być przedstawiony w postaci modelu rozmytego TSK, jako zbiór reguł.

$$R_{ijk}: \text{Jeżeli } x_{n-1} \text{ jest } A_i \text{ I } x_{n-2} \text{ jest } A_j \text{ To } x_n^* = a_k x_{n-1} + b_k x_{n-2} + c_k \quad (5.31)$$

Parametry  $a_k, b_k, c_k$  funkcji liniowych oblicza się w oparciu o znane obserwacje procesu wykorzystując metodę najmniejszych kwadratów.

Wnioskowanie przeprowadza się w kilku etapach.

Na etapie wyznaczania zapłonu reguł postępuje się podobnie jak w modelu Mamdaniego dla AR(2), znajdując np.  $\tau'$  i  $\tau''$  dla dwóch reguł aktywnych, przy założeniu wejść  $x_{n-1}^*, x_{n-2}^*$ . Następnie oblicza się wartości funkcji liniowych (następników)  $x_n'$  i  $x_n''$  dla wartości  $x_{n-1}^*$  i  $x_{n-2}^*$  oraz wyznacza się średnią ważoną wyjść (prognozę):

$$x_n^* = \frac{\tau' x_n' + \tau'' x_n''}{\tau' + \tau''} \quad (5.32)$$

W modelach TSK następnik może być funkcją liniową bez wyrazu wolnego lub dowolną funkcją nieliniową. Wybór właściwej funkcji następnika oraz podział przestrzeni zmiennych wejściowych mają decydujące znaczenie dla dokładności modelu, w przypadku systemów dynamicznych decydują o wyniku prognozy.

### 5.3. Adaptacyjny system neuronowo-rozmyty (ANFIS)

ANFIS (*ang. Adaptive Neuro-Fuzzy Inference System*) jako jeden z pierwszych systemów tego typu został opracowany przez Janga w latach 1992-96 (Jang, 1993). Obecnie ANFIS jest szeroko stosowany w systemie obliczeń MATLAB. Ze względu na architekturę, jest to wielowarstwowa sieć neuronowa. Z punktu widzenia funkcji działań system ANFIS jest równoważny systemowi wnioskowania rozmytego typu Takagi-Sugeno-Kanga. Przykładowa sieć neuronowa struktury ANFIS została przedstawiona na rysunku 5.8. Sieć realizuje funkcje systemu TSK o dwóch zmiennych wejściowych i bazie reguł o postaci:

$$R^{(1)}: \text{Jeżeli } x_1 \text{ jest } A_1^{(1)} \text{ I } x_2 \text{ jest } A_2^{(1)} \text{ To } y^{(1)} = f^{(1)}(x_1, x_2), \quad (5.33)$$

$$R^{(2)}: \text{Jeżeli } x_2 \text{ jest } A_1^{(2)} \text{ I } x_1 \text{ jest } A_2^{(2)} \text{ To } y^{(2)} = f^{(2)}(x_1, x_2). \quad (5.34)$$

Realizacja wnioskowania przebiega następująco:

a) sygnały wejściowe, numeryczne wartości  $x_{o1}$  i  $x_{o2}$ , są podawane do systemu za pomocą warstwy wejściowej;

b) w pierwszej warstwie ukrytej wyznaczane są stopnie przynależności

$$\tau_1^{(1)} = \mu_{A_1^{(1)}}(x_{o1}) \text{ i } \tau_2^{(1)} = \mu_{A_2^{(1)}}(x_{o1})$$

$$\tau_1^{(2)} = \mu_{A_1^{(2)}}(x_{o2}) \text{ i } \tau_2^{(2)} = \mu_{A_2^{(2)}}(x_{o2});$$

Zbiory rozmyte zawarte w przesłankach reguł są na ogół funkcjami gaussowskimi, nieliniowymi.

c) druga warstwa ukryta realizuje *t-normę*, dla wyznaczenia stopnia aktywacji każdej reguły. Operator t-normy określa sposób znalezienia części wspólnej dla zbiorów rozmytych.

$$\tau^{(1)} = t(\tau_1^{(1)}, \tau_2^{(1)})$$

$$\tau^{(2)} = t(\tau_1^{(2)}, \tau_2^{(2)});$$

d) w trzeciej warstwie ukrytej realizowane są znormalizowane współczynniki aktywacji  $n^{(1)}$  i  $n^{(2)}$  każdej reguły

$$n^{(1)} = \frac{\tau^{(1)}}{\tau^{(1)} + \tau^{(2)}},$$

$$n^{(2)} = \frac{\tau^{(2)}}{\tau^{(1)} + \tau^{(2)}};$$

e) w czwartej warstwie ukrytej realizowane są konkluzje poszczególnych reguł pomnożone przez znormalizowany stopień aktywacji danej reguły

$$y^{(1)} = n^{(1)} \cdot f_1(x_{o1}, x_{o2}) \text{ oraz } y^{(2)} = n^{(2)} \cdot f_2(x_{o1}, x_{o2});$$

f) warstwa wyjściowa realizuje sumę konkluzji poszczególnych reguł

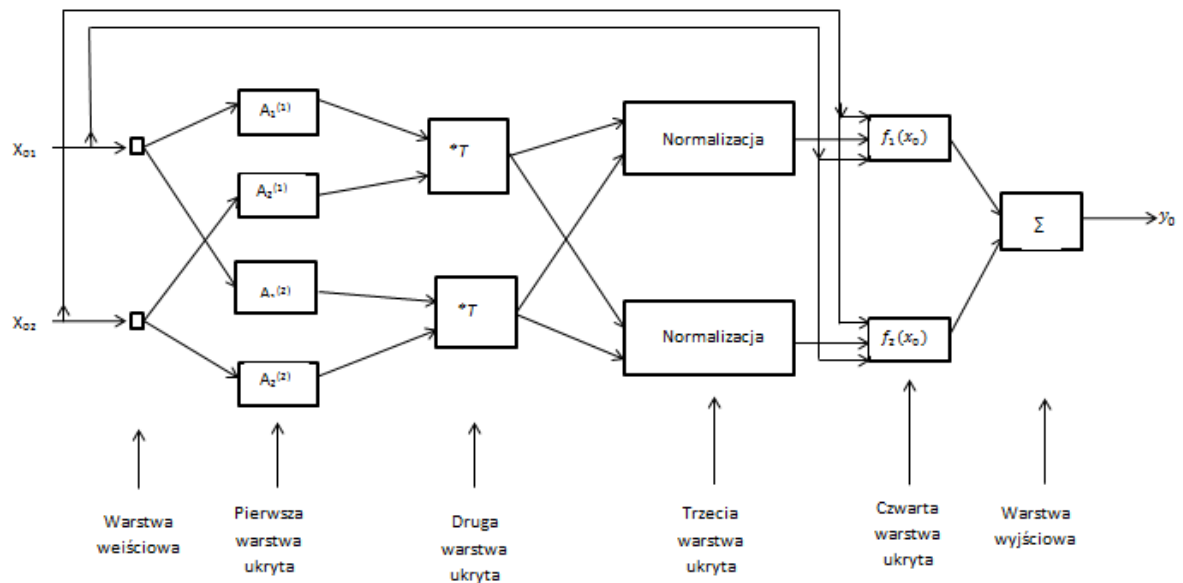
$$y_o = y^{(1)} + y^{(2)}.$$

Przedstawiona procedura wnioskowania jest realizowana w myśl zasady systemu TSK, czyli konkluzja systemu jest sumą ważoną konkluzji nierozmytych poszczególnych reguł. Wagi tej sumy są wyznaczone w oparciu o stopnie przynależności numerycznych wartości sygnałów wejściowych. Funkcje  $f_1$  i  $f_2$  są wielomianami pierwszego lub zerowego rzędu.

Modelowanie rozmyte w systemie MATLAB-ANFIS rozpoczyna się od wyboru początkowej struktury: liczby zmiennych wejściowych, zakresu zmiennej wyjściowej. W następnym kroku zadaje się liczbę zbiorów rozmytych i kształt funkcji przynależności. Liczba reguł jest równa liczbie obszarów pełnej siatki powstałej przez podział przestrzeni zmiennych wejściowych przez zbiory rozmyte. Ten rodzaj podziału przestrzeni ma tą wadę że liczba reguł zależy istotnie od wymiarowości systemu.

ANFIS jest siecią uczącą się. W trakcie procesu uczenia mogą być modyfikowane iteracyjnie parametry przesłanek reguł. Tak zwany współczynnik uczenia zależy od sumy kwadratów błędów, przy czym błąd jest różnicą pomiędzy wyrazami ciągu uczącego i

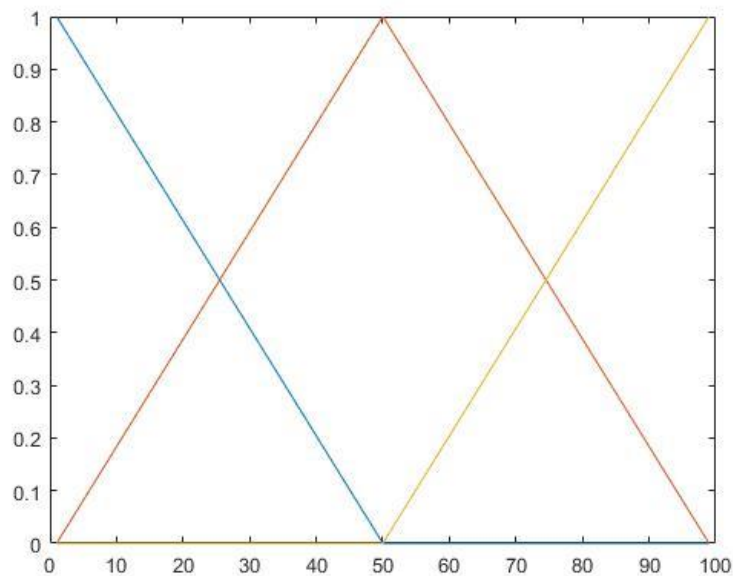
wyznaczoną wartością wyjścia systemu ANFIS. i parametry systemu np. parametry przesłanek reguł warunkowych, są dobierane iteracyjnie w trakcie uczenia.



Rys. 5.8. Sieć neuronowa typu ANFIS równoważna systemowi TSK o dwóch wejściach, jednym wyjściu i dwóch regułach warunkowych (Łęski, 2008, s. 486)

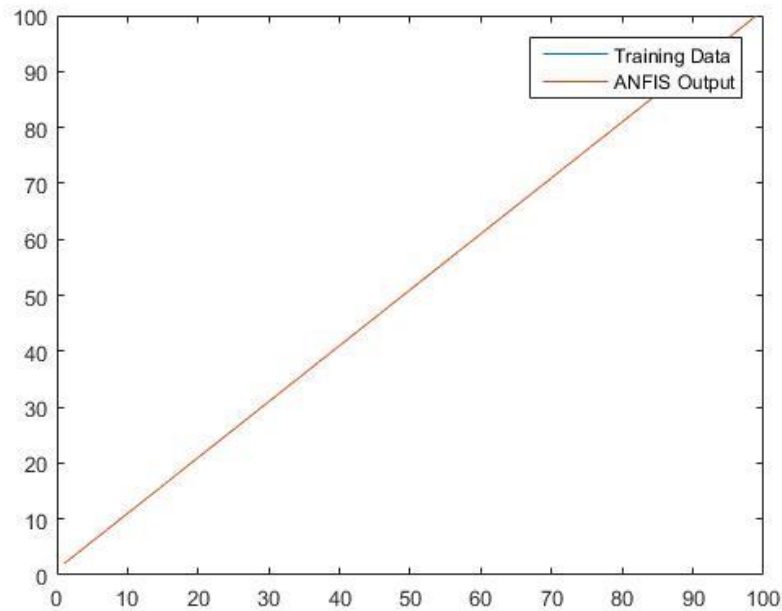
#### 5.4. Przykład identyfikacji danych testowych

Aby sprawdzić charakterystykę przejściową w systemie ANFIS wyjście  $W_y = f(W_e)$  przyjęto, jako ciąg danych wejściowych rosnący zbiór liczb o dużym zakresie:  $\{0, 1, 2, \dots, 100\}$ .

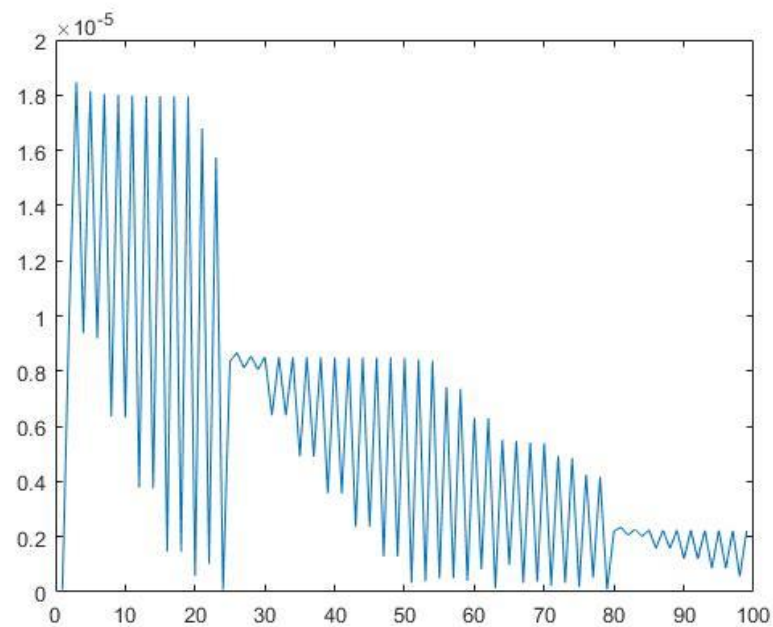


Rys. 5.9. Podział przestrzeni na zbiory rozmyte o funkcjach przynależności typu trójkątnego

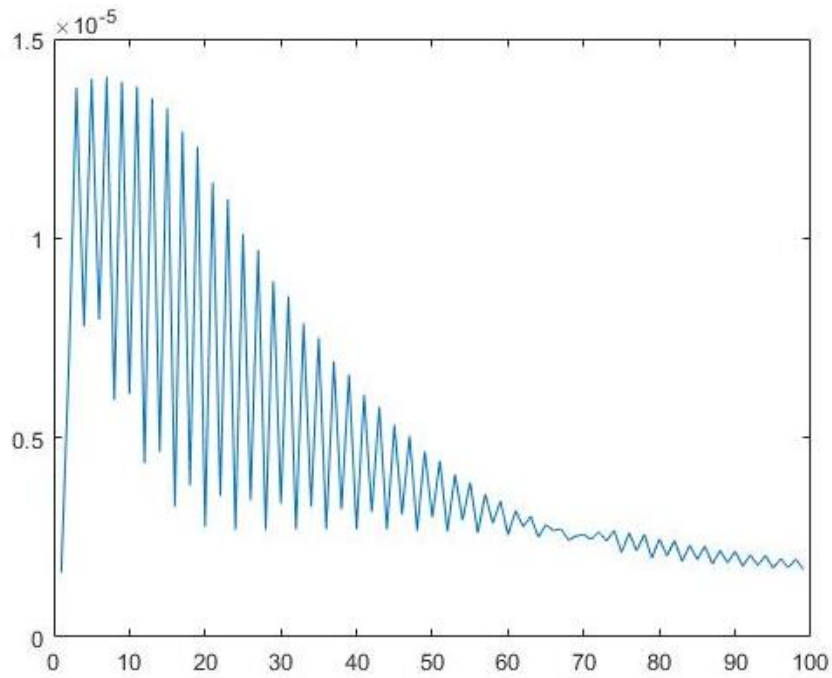
Celem testu było badanie liniowości charakterystyki dla różnych typów zbiorów rozmytych. Dla zbiorów rozmytych o trójkątnych funkcjach przynależności (rysunek 5.9.) charakterystyka przejściowa ma postać liniowa, jak na rysunku 5.10. Uwzględniając przebieg błędów uczenia dla różnej liczby zbiorów trójkątnych (rysunki 5.11 i 5.12.) można powiedzieć że błąd względny charakterystyki nie jest większy niż  $4 \times 10^{-3}\%$ .



Rys. 5.10. Wartości prognozy (ANFIS) ciągu danych testowych na tle danych wejściowych przy funkcjach przynależności typu trójkątnego

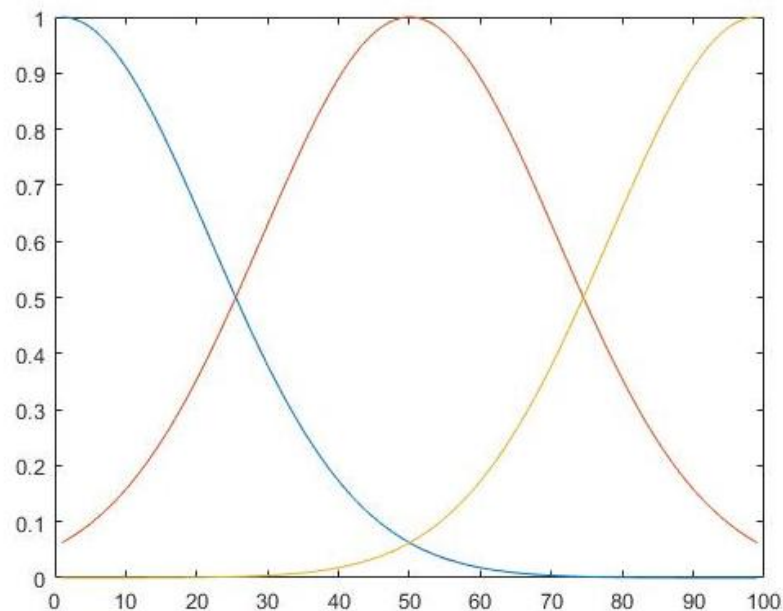


Rys. 5.11. Zmienność błędów w procesie uczenia prognozy dla danych testowych w systemie ANFIS (3 zbiory rozmyte)

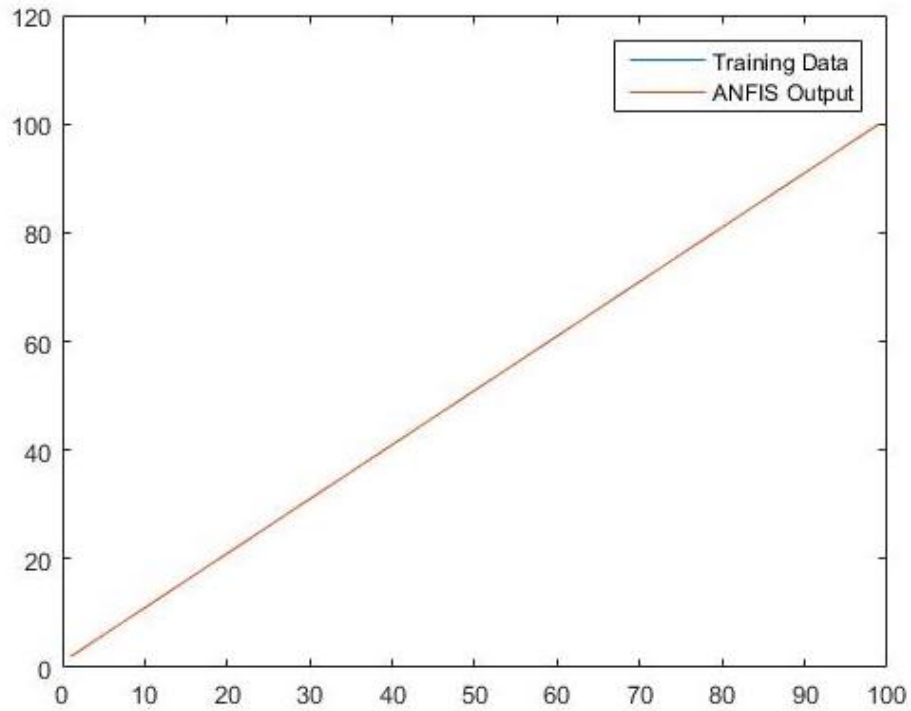


Rys. 5.12. Zmienność błędu w procesie uczenia prognozy dla danych testowych w systemie ANFIS (7 zbiorów rozmytych)

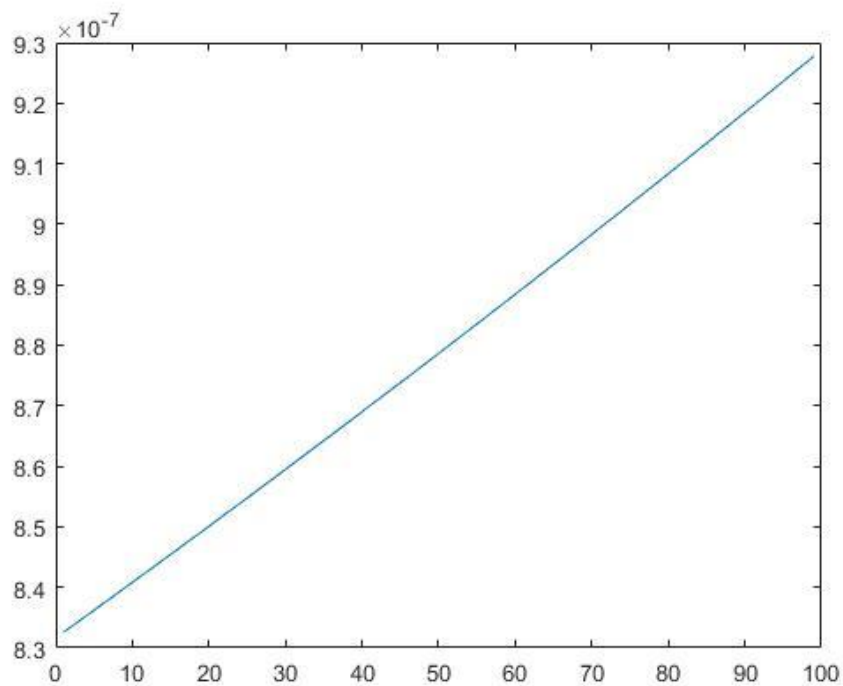
Następnie, dla tych samych danych testujących, przyjęto zbiory rozmyte o funkcjach przynależności typu gaussowskiego (rysunek 5.13.). Uzyskana charakterystyka wyjściowej jest również liniowa (rysunek 5.14.). Błąd względny, wynikający z procesu uczenia, zależny od punktu na charakterystyce jest rzędu  $9,3 \times 10^{-7}\%$  –  $1,4 \times 10^{-5}\%$ .



Rys. 5.13. Podział przestrzeni na zbiory rozmyte o funkcjach przynależności typu gaussowskiego

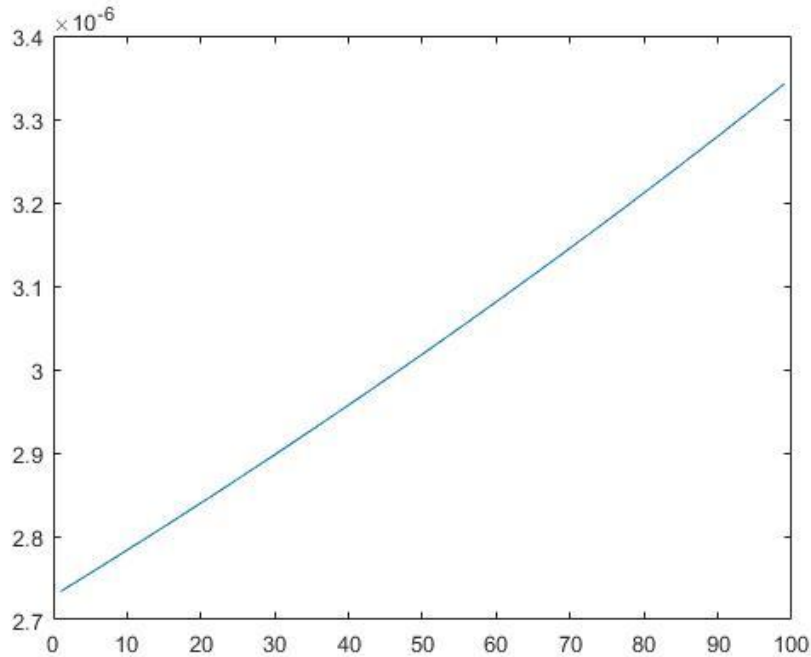


Rys. 5.14. Wartości prognozy (ANFIS) ciągu danych testowych na tle danych wejściowych przy funkcjach przynależności typu gaussowskiego



Rys. 5.15. Zmienność błędu w procesie uczenia prognozy dla danych testowych w systemie ANFIS (3 zbiory)





Rys. 5.16. Zmienność błędu w procesie uczenia prognozy dla danych testowych w systemie ANFIS (5zbiorów)

Można więc twierdzić że system ANFIS zapewnia liniową zależność sygnału wyjściowego od wejściowego w dużym zakresie zmian.

## 5.5. Identyfikacja ciągów czasowych za pomocą systemu neuronowo-rozmytego ANFIS

### Identyfikacja ciągów czasowych z zastosowaniem modelu ANFIS o jednym wejściu

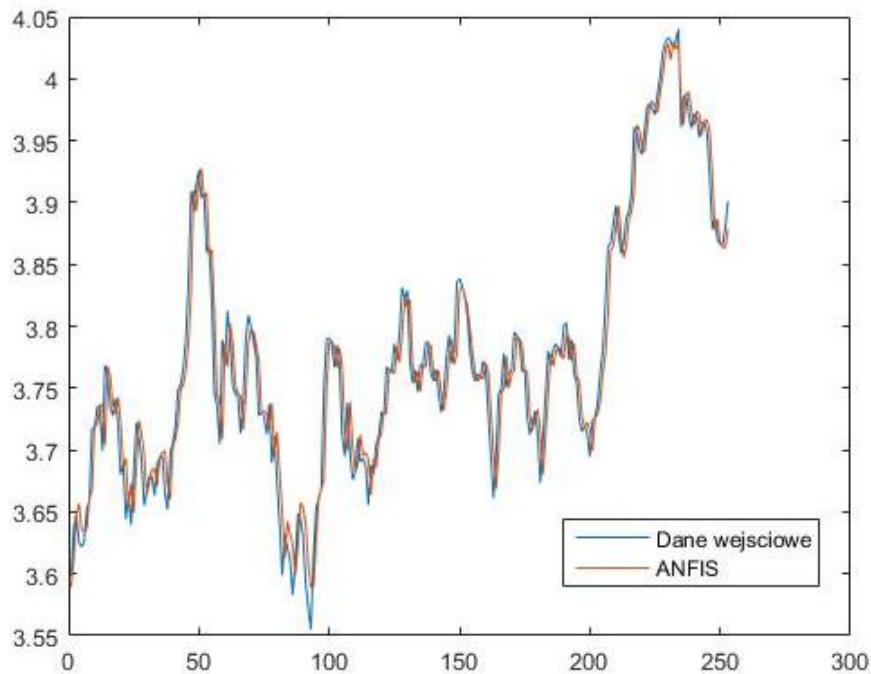
Model ANFIS o jednym wejściu odpowiada rozmytemu modelowi TSK o jednym wejściu, opisanemu w paragrafie 5.2.

Część systemu ANFIS realizująca proces uczenia jest układem sprzężenia zwrotnego. Układ ten dostosowuje parametry sieci neuronowej w taki sposób aby funkcja wyjściowa (5.18) jak najmniej różniła się od wartości ciągu uczącego  $\{x_n\}$  w każdym kroku  $n$ . W tym celu zmieniane są parametry zbiorów rozmytych w poprzednikach reguł (5.17) oraz parametry funkcji następników w tych regułach. W efekcie, wyjściowa funkcja (5.18) jest funkcją o parametrach zależnych od kroku  $n$ . Wykonane obliczenia identyfikacyjne ciągów czasowych pokazują wynik prognozy oraz zmienność błędów w procesie uczenia się systemu ANFIS. Wielkość i zmiana błędów zależą między innymi od liczby zbiorów rozmytych czyli od podziału przestrzeni wejścia (Kashpruk, 2018).

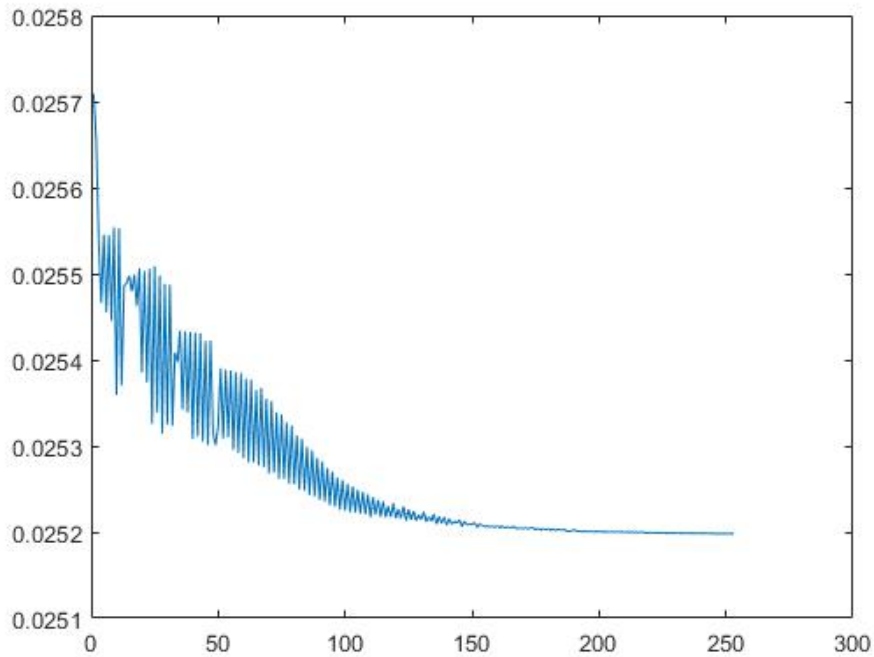
Model wypracowany przez system ANFIS nie jest znany z jawnej postaci.

Na rysunku 5.17. przedstawiono wyniki prognozy wyznaczone w systemie ANFIS dla ciągu notowań USD/PLN (jak dane w rozdziale 4.6.). Obliczenia powtarzano dla różnej

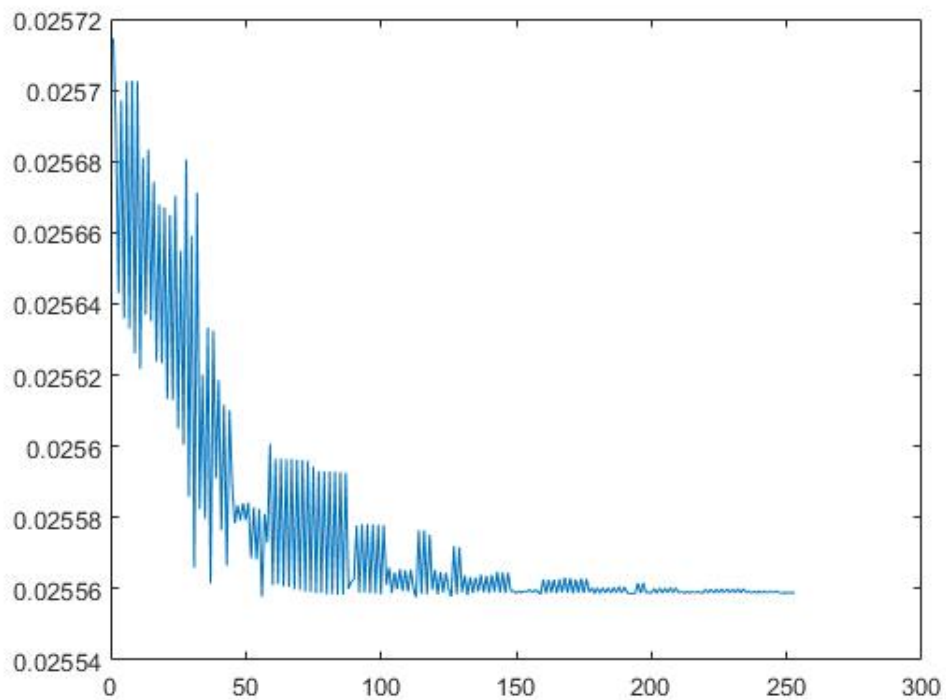
struktury modelu rozmytego TSK, zawierającego 3, 4, 5, 7 zbiorów rozmytych w przestrzeni danych wejściowych. Na rysunkach 5.18.–5.21. przedstawiono zmienność błędu bezwzględnego w procesie uczenia prognozy. Proces uczenia wymaga dużej liczby danych uczących, tym większej im większa liczba zbiorów rozmytych w strukturze modelu. Najmniejszy błąd, w stanie ustalonym uzyskuje się dla siedmiu zbiorów rozmytych.



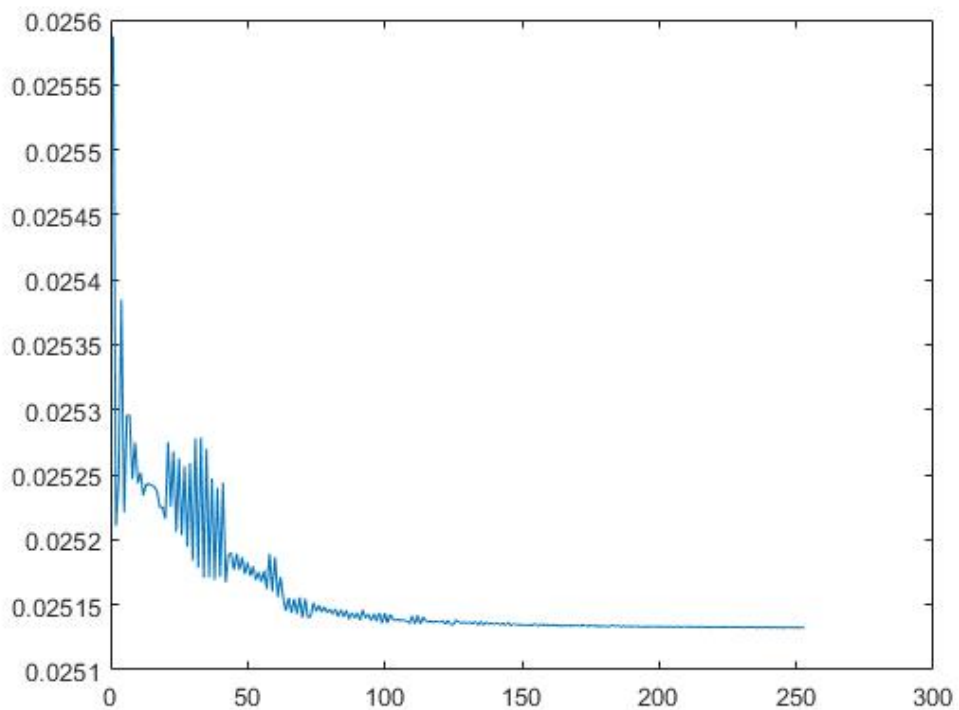
Rys. 5.17. Wartości prognozy (ANFIS) ciągu notowań USD/PLN na tle danych wejściowych



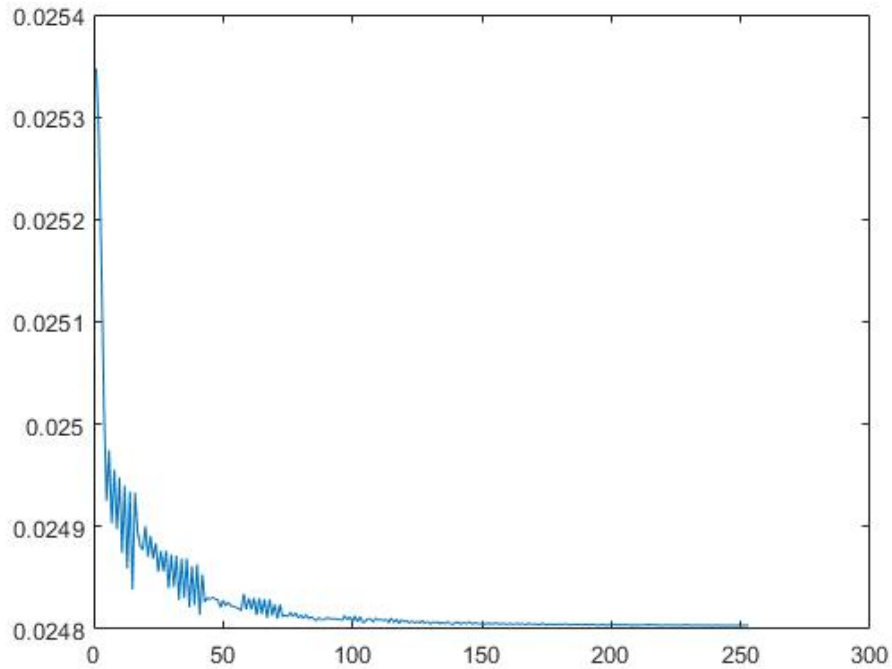
Rys. 5.18. Zmienność błędu w procesie uczenia prognozy dla notowań USD/PLN w systemie ANFIS (3 zbiory rozmyte)



Rys. 5.19. Zmienność błędu w procesie uczenia prognozy dla notowań USD/PLN w systemie ANFIS (4 zbiory rozmyte)

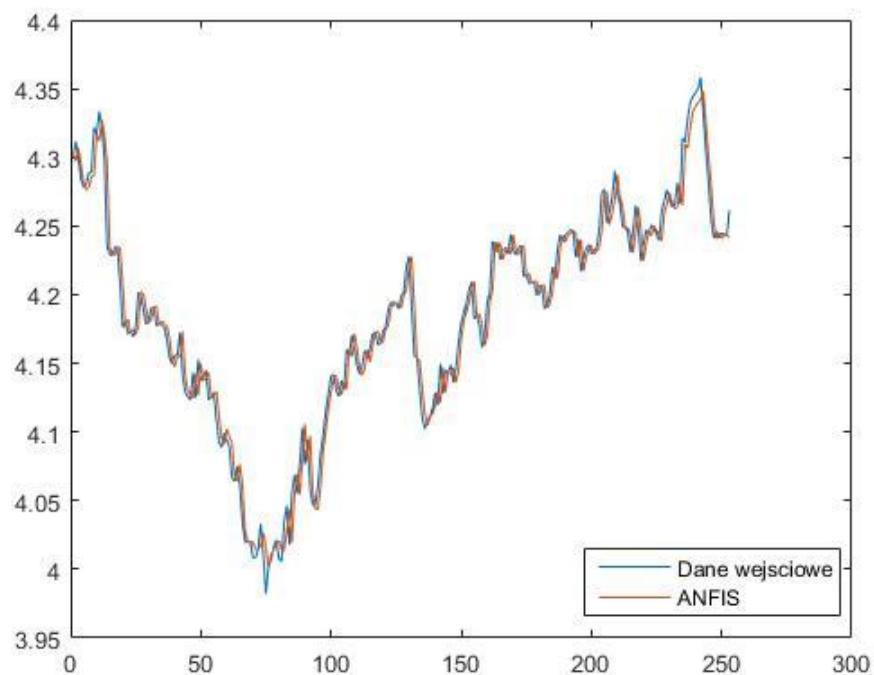


Rys. 5.20. Zmienność błędu w procesie uczenia prognozy dla notowań USD/PLN w systemie ANFIS (5 zbiorów rozmytych)

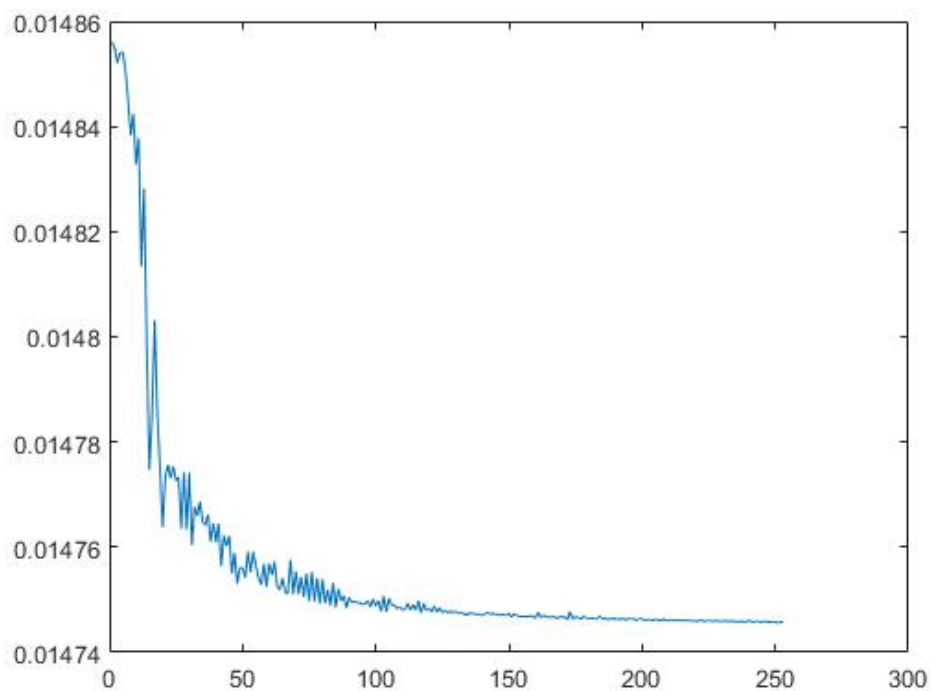


Rys. 5.21. Zmienność błędu w procesie uczenia prognozy dla notowań USD/PLN w systemie ANFIS (7 zbiorów rozmytych)

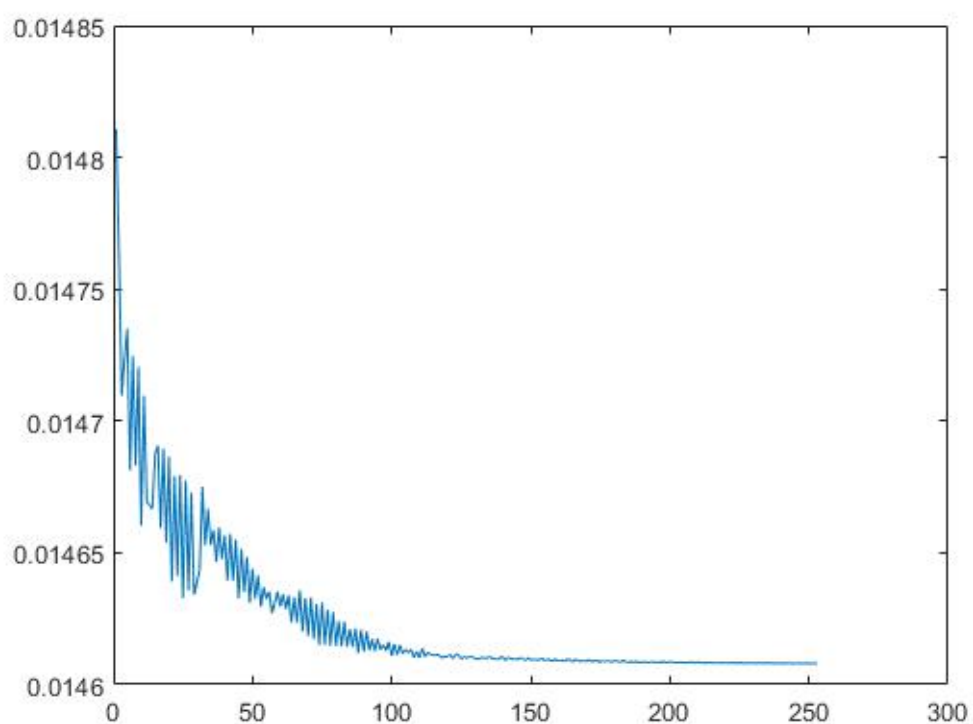
Analogicznie, jak dla notowań USD/PLN, wykonano obliczenia prognozy w systemie ANFIS dla notowań EUR/PLN (rysunek 5.22.). Podobnie też, zmieniano strukturę modelu rozmytego TSK, przyjmując 3,4,5,7,9 zbiorów w przestrzeni wejścia. zmienność błędu w procesie uczenia przedstawiono dla przykładu na rysunkach 5.23., 5.24. i 5.25. Błąd prognozy w procesie uczenia maleje w przybliżeniu, wykładniczo. Błąd w stanie ustalonym maleje ze wzrostem liczby zbiorów rozmytych.



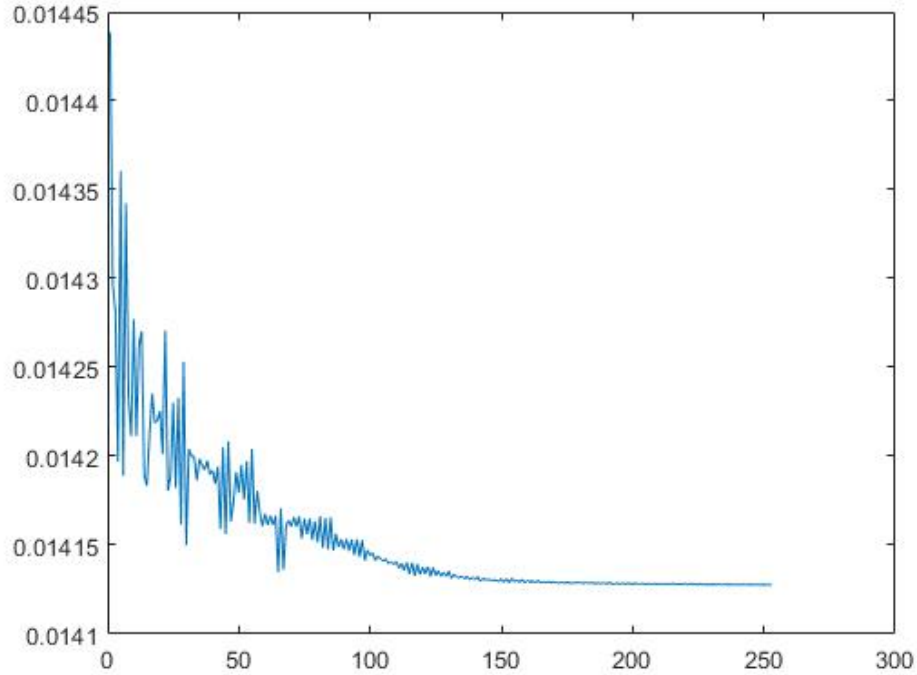
Rys. 5.22. Wartości prognozy (ANFIS) ciągu notowań EUR/PLN na tle danych wejściowych



Rys. 5.23. Zmienność błędu w procesie uczenia prognozy dla notowań EUR/PLN w systemie ANFIS (3 zbiory rozmyte)

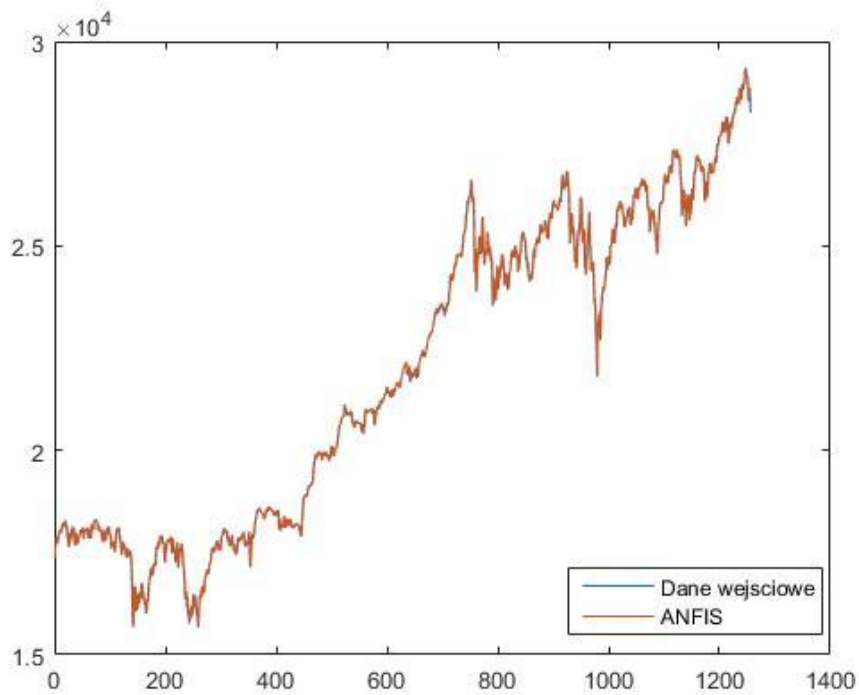


Rys. 5.24. Zmienność błędu w procesie uczenia prognozy dla notowań EUR/PLN w systemie ANFIS (4 zbiory rozmyte)

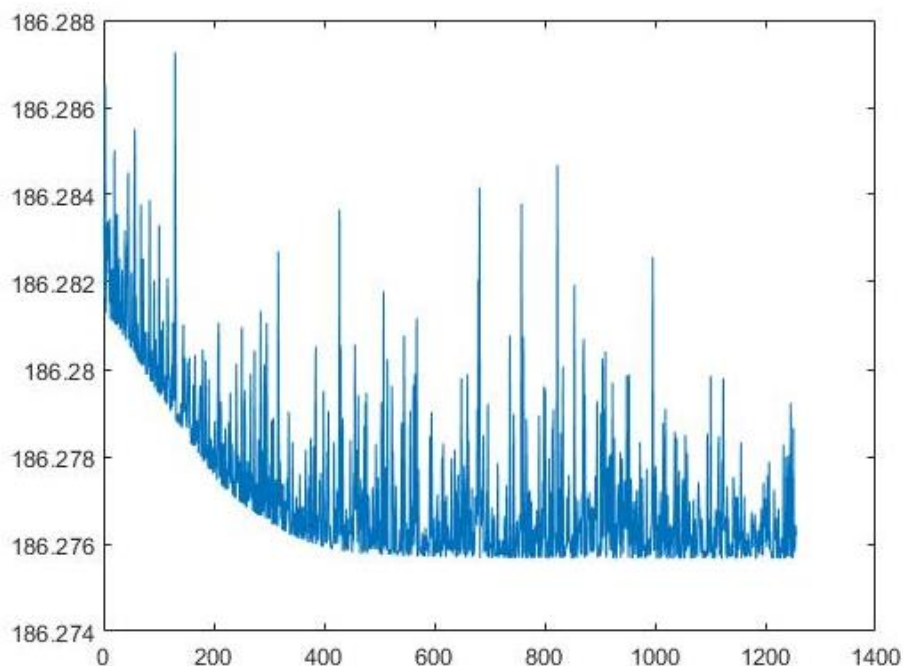


Rys. 5.25. Zmienność błędu w procesie uczenia prognozy dla notowań EUR/PLN w systemie ANFIS (9 zbiorów rozmytych)

Nie co inny charakter błędów obserwujemy przy procesie uczenia prognozy dla notowań DJIA (rysunki 5.26. i 5.27.). Przyczyna trudności w dopasowaniu modelu ANFIS do ciągu danych uczących jest ich duży zakres zmienności. Ciąg podstawowy notowań zmienia swoją wartość w granicach  $1.5$  do  $3 \times 10^4$ .

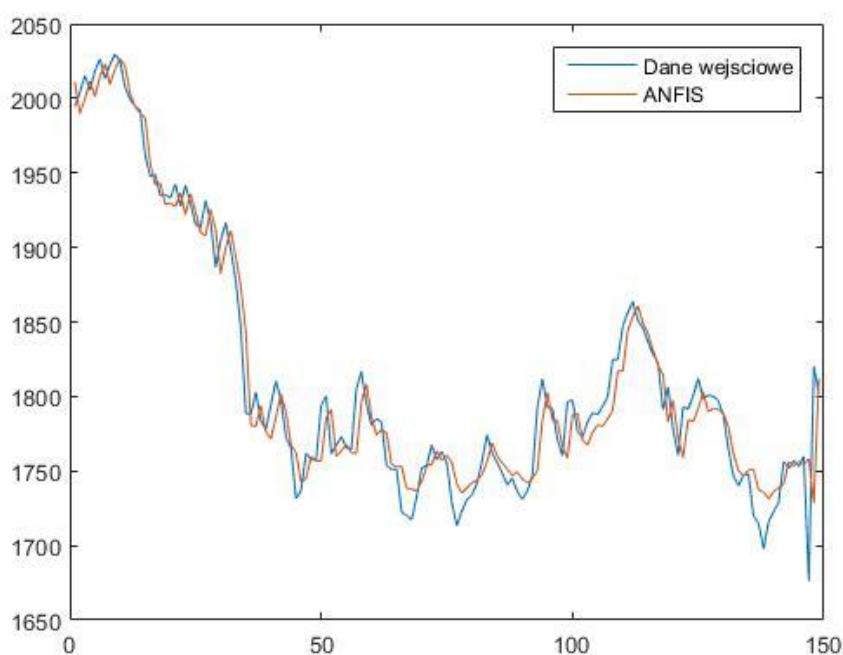


Rys. 5.26. Wartości prognozy (ANFIS) ciągu notowań DJIA na tle danych wejściowych

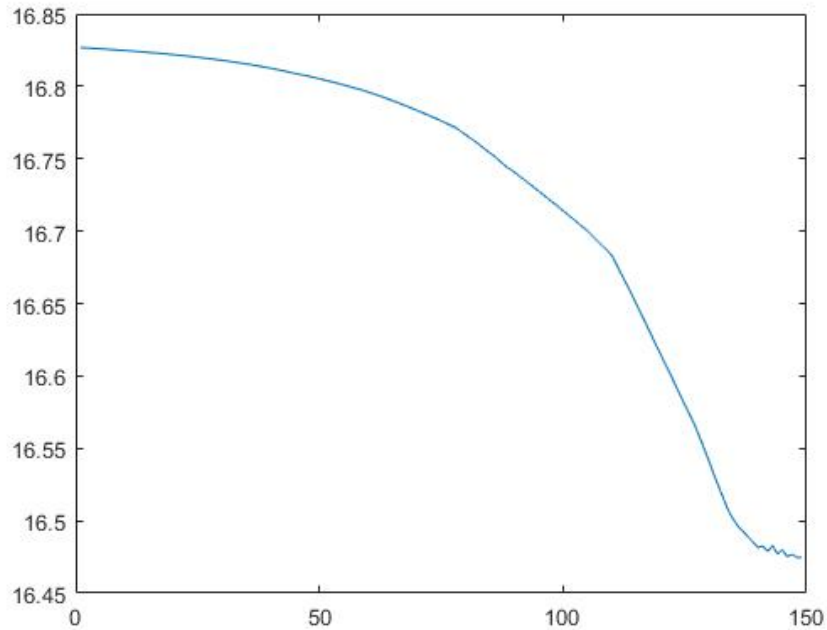


Rys. 5.27. Zmienność błędu w procesie uczenia prognozy dla notowań DJIA w systemie ANFIS (3 zbiory rozmyte)

Wyniki badań identyfikacyjnych dla ciągu notowań wskaźnika polskiego rynku WIG20 przedstawiono na rysunkach 5.28. i 5.29. Charakter zmian błędu w postaci gładkiej malejącej funkcji zaobserwowano dla 3wszystkich badanych struktur modelu zawierających: 3,4,5 i 7 zbiorów w przestrzeni zmiennej wejściowej. Wartość błędu względnego nie przekracza 8 %. Dla struktury z siedmioma zbiorami rozmytymi nie przekracza 8 %.

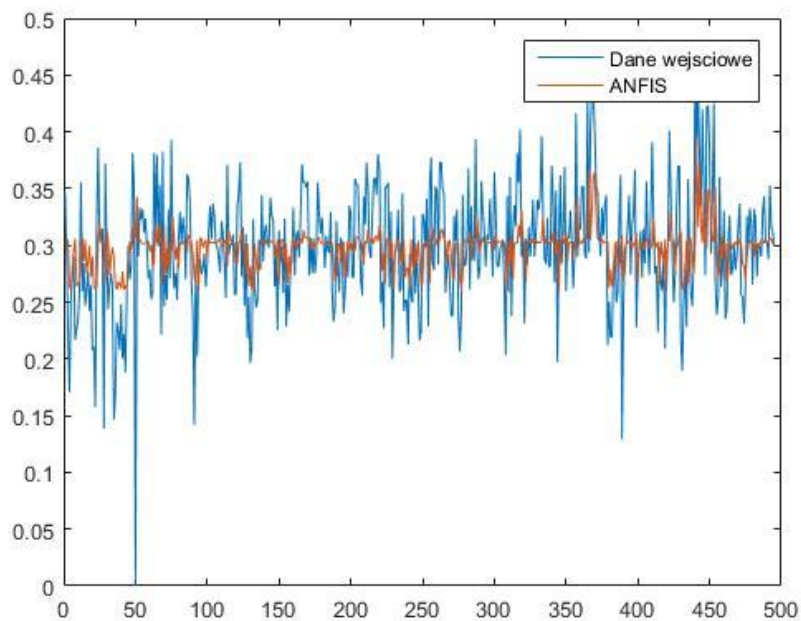


Rys. 5.28. Wartości prognozy (ANFIS) ciągu notowań WIG20 na tle danych wejściowych



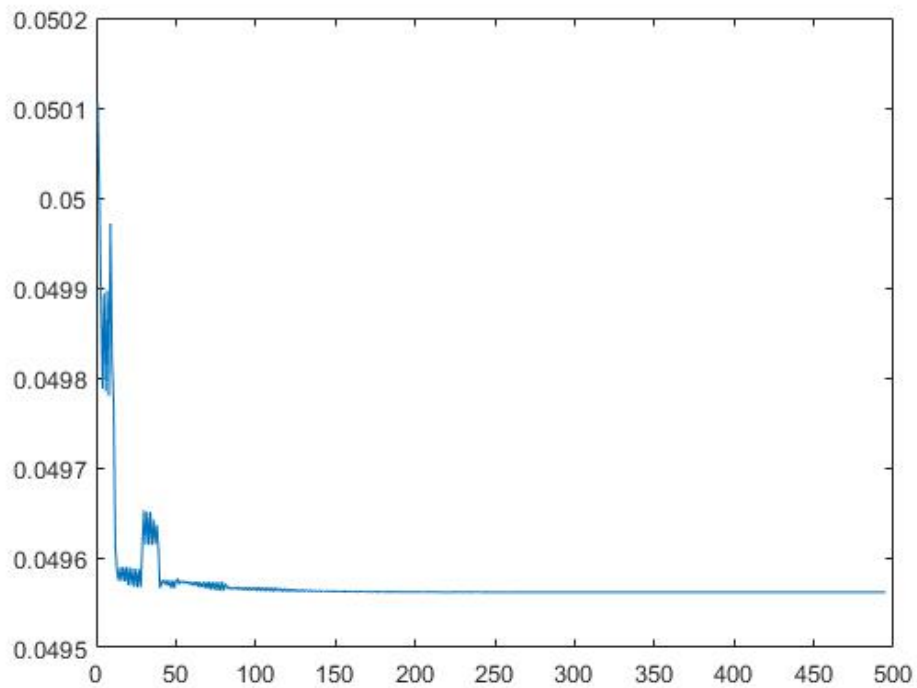
Rys. 5.29. Zmienność błędu w procesie uczenia prognozy dla notowań WIG20 w systemie ANFIS (7 zbiorów rozmytych)

Bardzo ciekawy przykład prognozy otrzymujemy dla ciągu obserwacji dwóch zmiennych  $X_4$  i  $X_1$  procesu technologicznego przedstawionego w paragrafie 4.6. Zmienność tych procesów ma charakter procesu białego szumu. Prognoza zrealizowana w systemie ANFIS jest bliska wartości średniej z niewielkimi wahaniami (rysunki 5.30. i 5.33.). Na rysunkach 5.31. i 5.32 oraz 5.34. i 5.35. przedstawiono błędy prognozy w trakcie uczenia dla poszczególnych procesów. Widzimy że błędy maleją wykładniczo w czasie 100 do 150 kroków. Wartość ustalona błędu maleje dla większej liczby zbiorów w strukturze modelu ANFIS. Jednak że wartości błędów pozostają znaczne, 15 do 20%.

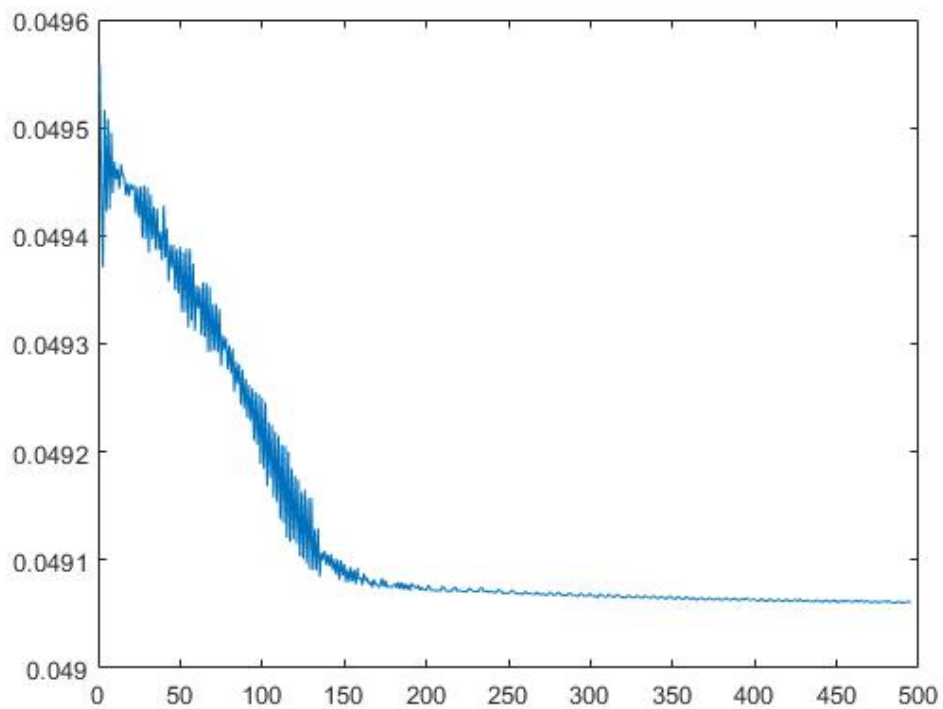


Rys. 5.30. Wartości prognozy (ANFIS) ciągu notowań  $X_4$  na tle danych wejściowych

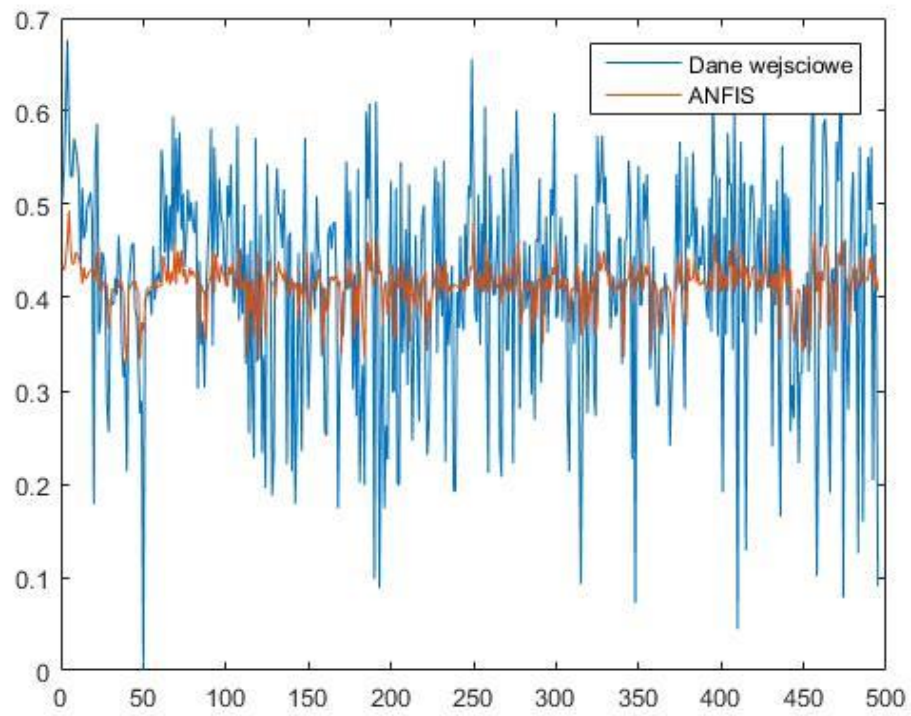




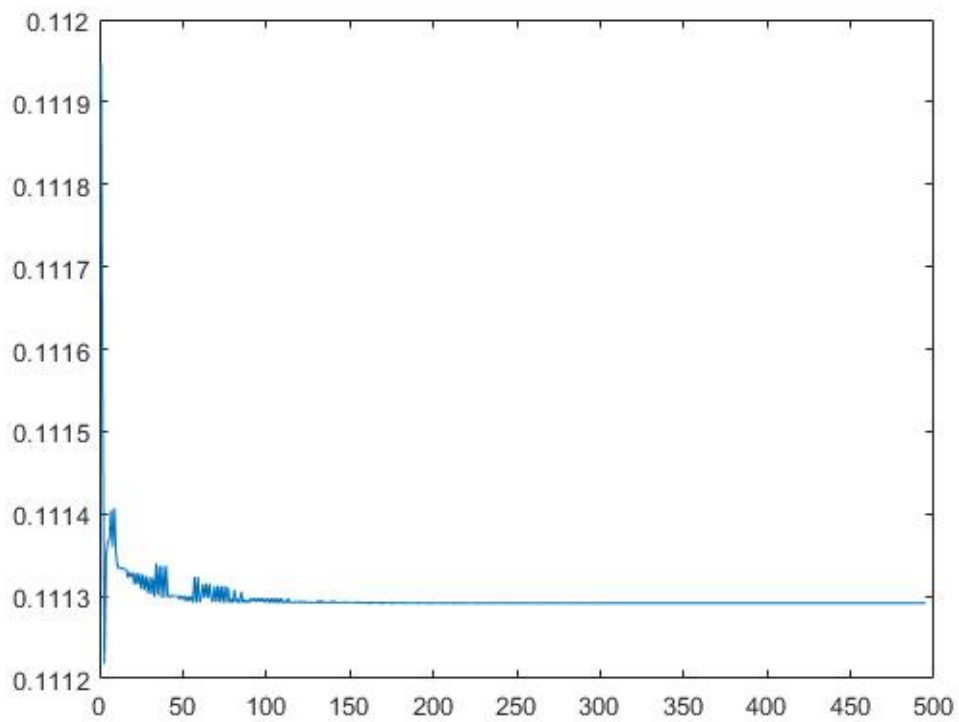
Rys. 5.31. Zmienność błędu w procesie uczenia prognozy dla notowań X4 w systemie ANFIS (3 zbiorów rozmytych)



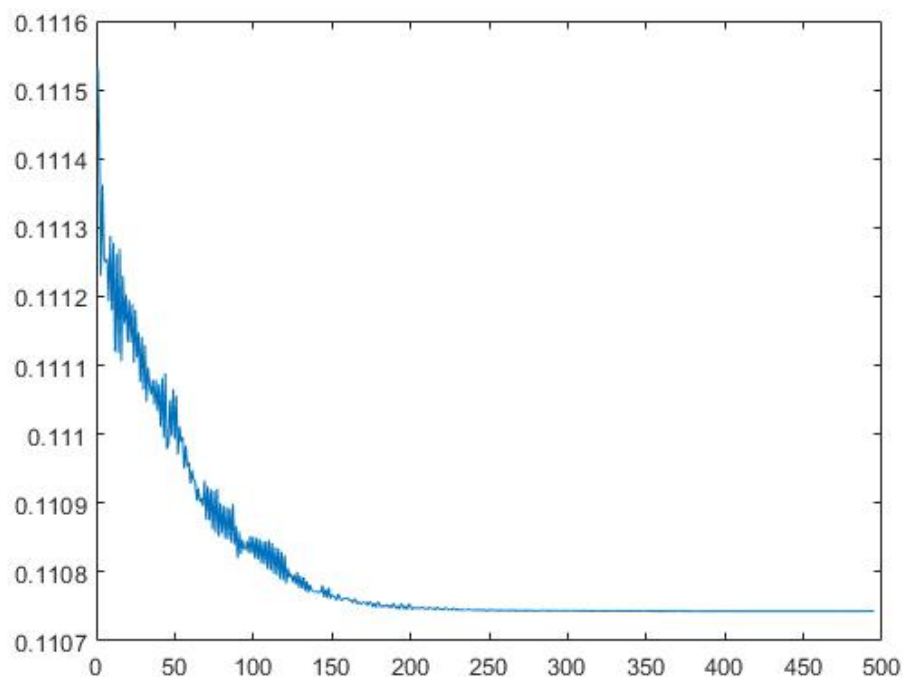
Rys. 5.32. Zmienność błędu w procesie uczenia prognozy dla notowań X4 w systemie ANFIS (7 zbiorów rozmytych)



Rys. 5.33. Dane wejściowe i ANFIS dla X1



Rys. 5.34. Zmienność błędu w procesie uczenia prognozy dla notowań X4 w systemie ANFIS (3 zbiory rozmyte)



Rys. 5.35. Zmienność błędu w procesie uczenia prognozy dla notowań X4 w systemie ANFIS (5 zbiorów rozmytych)

Pełny wykaz błędów od ilości zbiorów rozmytych przedstawiono w tabeli 5.1.

Tab. 5.1. Zależność błędów od ilości zbiorów rozmytych

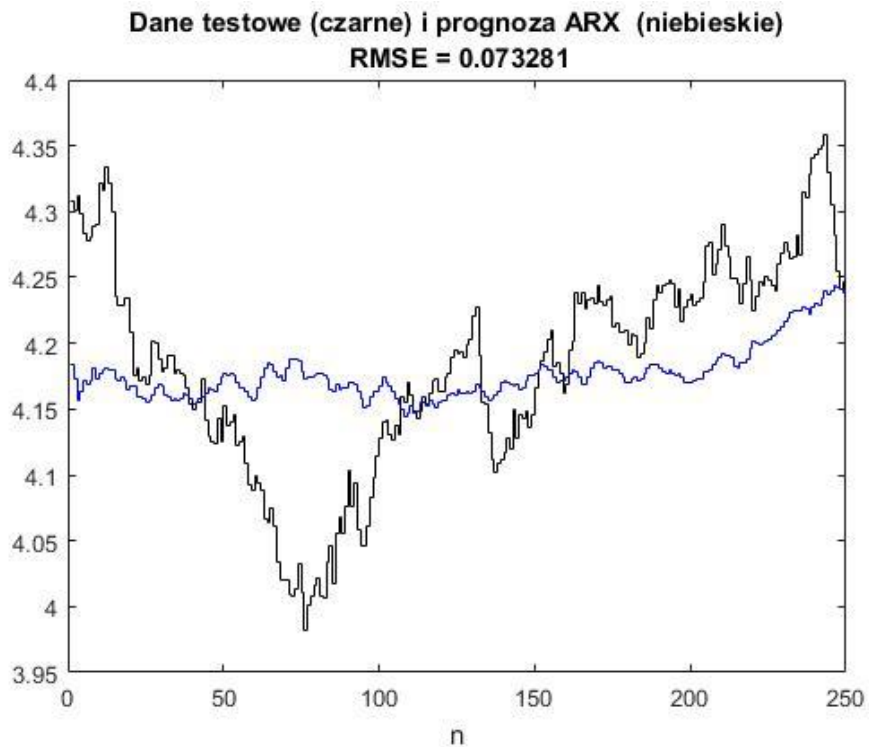
Dane	Liczby zbiorów rozmytych			
	3	4	5	7
USD2015	0,0252	0,0255	0,0251	0,0248
EUR2015	0,0147	0,0146	0,0144	0,0142
WIG20	19,26	18,07	17,62	16,47
DJIA	186,276	186,168	185,98	185,652
X1	0,1113	0,1111	0,1108	0,1098
X4	0,0496	0,0495	0,0494	0,0491

Liczba zbiorów rozmytych dzielących przestrzeń zmiennych wejściowych, a tym samym liczba reguł jest decydująca dla błędu prognoz

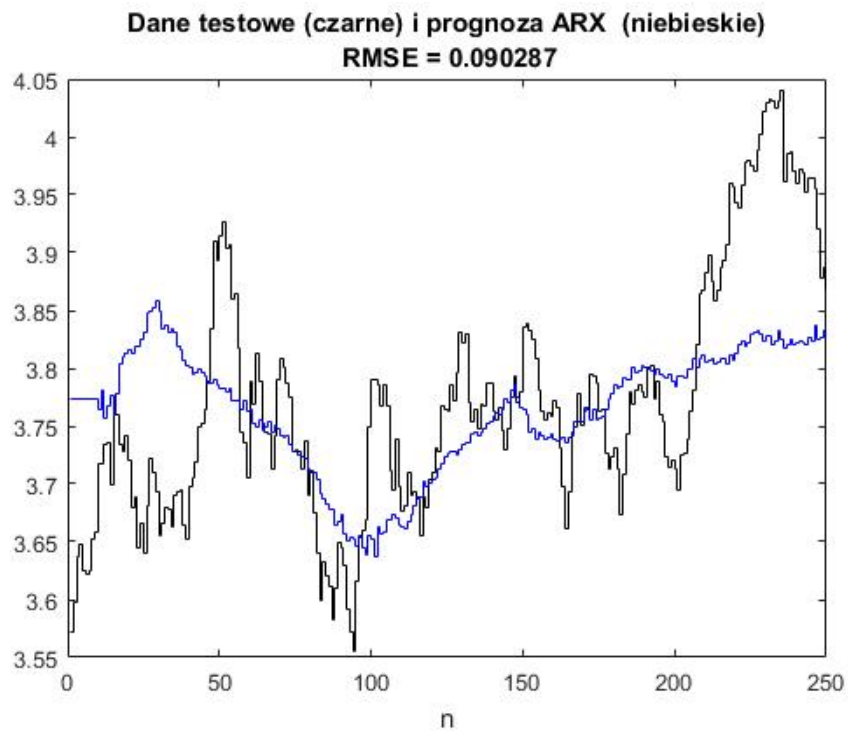
### Identyfikacja obiektów dynamicznych z zastosowaniem modelu ANFIS o dwóch wejściach

Jednym z obiektów modeli dynamicznych, jest model ARX. Wyjście obiektu w chwili  $n$ -tej jest liniowo zależne od poprzednich swoich wartości oraz opóźnionego wejścia. Do modelowania tego typu obiektów wybrano ciągi losowe obserwacji przedstawione w rozdziale 4. Prognoza jest więc wyliczana jako kombinacja liniowa poprzednich sygnałów wejścia i

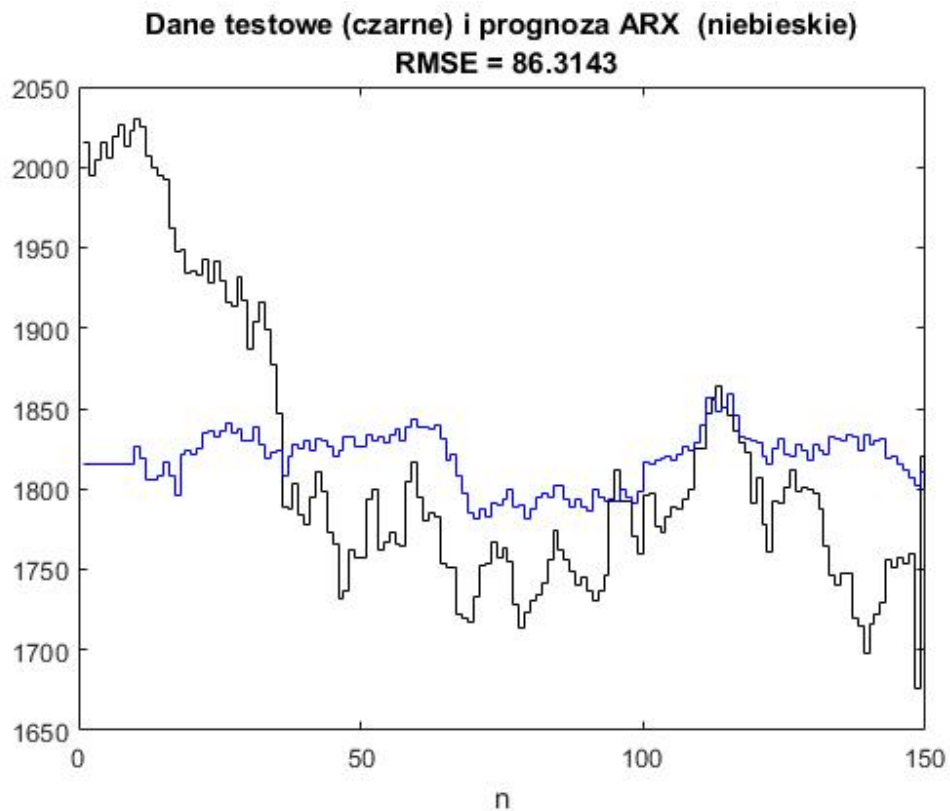
wyjścia. Na rysunkach 5.36. do 5.40. przedstawiono wyznaczone w systemie ANFIS przebiegi prognoz dla różnych kombinacji ciągów obserwacji.



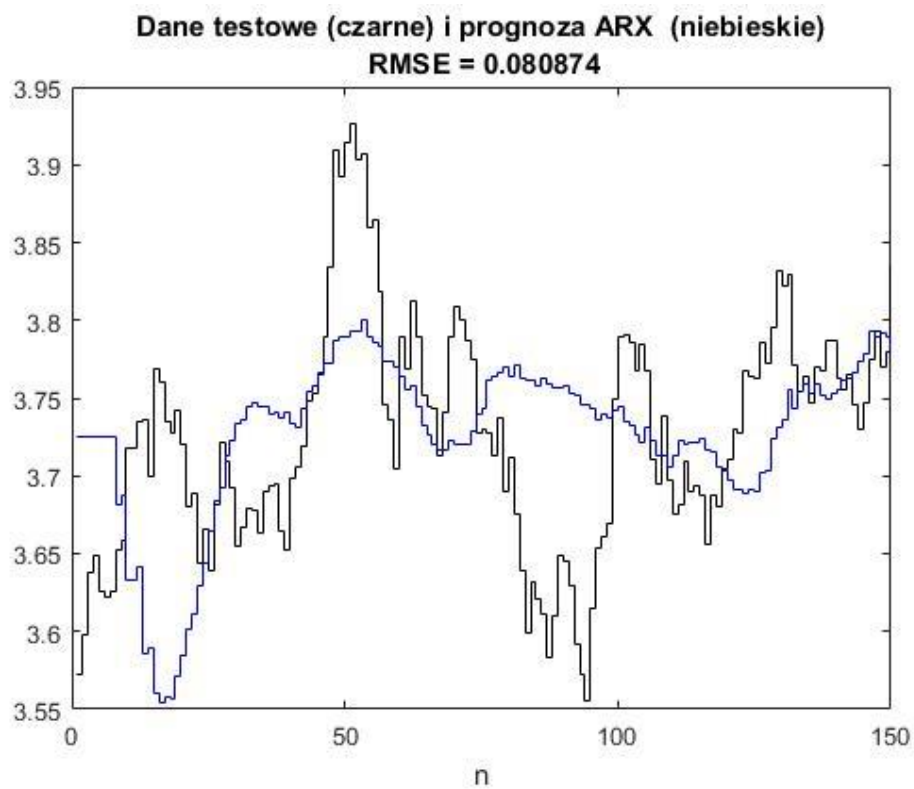
Rys. 5.36. Przebieg notowań USD/PLN prognozowanych według modelu ARX w zależności od notowań euro



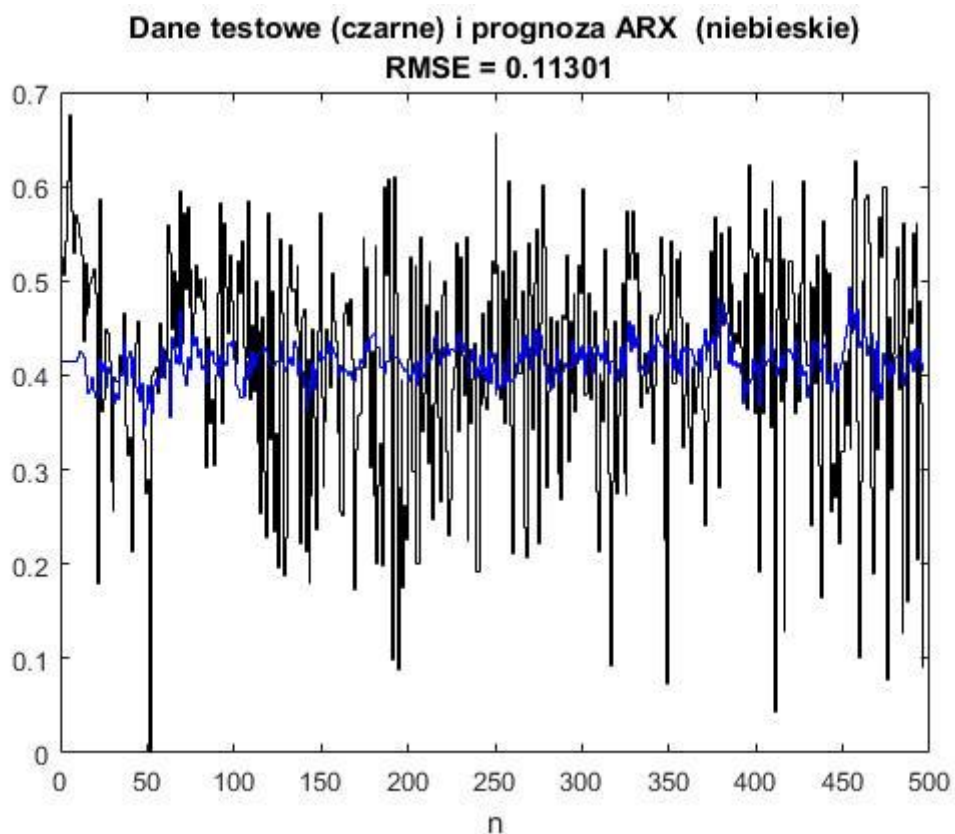
Rys. 5.37. Przebieg notowań EUR/PLN prognozowanych według modelu ARX w zależności od notowań dolara



Rys. 5.38. Przebieg notowań USD/PLN prognozowanych według modelu ARX w zależności od notowań WIG20



Rys. 5.39. Przebieg notowań WIG20 prognozowanych według modelu ARX w zależności od notowań dolara



Rys. 5.40. Przebieg prognoz zmiennej  $X_1$ (koncentratu) według modelu ARX w zależności od zmiennej  $X_4$

Model ARX jest w pewnym sensie średnią ważoną pewnej liczby kroków dwóch ciągów czasowych. Takie uśrednienie jest dobrze widoczne na zamieszczonych wykresach.

## 6. PODSUMOWANIA I DALSZY PLANY BADAWCZE

### 6.1. Podsumowania metod stosowanych w pracy

Modele stochastyczne ciągów czasowych mają długą tradycję w stosowaniu w wielu dziedzinach. Za ich początek można uważać powstanie pierwszych rynków finansowych, kiedy pojawiły się pierwsze próby przewidywania zysku i ochrony kapitału od zmian na rynku. Ten problem jest nadal bardzo aktualny. W obecnym czasie istnieje wiele organizacji które starają się przewidzieć zmiany na rynku finansowym. Do takich instytucji należą banki, fundusze inwestycyjne, spółki i organizacje finansowe. Bardzo popularnymi metodami są narzędzia prognozowania w czasie rzeczywistym zmian kursów akcji, obligacji i kursów zmiany walut, gdzie głównym celem jest kupowanie i sprzedawanie w celu optymalizacji zysku.

W szerokiej gamie zjawisk, problem prognozowania rozwiązuje się najczęściej przy stosowaniu metod prognozowania opartych na zastosowaniu szeregów czasowych z jedną zmienną tj. AR, ARMA, ARIMA. i inne.

Najważniejszym krokiem w wyborze odpowiedniego modelu jest określenie optymalnych parametrów tego modelu. Jednym z kryteriów jest to, że próbka ACF i PACF, obliczona na podstawie danych treningowych, musi odpowiadać odpowiednim wartościom. Inne dobrze znane kryteria informacyjne identyfikacji modeli AIC i BIC. W modelach statystycznych wartość prognozy jest wartością wyjścia modelu, czyli wartością funkcji wyznaczonej dla aktualnych obserwacji.

Modele miękkie (ang. *soft computing*) mają dużą swobodę budowania modeli ze względu na naturalną nieliniowość. Te podejście jest stosowane do rozwiązywania zadań poprzez rozpoznawanie wzorców do danych wejściowych, uczenie na podstawie doświadczenia i przedstawianie uogólnionych wyników w oparciu o wcześniejszą wiedzę. Do obliczeń miękkich zalicza się takie metody i techniki, jak: zbiory rozmyte, zbiory przybliżone, sztuczne sieci neuronowe, algorytmy genetyczne i szereg innych.

Modele rozmyte i neuronowo-rozmyte są tworzone w oparciu o reguły wiedzy eksperckiej i nie są modelami o jawnej postaci analitycznej. Często stanowią one część większego zadania ujętego w postaci systemu eksperckiego, jako knowledge-based system. W rozmytym modelowaniu regułowym, każda reguła stanowi model lokalny, zbudowany na pewnej części zbioru danych, a procedura agregacji spełnia rolę rozmytej tranzycji. Ze względu na postać bazy wiedzy i zasadę wnioskowania rozróżnia się dwa główne typy systemów rozmytych: lingwistyczne systemy typu Mamdaniego i rozmyte systemy Takagi-Sugeno-Kanga. Modele szeregów czasowych realizuje się numerycznie, zaś modele rozmyte można modelować lingwistycznie. Modele wnioskowania Mamdaniego charakteryzują się tym, że występujące koniunkcje, implikacje i alternatywy rozmyte zastępowane są operatorami *minimum* i *maximum*. Lingwistyczne modele rozmyte z wnioskowaniem typu Mamdaniego

oraz z wnioskowaniem uproszczonym dają, co do zasady, nieliniowe zależności typu wejście-wyjście ze względu na specyfikę procedur wnioskowania i defuzyfikacji.

W przypadku modeli progowych TAR każdy model AR(k) jest obliczany dla pewnej części przestrzeni  $\chi$ . O dokładności modelu TAR decydują takie czynniki jak: zbiór danych wejściowych, trafność podziału przestrzeni  $\chi$  na podzbiory, w których liczone są poszczególne modele oraz wybór klas modeli. Zarówno w modelach AR, jak i TAR decydujące znaczenie może mieć wiedza eksperta w zakresie wyboru klasy AR(k) oraz podziałem przestrzeni dla modeli TAR. Dopasowanie modeli do danych doświadczalnych, a więc także uporządkowanie w klasach modeli wprowadzają kryteria identyfikacji reszt. Modele TAR, STAR pozwalają na modelowanie skomplikowanych nieliniowych zjawisk, ale problemem jest w tym przypadku określenie nieliniowości.

Rozmyte modele regułowe typu TSK są najbardziej zbliżone do modeli TAR i STAR. Modele lokalne, liniowe, reprezentowane przez poszczególne reguły są agregowane jako suma ważona, przy czym współczynniki wagi nie są stałe, lecz są zależne od zmiennej wejściowej. Aby uzyskać liniowe odwzorowanie w modelach lokalnych TSK, wskazane jest zastosowanie trapezoidalnych zbiorów rozmytych w przestrzeni zmiennej wejściowej. W modelach rozmytych wartość prognozy jest wynikiem wnioskowania rozmytego wynikającego z kompatybilności aktualnej obserwacji i rozmytych przesłanek. Modele TSK i STAR są do siebie bardzo podobne, ale TSK mają lepiej określony sposób kształtowania nieliniowości.

W praktyce tworzenia informatycznych inteligentnych systemów modelujących panuje tendencja do tworzenia systemów hybrydowych, na przykład neuronowo-rozmytych, rozmytych z zastosowaniem algorytmów genetycznych optymalizujących model rozmyty itp. System ANFIS, który łączy elastyczność modeli rozmytych z automatycznym doбором funkcji przynależności to naturalny rozwój obydwu metodologii. Przedstawiona procedura wnioskowania, która została przedstawiona w Rozdziale 5, jest realizowana w myśl zasady systemu TSK, czyli konkluzja systemu jest sumą ważoną konkluzji nierozmytych poszczególnych reguł. Wagi tej sumy są wyznaczone w oparciu o stopnie przynależności numerycznych wartości sygnałów wejściowych.

ANFIS jest siecią uczącą się adaptacyjnie. W trakcie procesu uczenia mogą być iteracyjnie modyfikowane parametry reguł. Współczynnik uczenia zależy od błędu średniokwadratowego. W przypadku stosowania techniki ANFIS dokładność uzyskanej prognozy można poprawić, poprzez wprowadzenie podziału przestrzeni wejścia na więcej obszarów, czyli poprzez zwiększenie liczby reguł. Zbiór reguł rozmytych ukrytych w modelu ANFIS jest tworzony i optymalizowany automatycznie dzięki operacji sprzężenia zwrotnego realizowanego przez sieć adaptacyjną. Istotną wadą tego modelu jest brak modelowania niepewności.

Do współczesnych metod prognozowania należy system Prophet który łączy skomplikowany aparat matematyczny z możliwościami modelowania niepewności. Zaletami prognozowania z zastosowaniem Facebook Prophet jest szybkość wyznaczania modelu i łatwość prognozowania. Ten pakiet jest stosunkowo łatwy w użyciu, nauce i analizie, więc jest najlepszym rozwiązaniem w zadaniach wymagających szybkiej prognozy i nie wymagających najlepszej dokładności.



## 6.2. Wkład własny

Do oryginalnych osiągnięć pracy autorka zalicza:

- Przedstawiona krytyczna analiza dwóch różnych podejść do budowy i identyfikacji modeli ciągów czasowych: klasycznego podejścia uwzględniającego modele statystyczne oraz podejścia wykorzystującego obliczenia miękkie.
- Opracowano metodę strojenia modeli TAR – przedstawiono schemat doboru modeli oraz pokazano strojenie modelu TAR na przykładzie danych losowych.
- Modelowanie zmian kursu USD – przeprowadzono badania prognozy kursu USD w 2015 roku stosując modele statystyczne (AR, ARMA, ARIMA) oraz stosując modele rozmyte.
- Modelowanie zawartości poszczególnych frakcji węgla – zrobiono podbór parametrów AiC I BIC dla wyboru najlepszego modelu ARIMA dla prognozowania procesu technologicznego. Ponadto dla porównania zastosowano model ARX.
- Prognozowanie efektywności ekonomicznej pakietem Prophet, który pozwala na tworzenie dokładnych modeli predykcyjnych w trybie półautomatycznym. Dla przykładu przedstawiono prognozę sygnałów poboru Mocy i Stanu energii elektrycznej z liczników.
- Prognozowanie modelem ANFIS - wykonane obliczenia identyfikacyjne ciągów czasowych pokazują wynik prognozy oraz zmienność błędów w procesie uczenia się systemu ANFIS. Wielkość i zmiana błędów zależą między innymi od liczby zbiorów rozmytych czyli od podziału przestrzeni wejścia. Model wypracowany przez system ANFIS nie jest znany z jawnej postaci.

### 6.3. Kierunki przyszłych badań

Tematykę planowanej pracy badawczej przedstawiono w następujących punktach:

- Kontynuacja badań związanych z modelowaniem niepewności, w których zastosowane zostaną zaawansowane metody sztucznej inteligencji, oraz rozmyte bazy wiedzy dla systemów stochastycznych;
- Przeprowadzenie obliczeń przy wykorzystaniu metody Monte Carlo;
- Metody Bayesowskie, stosowane są dla wyznaczenia wiarygodności procesów i zdarzeń które opisują prawdopodobieństwo zdarzenia na podstawie okoliczności, które mogą być związane z tym wydarzeniem. Ta metoda ma liczne zastosowania w szerokim zakresie obliczeń probabilistycznych, nie tylko we wnioskowaniu Bayesowskim;
- Funkcjonalna analiza danych.

## SPIS LITERATURY

Acakpovi, A., Ternor, A. T., Asabere N.Y., Adjei P., Iddrisu A., (2020). Time Series Prediction of Electricity Demand Using Adaptive Neuro-Fuzzy Inference Systems, *Mathematical Problems in Engineering*, vol. 2020, ArticleID: 4181045. doi.org/10.1155/2020/4181045.

Adedeji, P.A., et al, (2020). Hybrid adaptive neuro-fuzzy inference system (ANFIS) for a multi-campus university energy consumption forecast. *International Journal of Ambient Energy*. doi.org/10.1080/01430750.2020.1719885.

Akaike, H., (1973). Information theory and an extension of the maximum likelihood principle. W: B. N. Petrov i F. Csaki (red.), *Second International Symposium on Information Theory*. Budapeszt: Akademiai Kiado.

Al-qaness, M.A., et al, (2020). Optimization method for forecasting confirmed cases of COVID-19 in China. *J. Clin. Med.* 9, 674. doi.org/10.3390/jcm9030674.

Anastassopoulou, C., Russo, L., Tsakris, A., Siettos, C., (2020). Data-based analysis, modelling and forecasting of the COVID-19 outbreak. *PLoS One* 15, e0230405. doi.org/10.1371/journal.pone.0230405.

Aznarte J.L., Benítez J.M., (2013). The Links Between Statistical and Fuzzy Models for Time Series Analysis and Forecasting. In: Pedrycz W., Chen SM. (eds) *Time Series Analysis, Modeling and Applications*. Intelligent Systems Reference Library, vol 47. Springer, Berlin, Heidelberg.

Bails, Dale G. and Peppers, Larry C., (1982). *Business Fluctuations; Forecasting Techniques and Applications*. Englewood Cliffs, New Jersey: Prentice Hall Inc.

Behnood, A., Golafshani, E. M. i Hosseini, S.M., (2020). Determinants of the infection rate of the COVID-19 in the U.S. using ANFIS and virus optimization algorithm (VOA), *Chaos, Solitons & Fractals*, Vol.139, 110051, ISSN 0960-0779. doi.org/10.1016/j.chaos.2020.110051.

Benvenuto, D., Giovanetti, M., Vassallo, L., Angeletti, S., and Ciccozzi, M., (2020). Application of the ARIMA model on the COVID-2019 epidemic dataset, *Data in Brief*, Volume 29. doi.org/10.1016/j.dib.2020.105340.

Box, G. Jenkins, G., (1983). *Analiza Szeregów Czasowych. Prognozowanie i Sterowanie*, PWN, Warszawa.

- Box, G., and Jenkins, G., (1970). *Time Series Analysis: Forecasting and Control*. San Francisco: Holden-Day.
- Box, G., Jenkins, G., Reinsel, G. and Ljung, G., (2016). *Time Series Analysis: Forecasting and Control*. 5th ed. John Wiley & Sons.
- Brockwell, P.J. and Davis, R.A., (1996). *Introduction to Time Series and Forecasting*. Springer-Verlag, New York.
- Brockwell, P.J., Davis, R.A., (2016). *Modeling and Forecasting with ARMA Processes*. In: *Introduction to Time Series and Forecasting*. Springer Texts in Statistics. Springer, Cham. doi.org/10.1007/978-3-319-29854-2\_5.
- Bubnicki, Z., (1974). *Identyfikacja Obiektów Sterowania*. Państwowe Wydawnictwo Naukowe, Warszawa.
- Castillo, O. and Melin, P., (2020). Forecasting of COVID-19 time series for countries in the world based on a hybrid approach combining the fractal dimension and fuzzy logic, *Chaos, Solitons & Fractals*, Vol.140. doi.org/10.1016/j.chaos.2020.110242.
- Chan, K.S., (1993). Consistency and Limiting Distribution of the Least Squares Estimator of a Threshold Autoregressive Model, *The Annals of Statistics*, vol. 21, 520–533.
- Chan, Ngai, (2011). *Autoregresyjne modele średniej ruchomej*. doi.org/10.1002/9781118032466.ch3.
- Chappell, D., and Eldridge, R.M., (1997). Non-Linear Characteristics of the Sterling / European Currency Unit Exchange Rate: 1984-1992. *The European Journal of Finance*, 3 (2), 159-182.
- Chatfield, C., (1975). *The Analysis of Time Series. Theory and Practice*. Chapman and Hall, London, 263p.
- Chen, J., (2019). *Autoregressive Integrated Moving Average (ARIMA)*. Investopedia. Updated Apr 13. www.investopedia.com/terms/a/autoregressive-integrated-moving-average-arima.asp (18.09.2020 g. 12.42).
- Chen, J., Chen, Z., (2008). Extended Bayesian information criteria for model selection with large model spaces. *Biometrika*. 95 (3): 759–771. doi.org/10.1093/biomet/asn034.
- Chen, N., Dai, T., Wang, L., Zhao, W., Lu, K., (2018). Overload Analysis of Distribution Transformers Based on Data Mining. *IOP Conference Series: Materials Science and Engineering* vol. 439, 032112. doi.org/10.1088/1757-899x/439/3/032112.
- Chen, R., Tsay R., (1993). Functional-Coefficient Autoregressive Models. *Journal of the American statistical association* 88, 298-308.

- Cory Maklin, (2019). ARIMA Model Python Example — Time Series Forecasting, [towardsdatascience.com/machine-learning-part-19](https://towardsdatascience.com/machine-learning-part-19)
- Czogała, E., Pedrycz W., (1985). Elementy i metody teorii zbiorów rozmytych. PWN.
- Dash, P.K., Dash, S., Liew, A.C., Rahman, S., (1995). A Fuzzy Engineering Approach for Time Series Forecasting of Electric Load. *J. Intell. Fuzzy Syst.* Vol. 3, 261–271.
- Dhiman, N., and Sharma, M., (2019): Mediative Sugeno's-TSK Fuzzy Logic Based Screening Analysis to Diagnosis of Heart Disease. *Applied Mathematics.* 10. doi.org/10.4236/am.2019.106032.
- Dhiman, N., and Sharma, M., (2020). Fuzzy Logic Inference System for Identification and Prevention of Coronavirus (COVID-19). *International Journal of Innovative Technology and Exploring Engineering*, 9. doi.org/10.35940/ijitee.F4642.049620.
- Doob, J.L., (1953). *Stochastic Processes*. Wiley, New York.
- Dubois, D., and Prade, H., (1980). *Fuzzy Sets and Systems. Theory and Applications.* Mathematics in Science and Engineering, vol. 144. Academic Press, New York.
- Dziedzelyuk, BO., Kostiv, L. I Rabyk, V., (2013). Building Arima Time Series Models for Weather Data Predicting Using R Programming Language. *Electronics and information technologies.* 2013. Issue 3. ISSN 2224-087X.
- Eckehard Olbrich, (2008). *Data analysis and modeling - course.* Two lectures on linear models and spectral analysis. Chapter 2. [personal-homepages.mis.mpg.de/olbrich/data\\_analysis.html](http://personal-homepages.mis.mpg.de/olbrich/data_analysis.html).
- Faggini, M., (2014). Chaotic time series analysis in economics: Balance and perspectives . *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 24, Issue 4. doi.org/10.1063/1.4903797.
- Fallahgoul, H., Focardi, S. and Fabozzi, F., (2016). *Fractional calculus and fractional processes with applications to financial economics: theory and application.* Academic Press.
- Fanelli, D., Piazza, F., (2020). Analysis and forecast of COVID-19 spreading in China, Italy and France. *Chaos, Solitons and Fractals* 134, 1–12. doi.org/10.1016/j.chaos.2020.109761.
- Fisz, M., (1967). *Rachunek Prawdopodobieństwa i Statystyka Matematyczna.* PWN, Warszawa.
- Hamilton, J.D., (1994). *Time Series Analysis.* Princeton University Press, Princeton. doi:10.2307/j.ctv14jx6sm.
- Hansen, B.E., (1997). Inference in TAR Models, *Studies in Nonlinear Dynamics and Econometrics*, vol. 2, no. 1, 1–14.

Hellendoorn, H., Driankov, D., (1997). *Fuzzy Model Identification. Selected Approaches.* Springer, Berlin.

Hiteshi Tandon, Prabhat Ranjan, Tanmoy Chakraborty, Vandana Suhag, (2020). Coronavirus (COVID-19): ARIMA based time-series analysis to forecast near future. [arxiv.org/abs/2004.07859](https://arxiv.org/abs/2004.07859).

Horikawa, S., Furuhashi, T., Okuma, S. and Uchikawa, Y., (1990). Composition Methods of Fuzzy Neural Networks. 16th Annual Conference of IEEE Industrial Electronics Society. [doi.org/10.1109/IECON.1990.149317](https://doi.org/10.1109/IECON.1990.149317).

Jang, J.S.R., (1993). ANFIS: Adaptive-Network-Based Fuzzy Inference System, *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 23, No.3, pp.665-685.

Jerry, Dr., Hintze, L., (2007). *Descriptive Statistics, Means, Quality Control, and Design of Experiments. User's Guide II.* NCSS Statistical Software, chapter470.

Jiang, P., Chen, C., Liu, X., (2016). Time series prediction for evolutions of complex systems: A deep learning approach. *IEEE International Conference on Control and Robotics Engineering*. [doi.org/10.1109/ICCRE.2016.7476150](https://doi.org/10.1109/ICCRE.2016.7476150).

Jianwei, E., Ye, J. and Jin, H., (2019). A novel hybrid model on the prediction of time series and its application for the gold price analysis and forecasting. *Phys A* 527:1–14

Kacprzyk, J., (1986). *Zbiory Rozmyte w Analizie Systemowej.* Państwowe Wydawnictwo Naukowe, Warszawa.

Kacprzyk, J., and Yager, R., (2001). Linguistic Summaries of Data Using Fuzzy Logic, *International Journal of General Systems*, 30:2, 133-154. doi: 10.1080/03081070108960702

Kashpruk, N., (2016a). Comparison Identification of Statistical and Fuzzy Models of Time Series. *IEEE International Conference on Systems, Man, and Cybernetics, SMC 2016 Budapest, Hungary*.

Kashpruk, N., (2016b). The Autocorrelation Function of Errors as a Criterion for The Selection of Time Series Models. *Oficyna Wydawnicza Politechniki Opolskiej, Elektryka*, z.72 Nr.360/2016 s.29-30, Opole – Pokrzywna.

Kashpruk, N., (2016c). Identyfikacja Modeli Autoregresji W Opisie Zmian Parametrów Węgla. *Oficyna Wydawnicza Politechniki Opolskiej, Elektryka* z.74 Nr.362/2016, s. 49-54, Opole.

Kashpruk, N., (2018). Identification of Time Series Fuzzy Models With an Application to Selected Technological Processes. *Oficyna Wydawnicza Politechniki Opolskiej, Elektryka*.

Khan, U.S., (2020). Improving the Performance of Model Predictive Control by Detecting Model Plant Mismatches. *International Bhurban Conference on Applied Sciences and Technology*. [doi.org/10.1109/IBCAST47879.2020.9044570](https://doi.org/10.1109/IBCAST47879.2020.9044570).

Kochanek, K., (2016), Zastosowanie Metody Uogólnionych Momentów do Estymacji Kwantyli Przepływów Maksymalnych Wybranych Rozkładów O Grubych Ogonach.

Koen Vos, (2013). A Fast Implementation of Burg's Method. [opus-codec.org/docs/vos\\_fastburg.pdf](https://opus-codec.org/docs/vos_fastburg.pdf).

Lazzaro, V., Silvia, A. and Massimo, C., (2020). Application of the ARIMA model on the COVID- 2019 epidemic dataset. Data in brief, 29: 105340. 5. World Health Organization (WHO), Coronavirus.

Łęski, J., (2008). Systemy Neuronowo-Rozmyte. Wydawnictwa Naukowo-Techniczne, Warszawa.

Li, J., Yang, L., Qu, Y., et al., (2018). An extended Takagi-Sugeno-Kang inference system (TSK+) with fuzzy interpolation and its rule base generation. Soft Computing 22. [doi.org/10.1007/s00500-017-2925-8](https://doi.org/10.1007/s00500-017-2925-8).

Li, Q., Feng, W., Quan, Y.H., (2020). Trend and forecasting of the COVID-19 outbreak in China. J. Inf. Secur. 80, 469–496. [doi.org/10.1016/j.jinf.2020.02.014](https://doi.org/10.1016/j.jinf.2020.02.014).

Liu, K., Chen, Y., and Zhang, X. (2017). An Evaluation of ARFIMA (Autoregressive Fractional Integral Moving Average) Programs. Axioms, 6, 16.

Lukhyswara, P., Putranto, L.M., Ariananda, D.D., (2019). Solar Irradiation Forecasting Uses Time Series Analysis. 11th International Conference on Information Technology and Electrical Engineering. [doi.org/10.1109/ICITEED.2019.8929990](https://doi.org/10.1109/ICITEED.2019.8929990).

Mamdani, E.H., Assilian, S., (1975). An Experiment in Linguistic Synthesis with a Fuzzy Logic Controller. International Journal of Man-Machine Studies, 7 (1), 1-13.

Mańczak, K., (1970). Metody Identyfikacji Wielowymiarowych Obiektów Sterowania. Wydawnictwa Naukowo-Techniczne, Warszawa.

Mańczak, K., Nahorski, Z., (1983). Komputerowa Identyfikacja Obiektów Dynamicznych. Państwowe Wydawnictwo Naukowe, Warszawa.

Mandelbrot, B., (1982). The Fractal Geometry of Nature. W.H. Freedman and Company, New York, Berlin, Heidelberg.

Mandelbrot, B., (1999). Multifraktale rządzą na Wall Street. „Świat Nauki”, nr 04, s. 64-67.

Mantegna, R.N., and Stanley, H.E., (2000). An Introduction to Econophysics. Correlations and Complexity in Finance. Cambridge University Press.

Marszałek, A., Burczyński, T., (2013). Financial fuzzy time series models based on ordered fuzzy numbers. In: W. Pedrycz, S.-M. Chen (Eds.), Time Series Analysis, Model & Applications, ISRL 47, pp. 77–95, Springer-Verlag, Berlin Heidelberg.

- Mills, T.C., (2019). Chapter 3 - ARMA Models for Stationary Time Series, Editor(s): Terence C. Mills, Applied Time Series Analysis, Academic Press. doi.org/10.1016/B978-0-12-813117-6.00003-X.
- Minakhi Rout, Babita Majhi, Ritanjali Majhi, Ganapati Panda, (2014). Forecasting of currency exchange rates using an adaptive ARMA model with differential evolution based training, Journal of King Saud University - Computer and Information Sciences, Volume 26. doi.org/10.1016/j.jksuci.2013.01.002.
- Mularczyk, A., (2008). Fraktalna Analiza Ryzyka w Działalności Przedsiębiorstw. Wydawnictwo Politechniki Śląskiej. Monografia.
- Niederliński, A., (1979). Systemy i Sterowanie. Skrypty Uczelniane Nr. 901, Gliwice, 1979.
- Niederliński, A., Kasprzyk, J., Figwer, J., (1997). MULTI-EDIP – Analizator Wielowymiarowych Sygnałów i Obiektów. Wydawnictwo Politechniki Śląskiej, Gliwice.
- Nowak-Brzezińska Agnieszka, (2012). Drzewa klasyfikacyjne. Konspekt do zajęć: Statystyczne metody analizy danych. [http://zsi.tech.us.edu.pl/~nowak/smad/smad\\_lab10a.pdf](http://zsi.tech.us.edu.pl/~nowak/smad/smad_lab10a.pdf)
- Papoulis, A., (1972). Prawdopodobieństwo, Zmienne Losowe i Procesy Stochastyczne. Wydawnictwa Naukowo-Techniczne, Warszawa.
- Pedregal, D.J., (2019). Time series analysis and forecasting with ECOTOOL. October 31. doi.org/10.1371/journal.pone.0221238.
- Pedrycz, W., and Chen, S.M., editors (2013). Time Series Analyses, A Modeling and Application. A computational Intelligence Perspective. Intelligent Systems Reference Library. Vol. 47. Springer.
- Pedrycz, W., and Gomide, F., (1998). An Introduction to Fuzzy Sets: Analysis and Design. Cambridge, MA : MIT Press.
- Piegat, A., (2009). Modelowanie i Sterowanie Rozmyte. Akademicka Oficyna Wydawnicza EXIT, Warszawa.
- Probierz Barbara, (2017). Automatyczna kategoryzacja wiadomości elektronicznych z zastosowaniem sieci społecznych oraz algorytmów mrowiskowych. Rozprawa doktorska. Uniwersytet Śląski, Wydział Informatyki i Nauki o Materiałach, Instytut Informatyki.
- Rashmi Bhardwaj, Aashima Bangia, (2020). Data driven estimation of novel COVID-19 transmission risks through hybrid soft-computing techniques, Chaos, Solitons & Fractals, Volume 140. doi.org/10.1016/j.chaos.2020.110152.
- Roda, W.C., Varughese, M.B., Han, D., Li, M.Y., (2020). Why is it difficult to accurately predict the COVID-19 epidemic? Infect. Dis. Model. 5, 271-281.



- Roosa, K., Lee, Y., Luo, R., Kirpich, A., Rothenberg, R., Hyman, J.M., Yan, P., Chowell, G. (2020). Real-time forecasts of the COVID-19 epidemic in China from February 5th to February 24<sup>th</sup>. *Infect. Dis. Model.* 5, 256–263. doi.org/10.1016/j.idm.2020.02.002.
- Rutkowski, L., (2005). *Metody i Techniki Sztucznej Inteligencji*. Wydawnictwo Naukowe PWN, Warszawa.
- Rutkowski, L., (2020). *Metody i techniki sztucznej inteligencji. Algorytm HCM*, Warszawa.
- Sahu, B.K., (2005). *Statistical Models in Earth Sciences*, BS Publications, Hyderabad.
- Sandya, H.B., et al (2013). Fuzzy Rule Based Feature Extraction and Classification of Time Series Signal. *International Journal of Soft Computing and Engineering (IJSCE)*, ISSN: 2231-2307, Volume-3, Issue-2, May.
- Schwarz, G., (1978): Estimating the dimension of a model. *Annals of Statistics* 6: 461–464.
- Silveira-Gontijo, T., Azevedo, Costa M., (2020). Forecasting Hierarchical Time Series in Power Generation. *Energies*, vol. 2020, 3722. doi.org/10.3390/en13143722.
- Sobczyk, M., (1996). *Statystyka*. Wydawnictwo Naukowe PWN.
- Soderstrom, T., Stoica P., (1989). *System Identification*. Prince Hall.
- Soderstrom, T., Stoica, P., (1997). *Identyfikacja systemów*. Wydawnictwo Naukowe PWN, Warszawa.
- Soudeep Deb, Manidipa Majumdar, (2020). A time series method to analyze incidence pattern and estimate reproduction number of COVID-19. [arxiv.org/abs/2003.10655](https://arxiv.org/abs/2003.10655).
- Stanisławski, R., Rydel, M., and Latawiec, K.J., (2017). Modeling of Discrete-Time Fractional Order State Space Systems Using the Balanced Truncation Method. *Journal of the Franklin Institute*, vol.354, no. 7, pp. 3008-3020. doi.org/10.1016/j.jfranklin.2017.02.003
- Stellwagen, E. and Tashman, L., (2013). ARIMA. *The Models of Box and Jenkins*. *Foresight: Int. J. Appl. Forecast.* 28-33.
- Sugeno, M. and Kang, G.T., (1988). Structure Identification of Fuzzy Model. *Fuzzy Sets and Systems*, vol. 28, pp. 15-33.
- Swiesznikow, A.A., (1965). *Podstawowe Metody Funkcji Losowych*. Wydawnictwo Naukowe PWN, Warszawa.
- Taieb, A., Soltani, M., Chaari, A., (2018). A fuzzy C-regression model algorithm using a new PSO algorithm. *Int. J. Adapt. Control Signal Process.* 32(1), 115–133.
- Takagi, T. and Sugeno, M., (1985). Fuzzy Identification of Systems and Its Applications to Modeling and Control,” *IEEE Transactions on Systems, Man and Cybernetics*, 15, pp.116-132.

Talar, J., (2007). Data mining methods – application in metallurgy — Metody eksploracji danych – zastosowanie w metalurgii / Archives of Metallurgy and Materials / Polish Academy of Sciences. Committee of Metallurgy. Institute of Metallurgy and Materials Science ; ISSN 1733-3490.

Taylor, Sean J., i Letham, Benjamin, (2018) Forecasting at Scale, *The American Statistician*, 72:1, 37-45. doi.org/10.1080/00031305.

Terasvirta, T. and Kock, A.B. (2010). Forecasting with Nonlinear Time Series Models, *SSRN Electronic Journal*. doi.org/10.2139/ssrn.1531092.

Terasvirta, T., (1994). Specification, Estimation and Evaluation of Smooth Transition Autoregressive models. *J. Am. Stat. Assoc.* 89, 208–218.

Teshome, H., Abebe, S., (2020). Forecasting the Number of Coronavirus (COVID-19) Cases in Ethiopia Using Exponential Smoothing Times Series Model. doi.org/10.1101/2020.06.29.20142489.

The MathWorks Inc. MATLAB—The Language of Technical Computing, Version R2018a. Natick, Massachusetts. URL [www.mathworks.com/products/matlab/](http://www.mathworks.com/products/matlab/); 2018.

Tong, R.M., (1978). Synthesis of Fuzzy Models for Industrial Processes – Some Recent Results. *International Journal of General Systems*, vol. 4, pp. 143-163.

Tong, R.M., (1979). The Construction and Evaluation of Fuzzy Models, in: *Advances in Fuzzy Set Theory and Applications*, Gupta M.M., Ragade R.K. and Yager R.R. (Eds.), North-Holland, Amsterdam, pp. 559-576, 1979.

Tsai, S.H., Chen, Y.W., (2018). A novel identification method for Takagi–Sugeno fuzzy model. *Fuzzy Sets and Systems*, Volume 338. doi.org/10.1016/j.fss.2017.10.012.

Tsay, R., (2002). *Analysis of Financial Time Series*. John Wiley & Sons, Inc., New York.

Tsung, F., Tsui, K.L., (2003). A mean-shift pattern study on integration of SPC and APC for process monitoring, *IIE Transactions*, vol. 35, 231-242. doi.org/10.1080/07408170304365.

Vinay Kumar Reddy Chimmula, Lei Zhang, (2020). Time series forecasting of COVID19 transmission in Canada using LSTM networks. *Chaos, Solitons and Fractals*, 135; 109864 2.

Walaszek-Babiszewska, A., (1992). *Modele Stochastyczne Opróbowania Węglu*. Górnictwo. Zeszyty Naukowe Politechniki Śląskiej, z.203, Gliwice.

Walaszek-Babiszewska, A., (2011). *Fuzzy Modeling in Stochastic Environment. Theory, Knowledge Bases, Examples*. LAP Lambert Academic Publishing, Saarbrücken.

Walaszek-Babiszewska, A., (2015). *Wykład Monograficzny dla Studiów Doktoranckich w Politechnice Opolskiej*, niepublikowany.

Walaszek-Babiszewska, A., Bryniarska, A., (2018). *Podstawy Teorii Systemów Rozmytych z Zadaniem*. Oficyna Wydawnicza Politechniki Opolskiej, Opole.

Walaszek-Babiszewska, A., Rudnik, K., (2013). Stochastic-fuzzy knowledge-based approach to temporal data modeling, w: *Time Series Analysis, Modeling and Applications; A Computational Intelligence Perspective*, W. Pedrycz i Shyi-Ming Chen (red.). *Intelligent Systems Reference Library*, vol.47, s. 97-118. Springer.

Walaszek-Babiszewska, A., Rydel, M. and Kashpruk, N. (2018). On the Equivalence Between AR Family Time Series Models and Fuzzy Models in Signal Processing. 2018 IEEE Second International Conference on Data Stream Mining & Processing (DSMP), Lviv, pp. 331-335.

Wang, L., Li, J., Guo, S., et al (2020). *J. ur of. Sci. Total Environ.*, 138394. doi.org/10.1016/j.scitotenv.2020.138394.

Welfe, A., (2019). *Ekonometria: Metody i Ich Zastosowania*. Państwowe Wydawnictwo Ekonomiczne.

Wiktorowicz, K., Krzeszowski, T., (2020). Training High-Order Takagi-Sugeno Fuzzy Systems Using Batch Least Squares and Particle Swarm Optimization. *Int. J. Fuzzy Syst.* 22, 22–34. doi.org/10.1007/s40815-019-00747-2.

Wu, J.T., Leung, K., Leung, G.M., (2020). Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: a modelling study. *Lancet* 395, 689–697. doi.org/10.1016/S0140-6736(20) 30260-9.

Xingde, Duan, Xiaolei Zhang, (2020). ARIMA modelling and forecasting of irregularly patterned COVID-19 outbreaks using Japanese and South Korean data. *Data in Brief*, Volume 31, Article 105779.

Yager, R.R., Filev, D.P., (1995). *Podstawy Modelowania i Sterowania Rozmytego*. Wydawnictwa Naukowo-Techniczne, John Wiley and Sons, Inc., Warszawa.

Yang, C., Wu Q., (2010). On stability analysis via Lyapunov exponents calculated from a time series using nonlinear mapping-a case study. *Nonlinear Dyn* 59, 239. doi.org/10.1007/s11071-009-9535-7

Yichi, L., Bowen, W., et al (2020). Mathematical Modeling and Epidemic Prediction of COVID-19 and Its Significance to Epidemic Prevention and Control Measures. *Annals of Infectious Disease and Epidemiology*.

Yonar, H., Yonar, A., Tekindal, M.A. and Tekindal, M., (2020). Modeling and Forecasting for the number of cases of the COVID-19 pandemic with the Curve Estimation Models, the BoxJenkins and Exponential Smoothing Methods. *EJMO*, 4(2):160–165.

Yong, W., Lingyun, P., and Jia, W. (2020). Statistical analysis and ARMA modeling for the big data of marathon score, *Science & Sports*. doi.org/10.1016/j.scispo.2020.01.009.

Zadeh, L.A. (1965). Fuzzy Sets. *Information and Control*, 8, pp. 338-353.

Zadeh, L.A., (1968). Probability measures of fuzzy events. *Journal of Mathematical Analysis and Applications*, Vol. 23, No 2, pp. 421-427.

Zadeh, L.A., (1973). Outline of a new approach to the analysis of complex systems and decision processes. *IEEE Transactions on Systems, Man and Cybernetics*, vol. SMC-3, pp. 28-44.

Zadeh, L.A., (1975). The Concept of a Linguistic Variable and its Application to Approximate Reasoning. *Information Sciences, Part I*: 8, pp. 199-249.

Zadeh, L.A., (1979). A theory of approximate reasoning, in: Hayes J.E., Michie D. and Mikulich L.I., (Eds.), *Machine Intelligence Vol. 9*, New York, pp.149-194.

Zadeh, L.A., (1996). Fuzzy Logic, Neural Networks, and Soft Computing. *Advances in Fuzzy Systems - Applications and Theory Fuzzy Sets, Fuzzy Logic, and Fuzzy Systems*, pp. 775-782. doi.org/10.1142/9789814261302\_0040.

Zeynep Ceylan, (2020). Estimation of COVID-19 prevalence in Italy, Spain, and France, *Science of The Total Environment*, Volume 729. doi.org/10.1016/j.scitotenv.2020.138817.

Zhang, Z., Moore, J.C. (2015). Chapter 8 - Autoregressive Moving Average Models. *Mathematical and Physical Fundamentals of Climate Change*, Elsevier. doi.org/10.1016/B978-0-12-800066-3.00008-5.

[en.wikipedia.org/wiki/Autoregressive\\_model#Yule–Walker\\_equations](https://en.wikipedia.org/wiki/Autoregressive_model#Yule–Walker_equations)

[facebook.github.io/prophet/](https://facebook.github.io/prophet/)

[metallurgprom.org/uk/novini/grnichorudna-galuz/3732-v-svjazi-s-sistemnym-krizisom-v-jenergeticheskoy-otrasli-ukrainy-dtjek-ostanavlivaet-shahty.html](https://metallurgprom.org/uk/novini/grnichorudna-galuz/3732-v-svjazi-s-sistemnym-krizisom-v-jenergeticheskoy-otrasli-ukrainy-dtjek-ostanavlivaet-shahty.html)

[money.pl/pieniadze/nbparch/srednie/?symbol=USD](https://money.pl/pieniadze/nbparch/srednie/?symbol=USD)

[shop.lumel.com.pl/pl/p/Miernik-licznik-energii-z-certyfikatem-MID-NMID30-2/14755?utm\\_source=shoper&utm\\_medium=shoper-cpc&utm\\_campaign=shoper-kampanie-google&shop\\_campaign=5705022066&gclid=CjwKCAiAqJn9BRB0EiwAJ1SztcR05twOdjBBeg0Pexl-W-hRqmDkpDPDy105XGc23Hre\\_BNaDm5ZjRoC57oQAvD\\_BwE](https://shop.lumel.com.pl/pl/p/Miernik-licznik-energii-z-certyfikatem-MID-NMID30-2/14755?utm_source=shoper&utm_medium=shoper-cpc&utm_campaign=shoper-kampanie-google&shop_campaign=5705022066&gclid=CjwKCAiAqJn9BRB0EiwAJ1SztcR05twOdjBBeg0Pexl-W-hRqmDkpDPDy105XGc23Hre_BNaDm5ZjRoC57oQAvD_BwE) (08.11.2020 g 10.10)