UNITED KINGDOM · CHINA · MALAYSIA

University of Nottingham

Doctoral Thesis

Interval Type-2 Atanassov-Intuitionistic Fuzzy Logic for Uncertainty Modelling

Imo Jeremiah Eyoh, BSc., MSc.

A thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy of the University of Nottingham.

School of Computer Science

April 27, 2018

Abstract

This thesis investigates a new paradigm for uncertainty modelling by employing a new class of type-2 fuzzy logic system that utilises fuzzy sets with membership and non-membership functions that are intervals. Fuzzy logic systems, employing type-1 fuzzy sets, that mark a shift from computing with numbers towards computing with words have made remarkable impacts in the field of artificial intelligence. Fuzzy logic systems of type-2, a generalisation of type-1 fuzzy logic systems that utilise type-2 fuzzy sets, have created tremendous advances in uncertainty modelling. The key feature of the type-2 fuzzy logic systems, with particular reference to interval type-2 fuzzy logic systems, is that the membership functions of interval type-2 fuzzy sets are themselves fuzzy. These give interval type-2 fuzzy logic systems an advantage over their type-1 counterparts which have precise membership functions. Whilst the interval type-2 fuzzy logic systems are effective in modelling uncertainty, they are not able to adequately handle an indeterminate/neutral characteristic of a set, because interval type-2 fuzzy sets are only specified by membership functions with an implicit assertion that the non-membership functions are complements of the membership functions (lower or upper). In a real life scenario, it is not necessarily the case that the non-membership function of a set is complementary to the membership function. There may be some degree of hesitation arising from ignorance or a complete lack of interest concerning a particular phenomenon. Atanassov intuitionistic fuzzy set, another generalisation of the classical fuzzy set, captures this thought process by simultaneously defining a fuzzy set with membership and non-membership functions such that the sum of both membership and non-membership functions is less than or equal to 1.

In this thesis, the advantages of both worlds (interval type-2 fuzzy set and Atanassov intuitionistic fuzzy set) are explored and a new and enhanced class of interval type-2 fuzzy set namely, interval type-2 Atanassov intuitionistic fuzzy set, that enables hesitation, is introduced. The corresponding fuzzy logic system namely, interval type-2 Atanassov intuitionistic fuzzy logic system is rigorously and systematically formulated. In order to assess the viability and efficacy of the developed framework, the possibilities of the optimisation of the parameters of this class of fuzzy systems are rigorously examined.

First, the parameters of the developed model are optimised using one of the most popular fuzzy logic optimisation algorithms such as gradient descent (first-order derivative) algorithm and evaluated on publicly available benchmark datasets from diverse domains and characteristics. It is shown that the new interval type-2 Atanassov intuitionistic fuzzy logic system is able to handle uncertainty well through the minimisation of the error of the system compared with other approaches on the same problem instances and performance criteria.

Secondly, the parameters of the proposed framework are optimised using a decoupled extended Kalman filter (second-order derivative) algorithm in order to address the shortcomings of the first-order gradient descent method. It is shown statistically that the performance of this new framework with fuzzy membership and non-membership functions is significantly better than the classical interval type-2 fuzzy logic systems which have only the fuzzy membership functions, and its type-1 counterpart which are specified by single membership and non-membership functions.

The model is also assessed using a hybrid learning of decoupled extended Kalman filter and gradient descent methods. The proposed framework with hybrid learning algorithm is evaluated by comparing it with existing approaches reported in the literature on the same problem instances and performance metrics. The simulation results have demonstrated the potential benefits of using the proposed framework in uncertainty modelling. In the overall, the fusion of these two concepts (interval type-2 fuzzy logic system and Atanassov intuitionistic fuzzy logic system) provides a synergistic capability in dealing with imprecise and vague information.

Declaration

The work presented in this thesis is based on a research carried out in the Automated Scheduling, Optimisation and Planning (ASAP) research group, in the School of Computer Science, University of Nottingham. No part of this thesis has been submitted elsewhere for any other examination, degree or qualification and it is solely the work of the author. All verbatim extracts have been distinguished by quotation marks, and all sources of information have been specifically acknowledged.

Signed:

Date: April 27, 2018

Acknowledgements

"For he that is mighty hath done to me great things; and holy is his name" Luke 1:49.

It has been a great pleasure working with some of the best intellectuals in the School of Computer Science, University of Nottingham.

I would like to thank my main supervisor, Prof. Robert John, who had been an understanding, supportive and patient supervisor, encouraging me in every way possible. He was not only concerned about my academic career but also about my total wellbeing. He gave me the confidence to succeed in the face of uncertainty. Thank you Prof. Bob for your open door policy that provided me the opportunity to meet with you irrespective of your tight schedules. In fact, I could not have had a better supervisor. I count it an honour and a great privilege to be listed as one of your PhD students. Thank you Bob. I would like to say a special thank you to Dr. Geert, my co-supervisor, who, on our first meeting gave me an advice that saw me through my three years of research. Geert, together with Bob, painstakingly read my thesis, dotting all the i's and crossing all the t's. What else could I have asked for? Thank you Geert. I would like to thank Deborah Pitchfork and Christine Fletcher for all the administrative deals. You did great jobs.

I am grateful to my PhD viva voce examiners, Prof. Jonathan Garibaldi (Editor-in-Chief, IEEE Transactions on Fuzzy Systems) and Dr. Simon Coupland for their valuable comments and suggestions that helped to position this thesis a step further.

I would like to thank the formidable ASAP C77 team - Wenwen (Mina), Raras, Haneen, Wasakorn, Peng, Binhui and Jonata (Captain). You all made my PhD journey an enjoyable and unforgettable experience, and to all my special friends, Susan, Faiza, Elissa and Osman, you were simply great.

A big thank you to RCCG, Rehoboth House family, Nottingham. Thanks to my friends in the Department of Computer Science, University of Uyo: Edward, Umoeka, Uduak Umoh and all the wonderful people of the department.

I am grateful to my wonderful sisters - Mrs Uduak Aimufua, Mercy Ukpong, Idongesit

Ben, Owoidighe-Abasi Usanga and Emem Abasi-ifreke, in-laws and other relatives, for all the love, prayers and encouragements. Thank you Veronica, my dear niece, for holding sway while I was away. I am indebted to Emefen and the family in Plymouth, who were always happy to give me and my children special Christmas treats. Your house was a home away from home.

Finally to my dear husband, everything I am today and who I will ever be, I owe them to you. I would not be where or who I am today without you. Thank you Precious for believing in me and pushing me to attain this great height. Thank you Kingsley, Henrietta, Priestley and Angel-Virtue for bringing smiles to Mummy's face and for creating a 'fuzzy' working space for Mummy to conduct her research. I have successfully completed the research, you can now have all the space with love.

This research was supported by the Government of Nigeria under the auspices of the Tertiary Education Trust Fund (TETFund).

To my best friend: Jerry

Contents

Abstract **ii**

Declaration **iv**

Acknowledgements **v**

1 Introduction **1**

1.1 Problem Statement 1

1.2 Background and Motivation 5

1.3 Contributions 8

1.4 Academic Publications 9

1.5 Thesis Outline 11

2 Related Work **13**

2.1 Introduction 13

2.2 Fuzzy Set Theory 14

2.2.1 Type-1 Fuzzy Set: Definition 15

2.2.2 Type-1 Fuzzy Logic Systems 16

2.2.3 Generalisation of a Fuzzy Set 17

2.3 Interval Type-2 Fuzzy Set 18

2.3.1 Comparison Between Interval Type-2 Fuzzy Set and Interval Valued Fuzzy Set 19

2.4 Interval Type-2 Fuzzy Logic Systems 20

2.4.1 Fuzzification Process 20

2.4.2 Rules 21

2.4.3 Fuzzy Inference Engine 21

2.4.4 Type Reduction 23

2.4.5 Defuzzification 23

2.5 Uncertainty Modelling 24

2.6 IT2FLSs Design Methodology 26

2.6.1 The Gradient Descent Methods 27

2.6.2 The Kalman Filter-based Methods 29

2.7 Application of IT2FLSs to Uncertainty Modelling 31

2.7.1 Application to Classification and Prediction Problems 31

2.7.2 Application to Pattern Recognition Problems 33

2.7.3 Application to Clustering Problems 34

2.7.4 Application to Control Problems 35

2.8 Drawbacks of IT2FLSs 36

2.9 Atanassov Intuitionistic Fuzzy Set 36

2.9.1 Type-1 AIFS: Definition 37

2.9.2 Atanassov Intuitionistic Fuzzy Logic Systems 38

2.10 Practical Applications of AIFSs 40

2.11 Drawbacks of AIFSs 42

2.12 Studies Involving Combination of AIFSs and IT2FSs 42

2.13 Summary 44

3 Model Formulation 46

3.1 Introduction 46

3.2 Generalised Type-2 A-Intuitionistic Fuzzy Set 46

3.3 Interval Type-2 Atanassov Intuitionistic Fuzzy Set 47

3.4 A Comparison Between Interval Valued Atanassov Intuitionistic Fuzzy Set and Interval Type-2 Atanassov Intuitionistic Fuzzy Set 50

3.5 TSK-based Interval Type-2 Atanassov-Intuitionistic Fuzzy Logic System Framework 51

3.5.1 Fuzzification 53

3.5.2 Rules 56

3.5.3 Inference 56

3.5.4 Output Processing 57

3.6 Summary 59

4 Gradient Descent Learning of IT2AIFLS with Application to Time Series and Regression Problems 60

4.1 Introduction 60
4.2 IT2AIFLS Rule Structure 61
4.3 Parameter Update Rule 61
4.3.1 Consequent Parameter Update 62
4.3.2 Antecedent Parameter Update 63
4.4 Experiments and Results 65
4.4.1 A Comparison of IT2AIFLS, FIS, IFIS and IT2FLS on Regression Problems 67
4.4.2 Friedman#2 70
4.4.3 Electrical Engineering Distribution Problems 72
4.4.4 Time Series Prediction 75
4.4.5 Complex High Dimensional Regression Problems 87
4.5 Summary 92

5 Extended Kalman Filter-based Learning of IT2AIFLS for System Identification and Time Series Predictions 93

5.1 Introduction 93
5.1.1 Rules 94
5.1.2 Inference 94
5.2 Parameter Updates 95
5.2.1 Extended Kalman Filter Parameter Update Rule 95
5.2.2 Decoupled Extended Kalman Filter 97
5.3 Antecedent Update Rule 97
5.4 Consequent Parameter Update 99
5.5 Experiments and Results 100
5.5.1 System Identification 100
5.5.2 NSW Electricity Load Forecast 101
5.5.3 Gas Compression System Time Series Prediction 105
5.6 Statistical Evaluation 107
5.7 Summary 110

6 **Hybrid Learning of IT2AIFLS as applied to Identification and Prediction Problems** **112**
6.1 Introduction . . . 112
6.2 IT2AIFLS Rule Structure . . . 113
6.3 Parameter Updates . . . 113
6.3.1 Consequent Parameter Updates . . . 114
6.3.2 Antecedent Parameter Updates . . . 114
6.4 Experimental Analysis and Evaluation . . . 115
6.4.1 Application to Artificially Generated Mackey-Glass Time Series . . . 115
6.4.2 System Identification Problem #1 . . . 119
6.4.3 System Identification Problem #2 . . . 123
6.4.4 System Identification Problem #3 . . . 124
6.4.5 Application to Real World Electricity Load Forecasting . . . 125
6.4.6 Gas Furnace Time Series . . . 127
6.4.7 Santa Fe A Time Series . . . 129
6.5 Summary . . . 131

7 **Conclusions and Discussion** **133**
7.1 Discussion . . . 133
7.2 Summary of Contribution . . . 135
7.3 Limitations of the Proposed Framework . . . 136
7.4 Future Research Directions . . . 136
7.4.1 Non-derivative Methods for Tuning the Parameters of IT2AIFLS . . . 137
7.4.2 Structure Learning of the Proposed Framework . . . 137
7.4.3 Use of Adaptive Learning Rate and IF-indices . . . 137
7.4.4 Application of other Membership and Non-membership Functions, $t - norms$ and Fuzzification Procedure . . . 138
7.4.5 Stability Analysis of IT2AIFLS . . . 139
7.4.6 Integration of AIFS with GT2FLS . . . 139
7.4.7 Formulating the Model based on Mamdani Fuzzy Inference . . . 139
7.4.8 Analysis of Data Stream . . . 140
7.5 Summary . . . 141

List of Figures

2.1 A Gaussian type-1 membership function 16

2.2 A type-1 FLS [1] 16

2.3 A Gaussian interval type-2 membership function 20

2.4 A T2FLS structure [1] 21

2.5 A-intuitionistic Gaussian membership and non-membership functions - AIFS 37

2.6 Structure of AIFLS 39

3.1 An IT2 A-intuitionistic Gaussian membership and non-membership functions - IT2AIFS [2] 48

3.2 Type-2 A-Intuitionistic Fuzzy Logic System [3] 52

3.3 Gaussian MF with uncertain standard deviation 54

3.4 IT2AIFS 55

3.5 An IT2AIFLS Structure - adapted from [4] 57

4.1 Histogram of additive white Gaussian noise for noisy Friedman problem . . 68

4.2 Actual and predicted outputs of Friedman with Gaussian white noise . . . 72

4.3 Adaptation of β values for electrical maintenance cost 76

4.4 Correlation analysis between the actual and predicted outputs for electrical maintenance cost 76

4.5 Actual and predicted values of sunspot time series using IT2AIFLS 79

4.6 Adaptation of β values for sunspot time series 80

4.7 Tree ring dataset 82

4.8 Actual and predicted outputs of tree ring time series 82

4.9 Original Canadian lynx time series (1821-1934) 85

4.10 Transformed Canadian lynx time series (log10) 85

4.11 Plot of Santa Fe A time series 86

4.12 Plot of abalone data inputs set 88

4.13 Plot of actual outputs of abalone dataset 88
4.14 Actual and predicted outputs of abalone dataset 90
4.15 Feature ranking of house sales data [5] 90

5.1 Actual and predicted output using IT2AIFLS for identification problem . . 101
5.2 Price prediction in summer, autumn, winter and spring using IT2AIFLS-DEKF and IT2AIFLS-GD respectively 104
5.3 GCS membership and non-membership functions before training with IT2AIFLS 106
5.4 GCS membership and non-membership functions after training with IT2AIFLS 106
5.5 GCS membership function before training with IT2FLS 107
5.6 GCS membership function after training with IT2FLS 107
5.7 Box-and-whisker plot showing the performance of IT2AIFLS, IT2FLS and AIFLS. 107

6.1 Hybrid learning procedure of IT2AIFLS using DEKF and GD 114
6.2 Actual and predicted output of Mackey-Glass time series 117
6.3 The adaptation of the parameter β for Mackey-Glass prediction problem . . 119
6.4 Time varying parameters for second-order system identification problem #1 120
6.5 Actual and predicted output using hybrid IT2AIFLS for second-order system identification problem #1 122
6.6 Actual and predicted values for system identification#2 using IT2AIFLS . . 124
6.7 Plot of Poland electricity load training data 126
6.8 Actual and predicted values of Poland electricity load with IT2AIFLS on test dataset 127
6.9 Actual and predicted output of gas furnace time series using IT2AIFLS trained with DEKF+GD 128
6.10 Actual and predicted values of Santa Fe A time series with IT2AIFLS . . . 131

List of Tables

4.1 Dataset characteristics 68

4.2 Excerpt from energy dataset 69

4.3 Excerpt from stock dataset 69

4.4 Excerpt from autoMPG6 dataset 69

4.5 RMSE and std of IT2AIFLS vs FIS/IFIS/IT2FLS on regression problems . 70

4.6 Performance comparison of IT2AIFLS with existing models on Friedman#2 72

4.7 Excerpt from low voltage line lengths 73

4.8 Performance comparison of IT2AIFLS with existing models on low voltage line length estimation problem 74

4.9 Performance comparison of IT2AIFLS with classical IT2FLS on voltage length estimation problem 74

4.10 Excerpt from electrical maintenance cost dataset 75

4.11 Performance comparison of IT2AIFLS with other models on electrical maintenance cost estimation problem 77

4.12 Performance comparison of IT2AIFLS with other appraches on Mackey-Glass time series forecasting 78

4.13 Comparison of IT2FLS-TSK and IT2AIFLS on Mackey-Glass time series . 78

4.14 Excerpt from sunspot time series data 79

4.15 Performance comparison of IT2AIFLS with other models on sunspot time series 80

4.16 Performance comparison of IT2AIFLS with classical IT2FLS using sunspot time series dataset 81

4.17 Excerpt from tree ring time series data 81

4.18 Performance comparison of IT2AIFLS with other models using tree ring time series forecasting 83

4.19 Excerpt from Canadian lynx time series data 83

4.20 Performance comparison of IT2AIFLS with non-fuzzy models on Canadian lynx time series . . . 84
4.21 Excerpt from Santa Fe A time series data . . . 85
4.22 Performance comparison of IT2AIFLS with other models on Santa Fe A time series dataset . . . 86
4.23 Excerpt from abalone dataset . . . 87
4.24 Comparison of IT2AIFLS with other models using abalone dataset . . . 89
4.25 Comparison of IT2AIFLS with classical IT2FLS using large and high dimensional house sales data . . . 91
4.26 Summary of results . . . 91

5.1 Comparison of IT2AIFLS vs AIFLS and IT2FLS on second-order identification problem . . . 101
5.2 NSW 2008 electricity price dataset partitions . . . 102
5.3 Performance of different models and algorithms during Summer season . . . 103
5.4 Performance of different models and algorithms during Autumn season . . . 103
5.5 Performance of different models and algorithms during Winter season . . . 103
5.6 Performance of different models and algorithms during Spring season . . . 103
5.7 Performance comparison of IT2FLS, AIFLS and IT2AIFLS using GCS dataset 106
5.8 Wilcoxon's test: IT2AIFLS and IT2FLS using test RMSE . . . 108
5.9 Wilcoxon's test: IT2AIFLS and AIFLS using test RMSE . . . 109
5.10 Wilcoxon's test: IT2FLS and AIFLS using test RMSE . . . 110
5.11 Summary of results . . . 110

6.1 Performance comparison of IT2AIFLS on Mackey-Glass time series forecasting with existing models . . . 118
6.2 Performance comparison of IT2AIFLS with other models on second-order system identification problem #1 . . . 121
6.3 Comparison of runtime of IT2AIFLS with other approaches on second-order identification problem #1 . . . 122
6.4 Performance comparison of hybrid IT2AIFLS with other models on non-linear system identification#2 . . . 123
6.5 A Comparison of IT2AIFLS, AIFLS and IT2FLS on second-order identification problem #3 . . . 125

6.6 Excerpt from Poland electricity data 126
6.7 Comparison of IT2AIFLS versus AIFLS and IT2FLS on Poland electricity load forecast 127
6.8 Excerpt from gas furnace data 128
6.9 Performance comparison of hybrid-IT2AIFLS on gas furnace time series . . 129
6.10 Performance comparison of hybrid-IT2AIFLS with other models on Santa Fe A time series dataset 130
6.11 Summary of results 131

List of Symbols/Abbreviations

A	A type-1 fuzzy set
$\tilde{A}$	A type-2 fuzzy set
A^*	A type-1 Atanassov intuitionistic fuzzy set
$\tilde{A}^*$	A type-2 Atanassov intuitionistic fuzzy set
β	Membership and non-membership functions weighting factor
E	Cost function of a fuzzy logic system
f	Firing strength of a type-2 fuzzy set
f^{μ}	Firing strength of membership function of a type-2 intuitionistic fuzzy set
f^{ν}	Firing strength of non-membership function of a type-2 intuitionistic fuzzy set
H	Derivative matrix
I	Identity matrix
K	Kalman gain
P	State covariance matrix
Q	Process noise covariance
R	Measurement noise covariance
R_k	The kth rule of a fuzzy logic system
y	Output of a type-2 fuzzy logic system
y^a	Actual output of a system
y_k^{μ}	Membership function output of the kth rule
y_k^{ν}	Non-membership function output of the kth rule

γ	Learning rate of a gradient descent back-propagation algorithm
θ	Generic parameter of a type-2 fuzzy logic system
$\mu_A(x)$	Type-1 membership degree
$\mu_{\tilde{A}}(x)$	Type-2 membership degree
$\mu_{\tilde{A}^*}(x)$	Type-2 intuitionistic membership degree
$\nu_{\tilde{A}^*}(x)$	Type-2 intuitionistic non-membership degree
ANN	Aritifical neural network
AIFS	Atanassov Intuitionistic fuzzy set
AIFLS	Atanassov Intuitionistic fuzzy logic system
ANFIS	Adaptive neuro-fuzzy inference system
BMM	Begian-Melek-Mendel
DEKF	Decoupled extended Kalman filter
EKF	Extended Kalman filter
FL	Fuzzy logic
FLS	Fuzzy logic System
FOU	Footprint of uncertainty
FS	Fuzzy set
GCS	Gas compression system
GD	Gradient descent
GT2FS	Generalised type-2 fuzzy set
GT2FLS	Generalised type-2 fuzzy logic system
GT2AIFS	Generalised type-2 Atanassov-intuitionistic fuzzy set
FIS	Fuzzy inference system
IFIS	Intuitionistic fuzzy inference system
IT2	Interval type-2
IT2FS	Interval type-2 fuzzy set
IT2FLS	Interval type-2 fuzzy logic system
IT2AIFS	Interval type-2 Atanassov-intuitionistic fuzzy set

IT2AIFLS	Interval type-2 Atanassov-intuitionistic fuzzy logic system
IVFS	Interval-valued fuzzy set
IVAIFS	Interval-valued Atanassov intuitionistic fuzzy set
KF	Kalman filter
K−M	Karnik-Mendel
LFM	Linguistic fuzzy modeling
MF	Membership function
MSE	Mean sqaured error
MAE	Mean absolute error
NDEI	Non-dimensional error index
NMF	Non membership function
NMSE	Normalised mean squared error
NT	Nie-Tan
NSW	New South Wales
PFM	Precise fuzzy modeling
RMSE	Root mean squared error
SNR	Signal to noise ratio
std	Standard deviation
T1	Type-1
T2	Type-2
TSK	Takagi-Sugeno-Kang
T1FLS	Type-1 fuzzy logic system
T2FS	Type-2 fuzzy set
T2FLS	Type-2 fuzzy logic system
TR	Type reduction
TRS	Type-reduced set
trn	Training dataset
tst	Testing dataset

UoD	Universe of discourse

Chapter 1

Introduction

> There is nothing worse than a sharp image of a fuzzy concept.
>
> — Ansel Adams

Real world problems are fraught with a great deal of uncertainties. Over the years, there has been a growing interest in the formulation of theories and concepts to handle in effective and better ways, the effects of these uncertainties. Indeed, the presence of high level of uncertainty in every aspect of human lives and from a variety of platforms has provided a paradigm shift in uncertainty modelling. However, the underlying concepts in this respect is the concept of fuzzy sets [6].

The intention is not to replicate the significant body of work done in the area of fuzzy sets and systems. Rather, the primary focus in this research is to advance the frontiers of uncertainty modelling by developing a new framework that fuses the concept of two important generalisations of fuzzy sets namely interval type-2 fuzzy set (IT2FS) [7] and intuitionistic fuzzy set (IFS) in the sense of Atanassov [8]. With this new framework, the three distinct states of a phenomenon namely: membership, non-membership and indeterminate states can be separately and simultaneously addressed with the capacity for incorporating uncertainties.

1.1 Problem Statement

A fuzzy set (FS) is a generalisation of the classical notion of a set where an element belongs to a set to a certain degree. Contrary to classical sets with 0 or 1 membership, fuzzy sets are characterised by membership functions which define the degree of membership of an element to a fuzzy set. Fuzzy sets largely reflect a paradigm shift from the computation involving classical binary sets to approximate reasoning and computing with words [9], with the potential to capture an abundance of information and model vagueness, imprecision and uncertainty. The rationale behind FS stems from the facts that most human concepts are complex in nature and these concepts are not binary and have no associated objective measure [10]. A fuzzy set allows an entity to gradually move from full membership (with membership degree 1) to non-membership (with membership degree 0) and including everything in-between (partial membership). For the classical fuzzy set of type-1, the degree of non-membership is the complement of the membership.

However, because the membership degrees of a type-1 fuzzy set are precise in a referential set [0,1], they are not robust and do not handle uncertainties well in many applications. Zadeh in [7], proposed the concept of a type-2 fuzzy set (T2FS) where membership functions are themselves type-1 FS and with a third dimension description. Generally speaking, these T2FSs, in many instances, are able to cope with and manage uncertainty better than their type-1 counterparts with precise membership grades [11, 12]. Gorzalczany [13, pp. 2] in support maintained that:

> *"... it is not always possible for a membership function of the type* $\mu : X \to [0, 1]$ *to assign precisely one point from the interval* $[0, 1]$ *to each element* $x \in X$ *without loss of at least a part of information."*

and according to Gehrke [14, pp. 1]:

> *"··· But an increasingly prevalent view is that models based on [0,1] are inadequate. Many believe that assigning an exact number to an expert's opinion is too restrictive, and that the assignment of an interval of values is more realistic."*

In Mendel and John [15], it is conjectured that the additional degrees of freedom provided by the third dimension of a T2FLS lead to improved performance of a T2FLS over its type-1 counterpart. Recently, Mendel [16, pp. 2] pointed out that:

> *"it is the greater sculpting of the state space that lets an IT2 fuzzy system usually outperform a T1 fuzzy system, and a T1 fuzzy system usually outperform a crisp system."*

Type-2 fuzzy sets and systems can be classified into general type-2 fuzzy sets (GT2FSs) and interval type-2 fuzzy sets (IT2FSs). For GT2FS, both primary and secondary membership functions are all fuzzy. The third dimension (secondary membership) of a GT2FS has different magnitudes or weights which associate the amount of uncertainty to every point within the footprint of uncertainty (bounded-two dimensional region) of a GT2FS. The third dimension provides GT2FSs with additional design degrees of freedom than IT2FS and therefore have the potential to outperform a system that uses IT2FSs in the rule base [17]. Both GT2FSs and IT2FSs are 3-dimensional (3-D) structures. The only difference is that for an IT2FS, the memberships on the third dimension (secondary memberships) all take the value 1. The traditional GT2FSs are computationally intensive, difficult to use and understand [15] because of the computation involving the secondary memberships. According to Coupland and John, the full GT2FS requires large computational resources [18], and this may be impractical in real time application systems. Nevertheless, Liu [19] and Mendel *et al.* [17] have simplified the use of GT2FSs through the decomposition of a GT2FS into a set of α-planes, which are horizontal slices equivalent to IT2FSs. In this way, it is possible to represent a GT2FS as a union of 2-D α-planes, *"each of which is an IT2FS"* [17]. Thus, IT2FS stands as the state-of-the-art in uncertainty modelling and has been widely used [15, 20]. The α-plane representation of a GT2FS demonstrates that the computational cost of a GT2FS can grow in a linear fashion in relation to that of the IT2FS. It therefore implies, based on this premise, that a simpler and straight forward way to understand and use the GT2FS, is to first of all understand the operations performed upon IT2FS which are quite easy to implement [86]. In recent years, research has focussed mostly on IT2FSs [20–23, 86] which are quite practical with manageable computational intricacies since the secondary membership grades all take the value 1 [24]. The uncertainties about an IT2FSs are completely captured only on the bounded two-dimensional region, otherwise known as the footprint of uncertainties (FOUs) which are intervals. With the FOUs of IT2FSs, more information is retained and the loss of information is greatly reduced [25] as compared to T1FS. Mendel [26] argued that using IT2FSs to model linguistic uncertainties is scientifically correct whereas T1FS is not. The research presented in this thesis adopts the principles of the simpler and widely used IT2FS. It is believed that by first understanding this research from the perspective of IT2FS, will pave the way for the use of GT2FS in the future to further explore the ideas presented in this thesis.

The use of IT2FSs to model uncertainties in data cannot be over-emphasized as there exists abundance of applications involving IT2FLSs which employ at least one IT2FS in the rule base. The reader is referred to [27–32] for a comprehensive review of IT2FLS applications. The key advantage of IT2FSs is that the membership functions of these sets are themselves fuzzy where the actual degrees of membership are assumed to belong. The IT2FSs have a greater capability to model imprecise and imperfect information because of the extra degrees of freedom provided by their footprints of uncertainties (FOUs). That is, IT2FSs are quite useful in cases where it is difficult to specify a single crisp numeric membership function value and where linguistic and numerical uncertainties abound, particularly in many real world applications.

According to Wu [33], one of the reasons for the wide spread use of IT2FLSs is the fact that T2FLSs provide a better way of modelling intrapersonal uncertainties (the uncertainty a person has about the word [26]) and interpersonal uncertainties (uncertainty that a group of people have about the word [26]) which are intrinsic to natural language because their membership functions are uncertain. This increases the robustness of the system. Also, IT2FLSs are adaptive with the ability to model complex input-output relationships better than their type-1 counterparts. For a more detailed advantages of using IT2FLSs, the readers are referred to [33, 34]. Despite the advantages of IT2FSs, the extensive use of IT2FLSs and their abilities to handle uncertainties in data better than their type-1 counterparts, they still make use of only the membership functions (upper and lower) to model these uncertainties. For IT2FS, the non-membership for the lower membership function is complementary to the upper membership function and non-membership for the upper membership function is complementary to the lower membership function. These kinds of fuzzy sets (type-1 and type-2) are also known as the complementary fuzzy sets [35].

In a real life scenario, it is not always the case that the non-membership grade of an element to a set is complementary to the membership (upper or lower). There tend to be some extra degrees that represent evidence of neither membership nor non-membership, otherwise known as hesitation or indeterminate degree of an element to a set.

Traditional IT2FLSs lack the mechanism for tackling this phenomenon. This research is an attempt to address this drawback by incorporating Atanassov's IFS (non-membership function and hesitation degrees) into IT2FS. Thus, with the ability of IT2FSs to adequately capture the uncertainties in their FOUs and the ability of IFS to separately cater for the membership and non-membership grades of elements with extra degrees of hesitancy, the integration of these two concepts is adopted to design a new type-2 fuzzy framework for uncertainty modelling.

1.2 Background and Motivation

During the first year of this research, experiments utilising different machine learning approaches, particularly, decision tree, support vector machine and artificial neural network (ANN) were conducted. Suffice it to say that towards the end of the first year, a paper titled "Machine Learning and Statistical Approaches to Classification - A Case Study" was presented at the 15th UK Workshop on Computational Intelligence, UKCI 2015, Exeter. Based on the outcome of these experiments, ANN was adopted as a viable learning method for the intended model. The main focus of this thesis is to advance the frontiers of uncertainty modelling by integrating Atanassov's intuitionism (non-membership and intuitionistic fuzzy indices) into interval type-2 fuzzy logic system (IT2FLS) with the aim of investigating the capacity of uncertainty modelling using both the membership and non-membership function FOUs of a set.

Conventional fuzzy systems make use of type-1 or type-2 FSs. A Type-1 FS with precise membership grades cannot fully handle the level of uncertainty inherent in many real world applications. The reason is that once the membership grades of a type-1 FS are chosen, uncertainty disappears, leaving crisp numerical values. The type-2 FSs with upper and lower membership functions do handle uncertainties in many applications better than their type-1 counterparts. These typical approaches to uncertainty modelling rely solely on the membership function of an entity, x - i.e. $\mu(x)$. The underlying assumption being that the non-membership function ($\nu(x)$) of the element is complementary to the membership function. (i.e. $\nu(x) = 1 - \mu(x)$). As earlier discussed, it is not always the case that the non-membership grade of an element to an FS is complementary to the membership.

Before continuing the discussion, let us consider a typical scenario to drive home this thought process and to elaborate more on the motivation. Perhaps the best way to illustrate this is in the words of Zadeh [36, pp. 4] in the $50th$ anniversary of FS:

> *""Fuzzy set", ..., What is of historical interest is that initially – and for sometime thereafter – my paper was an object of indifference, skepticism and derision. ⋯ In contrast, my ideas were welcomed with open arms in Japan".*

Fuzzy set theory was a new theory where people had to express their individual opinions concerning the theory and its veracity. Apparently, some people were in support of the theory to a certain degree (Group 1 - those in Japan), some were in opposition of the theory to a certain degree (Group 2 - skeptism and derision) while some people abstained from making comments on FS either totally or partially (Group 3 - indifference). The assumption for these criticisms, according to Zadeh could have been due to lack of understanding of the theory of FS [37], or entire lack of interest in the concept. As reflected in the scenario, using a binary logic of [0,1] to classify these three groups of people simply as supporters or opponents of the theory will be too hasty and misleading. Furthermore, by insisting that the assessment (supporters and opponents) is exactly complementary is arguably too committing [38]. As earlier pointed out, for classical fuzzy sets, the single membership degree is assumed to include not only the state of membership for an entity but also the state of non-membership. This is arguably an unrealistic assumption. Thus, FS with only membership function definition, apart from lacking the mechanism of separately capturing the degree of non-membership, also cannot represent the state of "neither support nor opposition" [39] of an entity to a fuzzy set, a characteristic which is termed the indeterminate state (degree of indeterminacy or hesitation). In this thesis, the terms intuitionistic fuzzy indices, degrees of indeterminacy, and hesitation degrees are used interchangeably.

Atanassov [8] in 1986 introduced a new kind of fuzzy set, the so-called Atanassov-intuitionistic fuzzy set (AIFS), which is characterised by independently defined membership function and non-membership function together with some degree of indeterminacy. The AIFS therefore defies the claim of the FS that $\mu + \nu = 1$[1] "$\cdots$ *thus relaxing the enforced duality that* $\nu = 1 - \mu$ *from fuzzy set theory*" [38, pp. 1] and maintains a set whose sum of membership function and non-membership function is less than or equal to 1. With the degrees of membership function, non-membership function and hesitation, the AIFS becomes more meaningful in the context of human reasoning and natural language representation [40]. A typical example is voting, where some people will vote for, vote against or abstain from voting [41].

It is argued here that by exploiting and integrating the capabilities of IT2FS and AIFS in a FLS, a framework that is more robust and more efficient for uncertainty modelling can be realised which could lead to the attainment of as accurate an estimate as possible under uncertainty. The salient discussions above have motivated the design of an interval type-2 Atanassov intuitionistic fuzzy logic system (IT2AIFLS) with the aim of modelling uncertainty better. The marriage of these two concepts - AIFS and classical IT2FS - is able to

[1] μ represents degree of membership, ν represents degree of non membership

provide a synergistic capability in dealing with imprecise and vague information [3]. What is more, with this approach, the evaluation of concepts becomes more precise and close to human reasoning than classical type-1 FLSs and T2FLSs. The proposed IT2AIFLS framework is a Takagi-Sugeno-Kang (TSK) inference system that employs modified Gaussian function with uncertain standard deviation. Kayacan and Khanesar [42] have pointed out that the Gaussian membership functions with uncertain standard deviations is the only known membership function that is differentiable at all points and they have been one of the first choices in the design of T2FLSs for many applications. In this research, the number of membership and non-membership of Gaussian functions is restricted to two in order to ease the computational burden of the system. The number of membership (non-membership) can be increased up to eleven depending on the application and the number of inputs [43]. However, Chopra *et al.* [43] argued that increasing the number of membership (non-membership) functions as well as the rules beyond a certain limit is useless; as doing this only increases the complexity of the system with almost no effect on the output.

To harness the performance of fuzzy logic systems (FLSs), a variety of enabling technologies such as ANNs have been incorporated. The two approaches - fuzzy logic and ANN are known to be universal approximators [44, 45] that can identify and approximate any nonlinear systems to any arbitrary degree of accuracy. The integration of fuzzy logic and ANN merges the advantages of both approaches in a synergistic manner in terms of the excellent generalisation and learning capability of ANN and the ability of FLS to simultaneously and effectively handle uncertainties and imprecise information; and to approximately reason with these information. To show the efficiency of the proposed approach, different simulation studies have been considered using publicly available benchmark datasets (artificial and real world) and another real world dataset (commercially sensitive) obtained from a Nigerian-based power plant. It is worth mentioning at this point that most of the datasets used in this thesis are time series datasets. The reason is to aid comparison with previous works in the literature which are mostly based on time series analysis. However, it will be interesting to investigate the effects of this new model from a human knowledge modelling perspective using survey data that captures all three concepts namely: membership, non-membership and intuitionistic fuzzy indices. The intuitionistic fuzzy indices is of great importance because from the point of view of say, voter behaviour analysis, for instance, indeterminate voters (those who abstain) after proper enhancements and supports can finally vote for or vote against a product or proposal.

1.3 Contributions

The key contributions of this thesis are:

- A general framework that introduces Atanassov's non-membership functions and intuitionistic fuzzy indices (IF-indices) into IT2FS with the aim of capturing more uncertainties in data and enabling hesitation. This framework is henceforth referred to as the interval type-2 Atanassov's intuitionistic fuzzy set (IT2AIFS).
- Formulation of a new and enhanced class of IT2FLS based on Takagi-Sugeno-Kang (TSK)-fuzzy inference using IT2FS and Atanassov IFS otherwise known as the interval type-2 Atanassov-intuitionistic fuzzy logic system (IT2AIFLS-TSK).
- Exploiting the use of both membership and non-membership functions that are intervals with intuitionistic fuzzy indices for uncertainty modelling.
- Investigating the possibility of embedding ANN into the new fuzzy logic framework in order to assess for the first time its applicability in the learning process of the proposed framework. The capabilities of gradient descent (GD), a first-order derivative based learning method is exploited for the first time in tuning the parameters of the new framework. The developed approach is applied to well-known publicly available benchmark time series and regression problems of diverse instances and domains. Detailed description of these procedures and applications are presented in Chapter 4.
- Encouraged by the previous results, focus is shifted to the second-order derivative optimisation methods in order to tackle the drawbacks of the first-order derivative method and improve on the efficiency of the system. A variant of extended Kalman filter (EKF), a second-order derivative-based optimisation method, known as the decoupled EKF (DEKF) is exploited for the first time to assess the efficiency of the proposed model in terms of convergence and prediction accuracy. The full detail of this learning procedure and evaluation is presented in Chapter 5.
- Evaluation of the proposed model with alternative models such as classical IT2FLS and type-1 AIFLS to assess their statistical significance. Detailed desription and evaluation are presented in Chapter 5.

- Fusion (hybridisation) of two FLS optimisation methods (DEKF and GD) to assess their combined effects on the parameter tuning of the new proposed methodology for the first time. A detailed description of this hybrid approach is presented in Chapter 6.
- Tuning the contributions of the membership and non-membership in order to manage varying degrees of uncertainties in the rule base of the proposed framework.

1.4 Academic Publications

The following publications were produced as a direct result of the work undertaken during the course of conducting this research:

1. I. Eyoh, R. John and G. De Maere, "Interval type-2 intuitionistic fuzzy logic system for non-linear system prediction," in *2016 IEEE International Conference on Systems, Man and Cybernetics (SMC)*, Budapest, Hungary, pp. 1063-1068, 2016 [2].

   This paper presents the first published results on the evaluation of the proposed model with gradient descent (GD) learning algorithm. The simulation studies are done using two well known benchmark datasets namely Mackey-Glass time series and a synthetic dataset. These experiments show the effectiveness of the proposed approach on non-linear prediction problems as it closely modelled the input-output relationship of the data well with reduced root mean squared error. Detailed description is presented as part of Chapter 4.

2. I. Eyoh, R. John and G. De Maere, "Interval Type-2 A-Intuitionistic Fuzzy Logic for Regression Problems," *IEEE Transactions on Fuzzy Systems, DOI:10.1109/TFUZZ.2017.2775599* [3].

   Encouraged by the first experimental studies, the paper is extended for a journal publication. In this paper, analysis of publicly available regression datasets is considered. The results are compared with existing studies using the same benchmark datasets and computational set-ups. The discussion on the strength of the proposed model and its weaknesses are provided in this paper together with ways of enhancing the model further for the full utilisation of its capabilities. Detailed analysis is presented in Chapter 4.

3. I. Eyoh, R. John and G. De Maere, "Time Series Forecasting with Interval Type-2 Intuitionistic Fuzzy Logic Systems," in *2017 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*, Naples, Italy, pp. 1-6, 2017. **(Recommended for the best paper award**) [46].

   This paper presents the effectiveness of the proposed framework on time series problems. Three time series problems are analysed and results are compared with similar studies in the literature. Analysis of simulation results reveal an improvement in the performance of the proposed approach. Detailed description is presented in Chapter 5.

4. I. Eyoh, R. John and G. De Maere,"Extended Kalman Filter-based Learning of Interval Type-2 Intuitionistic Fuzzy Logic Systems," in *2017 IEEE International Conference on Systems, Man and Cybernetics*, Banff Center, Banff, Canada, pp. 728-733, 2017 [47].

   This publication introduces the decoupled extended Kalman filter (DEKF) for the tuning of the parameters of IT2AIFLS-TSK fuzzy inference for the first time. The analysis is conducted using real world dataset from Australia's electricity market. The IT2AIFLS-DEKF is compared with its type-1 variant and classical IT2FLS. Analysis of results reveal performance superiority of IT2AIFLS trained with DEKF over IT2AIFLS trained with gradient descent. The proposed IT2AIFLS-DEKF also outperforms its type-1 variant and IT2FLS on the same learning platform. Detailed description is presented in Chapter 5.

5. I. Eyoh, R. John, G. De Maere and K. Erdal, "Hybrid Learning for Interval Type-2 Intuitionistic Fuzzy Logic System as Applied to Identification and Prediction problems," *IEEE Transactions on Fuzzy Systems, DOI 10.1109/TFUZZ.2018.2803751*

   In this publication, the DEKF and GD are combined to produce a hybrid learning algorithm for tuning the parameters of the proposed model for the first time. The learning strategy is applied to the identification and prediction of well known and widely used benchmark datasets and results are compared with similar studies in the literature using the same computational settings. Detailed description is presented in Chapter 6.

6. I. Eyoh, R. John and G. De Maere, "Interval Type-2 Intuitionistic Fuzzy Logic Systems: A Comparative Evaluation" (*17th Information Processing and Management*

of Uncertainty in Knowledge-Based Systems Conference (IPMU), 2018. Accepted).

In this paper, the assessment of IT2AIFLS with alternative fuzzy logic systems such as classical IT2FLS and AIFLS is considered with the aim of evaluating their statistical significance. The parameters of the models are tuned using the DEKF algorithm. From the simulation results, IT2AIFLS performs significantly better than the classical IT2FLS and AIFLS. Analysis of results also shows that there is no significant difference between the classical IT2FLS and AIFLS. Detailed description is presented in Chapter 5.

7. I. Eyoh and R. John, "Machine Learning and Statistical Approaches to Classification: A case Study," in *proceedings of the 15th UK Workshop on Computational Intelligence,*" UKCI 2015, Exeter, UK.

   This paper presents a prerequisite study to the understanding of the workings of ANN - a popular machine learning approach that allows for the adaptive tuning of the parameters of FLSs.

1.5 Thesis Outline

The discussion in the remaining chapters is outlined as follows:

- In Chapter 2, a précis of the techniques exploited in building a concrete realisation of the framework proposed in this thesis is exploited. A survey of related work in uncertainty modelling using IT2FLS and AIFLS is provided. Discussions on the general drawbacks of classical IT2FLSs and AIFLS to uncertainty modelling are put forward. Existing fuzzy logic approaches that attempt to address these drawbacks are also discussed. These techniques underpin the model presented in Chapter 3.

- In Chapter 3, the proposed interval type-2 Atanassov intuitionistic fuzzy logic system is formulated. The differences between existing interval-valued Atanassov intuitionistic fuzzy sets (IVAIFS) and the new framework proposed in this thesis, the so-called IT2AIFS are highlighted. The different components of the developed architecture are also discussed.

- In Chapter 4, a critical evaluation of the proposed model is performed. The model is evaluated on same datasets and computational set-ups similar to other works in the literature. The analysis of the model is done using first-order derivative-based learning algorithms namely GD. It is shown that by using interval membership and non-membership functions with embedded hesitation indices, the error of prediction can be significantly reduced. The assessment is done using publicly available benchmark time series and regression problems (both artificial and real world).

- In Chapter 5, the decoupled extended Kalman filter (DEKF) is used to optimise the parameters of the proposed model. The new proposed model trained with DEKF is evaluated on a synthetic dataset and two real world datasets namely, New South Wales electricity load and a gas compression system (GCS) dataset of a gas turbine obtained from a Nigerian-based power plant. Also, in this chapter, the statistical significance between the model proposed in this thesis and other alternative models such as classical IT2FLS and type-1 AIFLS is investigated.

- In Chapter 6, the parameters of the proposed framework are tuned using hybrid algorithm of DEKF and GD. The resulting hybrid model is applied to system identification and prediction problems with encouraging results.

- In Chapter 7, a critical discussion of the research conducted in this thesis is presented. The contributions to knowledge contained in this thesis are highlighted. A reflection on ways to improve and assess them further is provided, and finally wrapped-up with a summary of the thesis.

Chapter 2

Related Work

It is better to be approximately right
than precisely wrong.

Warren Buffett

2.1 Introduction

To make this thesis self-contained, important underlying concepts exploited in building a concrete realisation of the proposed framework are reviewed in this chapter. These concepts underpin the new model proposed in this thesis. In Section 2.2, some important background on the notion of type-1 FSs are discussed. A précis of some of the generalisations of fuzzy sets, namely type-2 fuzzy sets (T2FSs), Atanassov intuitionistic fuzzy sets (AIFSs), interval-valued fuzzy sets (IVFSs) and interval valued Atanassov IFSs (IVAIFS) are also provided in order to differentiate the specific concepts of interval-valued (classical and intuitionistic) from the much broader concept of interval type-2 (classical and intuitionistic) fuzzy sets. Section 2.3 provides a detailed theoretical study of interval type-2 fuzzy sets, after which Section 2.4 presents the different aspects that constitute the interval type-2 FLSs. Section 2.5 discusses uncertainty modelling detailing different forms and sources of uncertainty. In Section 2.6, a brief overview of the different design methodologies for optimising the parameters of interval type-2 FLSs is presented, with particular focus on gradient descent and Kalman filter-based methods. In Section 2.7, some application areas of IT2FLSs are reviewed, followed by the drawbacks of IT2FLSs in Section 2.8. Atanassov intuitionistic fuzzy set is discussed in Section 2.9, followed by its practical applications in Section 2.10 and possible drawbacks in Section 2.11. As this research involves the integration of AIFS and IT2FS, a review of existing studies using both AIFS and IT2FS is presented in Section 2.12. Finally, Section 2.13 presents a summary and critique of the chapter. The intention is to motivate the model described in Chapter 3.

2.2 Fuzzy Set Theory

Fuzzy set (FS) was introduced by Zadeh [6] as a generalisation of the classical notion of a set. Belohlavek *et al.* [48] argued that the main motivation behind the generalisation of classical set to FS is to allow representations of concepts that have no sharp boundaries in a rigorous, mathematical way. According to Zadeh [49, pp. 3176],

> *"Fuzzy logic is a precise conceptual system of reasoning, deduction and computation in which the objects of discourse and analysis are, or are allowed to be, associated with imperfect information. Imperfect information is information which in one or more respects is imprecise, uncertain, incomplete, unreliable, vague or partially true."*

Zimmermann [50, pp. 318] lent credence to this when he pointed out that:

> *"Fuzzy set theory provides a strict mathematical framework (there is nothing fuzzy about fuzzy set theory!) in which vague conceptual phenomena can be precisely and rigorously studied. It can also be considered as a modelling language, well suited for situations in which fuzzy relations, criteria, and phenomena exist."*

For decades now, FS has served as an effective tool for handling uncertainty (fuzziness or vagueness) and computing with words [9]. The key idea underlying fuzzy logic is the use of linguistic variables rather than numbers to describe natural language phenomena such as voter turnout, age and temperature. A linguistic variable is a variable having words as their values rather than numbers [37] and the collection of these linguistic variable names (words) are called linguistic terms. For instance, the linguistic variable "vote" received each day on a proposal can have linguistic terms such as "low, medium and high."

Thus FS presents a paradigm shift from the use of numbers to the use of words. Ever since its introduction, FSs and systems have been successful in many fields and application domains ranging from real life applications to commercial products. It has found usefulness in a wide range of problems such as control [51–53], time series [54–56], classification and prediction [57, 58], decision making [59–61], load forecasting [62–65] and more. Excellent reviews of applications of FLSs can be found in [66–69]

2.2.1 Type-1 Fuzzy Set: Definition

A type-1 fuzzy set, A, is characterised by a membership function that determines the degree of membership of every element $x \in X$ and is represented as [6]:

$$A = \{(x, \mu_A(x)) \quad | \, \forall x \in X\} \tag{2.1}$$

The associated non-membership function degree of x in a FS

$$A = \{(x, \mu_A(x), \nu_A(x)) \quad | \, \forall x \in X\} \tag{2.2}$$

may therefore be formulated as a complement of the membership degree as below:

$$A = \{(x, \mu_A(x), 1 - \mu_A(x)) \quad | \, \forall x \in X\} \tag{2.3}$$

Hence, the non-membership degree of a classical FS is complementary to the membership degree. Alternatively, the FS A can be represented as:

$$A = \int_{x \in X} \mu_A(x)/x \tag{2.4}$$

for a continuous universe of discourse (UoD) or

$$A = \sum_{x \in X} \mu_A(x)/x \tag{2.5}$$

for a discrete UoD.

where $\int$ and $\sum$ denote a collection of all admissible points in the UoD. As shown in Figure 2.1, fuzzy sets of type-1 are two dimensional. Once the membership function value is chosen, the uncertainty disappears because the membership degrees of T1FS are completely precise. For example, the FS 'vote' received each day on a proposal may be represented as:

$$vote = 0.58/Day1 \; + \; 0.85/Day2 \; + \; 0.24/Day3 \tag{2.6}$$

where the + sign denotes the collection of all points in the UoD while the / sign links each element (Day) with its corresponding membership grade. A fuzzy logic system that utilises T1FSs is referred to as a type-1 FLS.

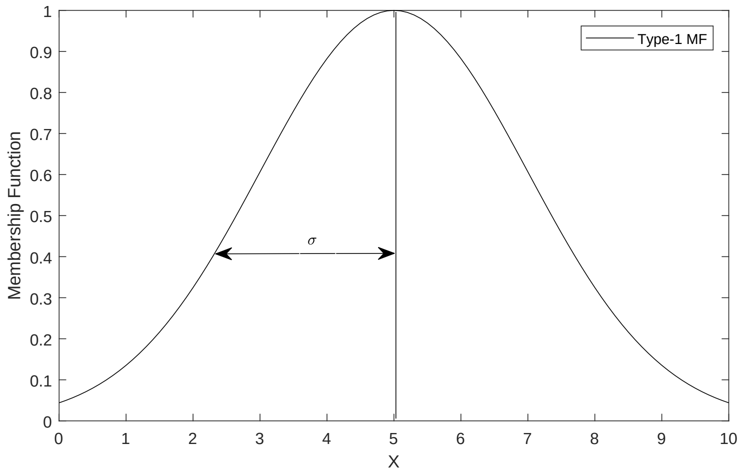

Figure 2.1: A Gaussian type-1 membership function

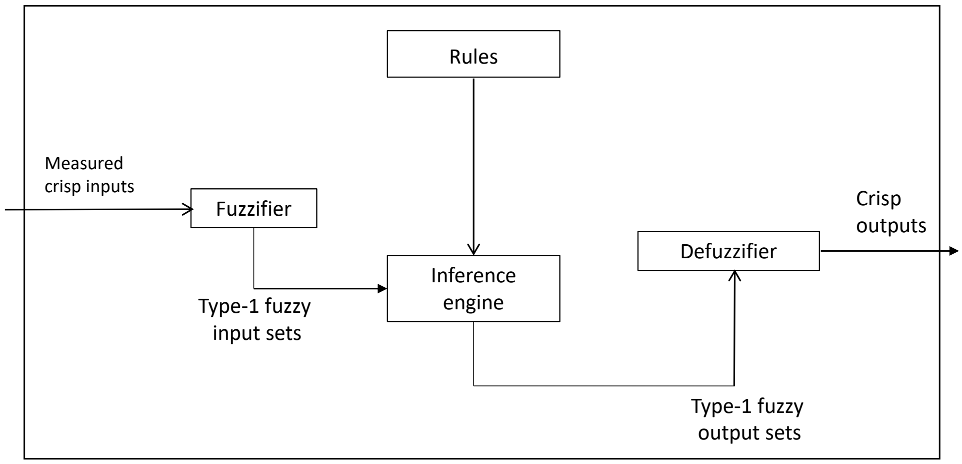

Figure 2.2: A type-1 FLS [1]

2.2.2 Type-1 Fuzzy Logic Systems

A type-1 FLS consists of a fuzzifier, a rule base, a fuzzy inference engine and a defuzzifier (see Figure 2.2). The fuzzifier takes every crisp input $x \in X$ and maps them into a fuzzy set. The problem is broken down into sets of rules. Generally, the rule for a type-1 FLS may be represented as:

$$R_k : \text{IF } x_1 \text{ is } A_{1k} \text{ and } x_2 \text{ is } A_{2k} \text{ and } \cdots \text{and } x_n \text{ is } A_{nk} \text{ THEN } y_k \text{ is } F_k \tag{2.7}$$

for Mamdani fuzzy inference systems and

$$R_k : \text{IF } x_1 \text{ is } A_{1k} \text{ and } x_2 \text{ is } A_{2k} \text{ and } \cdots \text{and } x_n \text{ is } A_{nk} \text{ THEN } y_k = \sum_{i=1}^{n} w_{ik}x_i + b_k \tag{2.8}$$

for TSK-type fuzzy inference systems.

where x_i's $(i = 1 \cdots n)$ are inputs, k is the number of rules, A_{nk} are antecedent type-1 FSs and y_k is the output representing a lingustic term (Mamdani FLS) and a function (TSK FLS) respectively. The inference engine combines these rules using any $t - norm$ usually a product or minimum $t - norm$ to produce a mapping from a type- 1 fuzzy input sets to a type-1 fuzzy output sets which are defuzzified into a final crisp output in the case of Mamdani FLS. Ever since the introduction of FS, many generalisations of FS theory have been proposed. According to Deschrijver and Kerre [70, pp. 227], *"Some of these theories are extensions of fuzzy set theory, others try to handle imprecision and uncertainty in a different (better?) way."*

2.2.3 Generalisation of a Fuzzy Set

Over the years, several generalisations or extensions of a fuzzy set have emerged. The focus of many such FSs have been on the need for appropriate representation of concepts described through imperfect information, as well as the representation of the lack of knowledge or uncertainty of the experts in a different way [71]. These FS extensions include Ł-fuzzy set [72], type-2 fuzzy set (T2FS) [7], interval-valued fuzzy set (IVFS) [13], Atanassov intuitionistic fuzzy set (AIFS) [8], interval-valued intuitionistic fuzzy set (IV-IFS) [73], grey set [74], vague set [75], hesitant fuzzy set [76] and neutrosophic fuzzy set [77]. While a few of these generalisations of FS are listed, a comprehensive and detailed discussion of the FS extensions are reported elsewhere. Bustince *et al.* [71], for example, discussed the history of fuzzy set extensions as well as their relationships. Studies show that IVFS are isomorphic to AIFS [73, 78, 79] and AIFSs are sometimes referred to as grey set [74]. After the introduction of vague set in [75], Bustince and Burillo [80] pointed out that vague sets are AIFSs. In the literature, IVFSs are also regarded as special cases of T2FSs [70, 71, 78, 81–84]. Specific discussion of the varieties of these generalisations of FS is beyond the scope of this research. Rather, the focus is on T2FS (IT2FS) and AIFS; the generalisations of FSs that underpin the contributions of this research.

As earlier discussed in Section 1.1, a T2FS consists of two instances namely: GT2FS and IT2FS. Next, a formal definition of a GT2FS is given.

Definition 2.2.1 A GT2FS is characterised by a type-2 membership function, $\mu_{\tilde{A}}(x, u)$ for all $x \in X$ and $u \in J_x \subseteq [0, 1]$ [15]

$$\tilde{A} = \left\{((x, u), \mu_{\tilde{A}}(x, u)) \mid \forall x \in X, \forall u \in J_x \subseteq [0, 1]\right\} \tag{2.9}$$

where x is the primary variable, u is the secondary variable and J_x is the support of the secondary membership function in the third dimension of x. The GT2FS is a three dimensional structure, where the third dimension provides extra degrees of freedom to GT2FS to directly model uncertainties [12, 15]. The main characteristic of a GT2FS is that the third dimension is fuzzy, otherwise known as the secondary membership function. This representation for GT2FS makes it very difficult to manage and understand with increased computational cost [15]. Although the FOUs of the GT2FSs together with the third dimension of membership functions provide GT2FS with extra design degrees of freedom to handle uncertainties effectively, they are very complex and rarely used in many applications [15]. Coupland and John [85] opined that GT2FS are powerful modelling tool, yet they remain impractical for approximate reasoning (until recently see Section 1.1). Having said these, there exist a simplified version of a T2FS, the so-called IT2FS. Hence, when all the secondary membership, $\mu_{\tilde{A}}(x, u)$ of a T2FS is equal to 1, an IT2FS is obtained.

2.3 Interval Type-2 Fuzzy Set

A simpler version of the T2FS, called the interval type-2 fuzzy sets (IT2FSs), is a fuzzy set where the secondary membership values are all unity, thus reducing the burden of working with the third dimension values and reducing the computational cost. With the interval representation of a T2FS, it is possible to project the interval T2FS onto a two dimensional (2-D) plane by capturing the uncertainties using only the FOU. The FOU which is the union of all primary memberships is a bounded region that represents the uncertainty in the primary memberships of an IT2FS (see Figure 2.3). The FOU size determines the amount of uncertainty captured by the IT2FSs. The wider the FOU, the more uncertain there is about the primary memberships. An upper membership function and a lower membership function are two type-1 membership functions that form the bounds for the FOU of an IT2FS [12]. The definition of the IT2FS by only the membership function on a 2-D plane simplifies its usage and in the words of Mendel [86, pp. 22], *"Almost all applications use IT2 FSs because, to date, it is only for such sets (and systems that* *use them) that all calculations are easy to perform."* Mendel *et al.* [24] provided a sound mathematical framework that simplifies the use of IT2FSs.

Definition 2.3.1 An IT2FS is specified by a footprint of uncertainty circumscribed by a lower membership function, $\underline{\mu}_{\tilde{A}}(x,u)$ and an upper membership function, $\bar{\mu}_{\tilde{A}}(x,u)$ for all $x \in X$.

$$\tilde{A} = \left\{ ((x,u), \underline{\mu}_{\tilde{A}}(x,u), \bar{\mu}_{\tilde{A}}(x,u)) \mid \forall x \in X, \forall u \in J_x \subseteq [0,1] \right\} \tag{2.10}$$

where $\underline{\mu}_{\tilde{A}}(x,u) = 1$ and $\bar{\mu}_{\tilde{A}}(x,u)$ =1. Thus, the IT2FS can also be expressed as:

$$\tilde{A} = \int_{x \in X} \int_{u \in J_x} 1/(x,u) \qquad J_x \in [0,1] \tag{2.11}$$

or

$$\tilde{A} = \sum_{x \in X} \sum_{u \in J_x} 1/(x,u) \qquad J_x \in [0,1] \tag{2.12}$$

where $\int$ and $\sum$ represent the union of all admissible points in a continuous and discrete UoD respectively [24]. For instance, the interpretation of IT2FS for 'vote' in Subsection 2.2.1 maybe expressed as:

$$vote = Medium/Day1 \;\; + \;\; High/Day2 \;\; + \;\; Low/Day3 \tag{2.13}$$

where the linguistic terms $Medium$, $High$, and Low are themselves fuzzy sets (two type-1 FSs each) signifying medium, high and low number of votes for the three days repectively. In this thesis, a finite UoD is assumed.

2.3.1 Comparison Between Interval Type-2 Fuzzy Set and Interval Valued Fuzzy Set

In the literature, IVFSs [13] are regarded as the special cases of IT2FSs [71, 81, 83, 84]. Specifically, and more notably is the work of Bustince *et al.* [84] which demonstrates in-depth, a wider and general view of the relationship between IT2FSs and IVFSs. Many people often believe that IVFS is equivalent to IT2FS, but according to [84], IVFSs are a special case of IT2FSs and as such both kinds of fuzzy sets should be treated differently. In their paper, four representations are defined for the primary membership functions of IT2FSs namely, as type-1 fuzzy sets, as interval-valued fuzzy sets, as multi-fuzzy sets and as multi-interval fuzzy sets. Thus, IT2FSs can easily be used to model other concepts, a

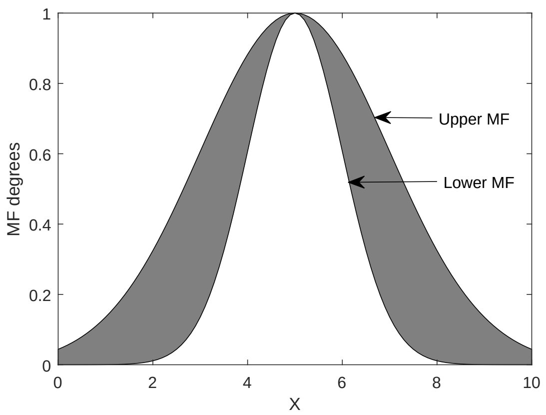

Figure 2.3: A Gaussian interval type-2 membership function

capability not obtainable with IVFSs [84, 87]. In this thesis, the interval-valued representation of the IT2FSs is adopted. A FLS that utilises one or more IT2FSs is referred to as an IT2FLS.

2.4 Interval Type-2 Fuzzy Logic Systems

The architectural block of an IT2FLS, shown in Figure 2.4, consists of the fuzzifier, fuzzy inference, fuzzy rule base and the output processing block. This is similar to the T1FLS. The only difference is the output processing module of the T2FLS which consists of the type-reducer and the defuzzifier as opposed to only the defuzzifier found in the T1FLS architecture. As shown in the block diagram of a T2FLS, the external crisp inputs are first fuzzified into T2FSs (IT2FSs in this case). The IT2FSs generated activate the inference engine and the rule base to produce IT2FSs as the outputs. These IT2FSs are then reduced to T1FSs which are finally defuzzified into crisp outputs. Below are the detailed description of the workings of each process module.

2.4.1 Fuzzification Process

There are two fuzzification procedures namely: singleton and non-singleton. The fuzzification process involves the mapping of a crisp numeric input vector with multiple inputs $x \in X$ into IT2FSs $\tilde{A}$ in X which activate the inference engine. In IT2FLSs, the join ($\sqcup$)

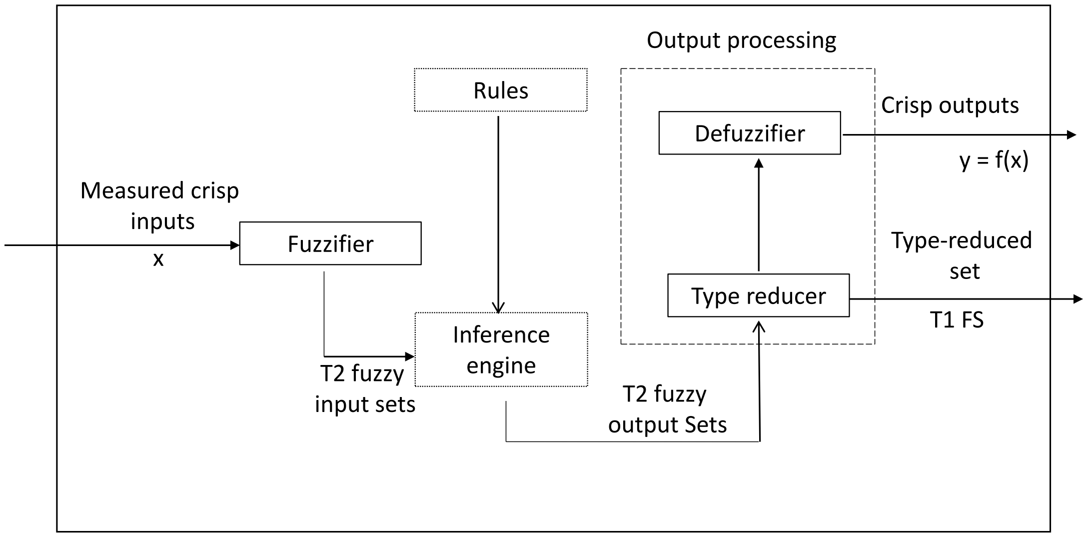

Figure 2.4: A T2FLS structure [1]

and meet operators ($\sqcap$), replace the union and intersection operators of TIFLS.

2.4.2 Rules

The rule representation of an IT2FLS is similar to a T1FLS. The only difference between these two types of fuzzy sets is with the nature of the membership function and this is not relevant during rule formation [12]. However, for IT2FLSs, IT2FSs are used in the antecedent and/or consequent parts of the rules. A general type-2 rule can be expressed as:

R_k: IF x_1 is $\tilde{A}_{1k}$ and x_2 is $\tilde{A}_{2k}$ and $\cdots$ and x_n is $\tilde{A}_{nk}$ THEN y_k is $\tilde{F}_k$

where $\tilde{A}_{1k}$,$\tilde{A}_{2k}$, $\cdots$,$\tilde{A}_{ik}$,$\cdots$,$\tilde{A}_{nk}$ are antecedent IT2FSs and y_k is the output of the kth rule which is another consequent IT2FS ($\tilde{F}_k$), w_{ik}'s are the consequent coefficient with offset b_k ($k = 1 \cdots M$). An IT2FLS has at least one IT2FS in the antecedent or consequent parts of the “IF ... THEN” rule.

2.4.3 Fuzzy Inference Engine

The inference engine combines rules and maps input IT2FSs to output IT2FSs. There are generally two main types of fuzzy inferencing namely: Mamdani and TSK which differ in their representation and output evaluation and ultimately influence their level of accuracy and interpretability. The consequent parts of Mamdani fuzzy inference are fuzzy sets while the consequent part of TSK fuzzy inferencing are linear functions of the inputs. Depending on the user's requirements, two fuzzy modelling suffice namely linguistic fuzzy modelling (LFM - Mamdani) and precise fuzzy modelling (PFM - TSK) [88, 89]. Whilst Mamdani uses defuzzification to obtain the final output of the fuzzy systems, TSK-type fuzzy inference uses weighted average to compute the final output. Thus, Mamdani fuzzy inference entails substantial computational cost because of the time consuming defuzzification procedure. The TSK inference is therefore more computationally effective and particularly works well with optimisation and adaptive techniques such as ANN used in this thesis. Based on these premises, a TSK fuzzy inference is adopted in this research.

There are basically three models for generating the output of a type-2 TSK inference system namely [42, 90]:

- Model I: The antecedent parts are type-2 fuzzy set while the consequent parts are type 1 fuzzy sets denoted by A2-C1.
- Model II: The antecedent parts are type-2 fuzzy sets with crisp numbers as consequents denoted by A2-C0.
- Model III: Both the antecedent and consequent parts are T1 fuzzy sets represented as A1-C1.

Models I and II use IT2FSs in the antecedent parts and thus have more degrees of freedom to model uncertainties and ultimately minimise their effects in data modelling. In this work, an A2-C0 TSK fuzzy inferencing is assumed. The Takagi-Sugeno [91] and Sugeno-Kang [92] - TSK fuzzy model have been extensively adopted for fuzzy modelling with great success. A Type-2 TSK fuzzy logic system first proposed in [90] makes it possible to handle linguistic uncertainties effectively. In particular, the intention is to use the proposed model to closely approximate the input-output relationship of a system, hence, TSK fuzzy inferencing becomes the most suitable and appropriate for the proposed framework. The IT2-TSK fuzzy models use IT2FS to capture uncertainty with respect to the assignment of membership function and describe the level of uncertainty in the antecedent and/or consequent parts of a fuzzy logic system. For A2-C0 fuzzy model, IT2FS are used in the antecedent while the consequent is expressed as a linear combination of the inputs. The A2-C0 can be described by IF $\cdots$ THEN rules as:

$$R_k : IF\, x_1 \;\; is \;\; \tilde{A}_{1k} \text{ and } \; x_2 \;\; is \;\; \tilde{A}_{2k} \;\; \text{and} \;\; \cdots \text{and} \;\; x_n\, is\, \tilde{A}_{nk} \;\; \text{THEN}\, y_k = \sum_{i=1}^{n} w_{ik}x_i + b_k \tag{2.14}$$

where $x_1 \cdots x_n$ are the inputs variables, w_{ik} and b_k represent the consequent parameters ($i = 1 \cdots n, k = 1 \cdots M$), y_k is the output variable and $\tilde{A}_{ik}$'s are T2FS.

After obtaining the rules (either from expert or from numerical data), the rules are combined using appropriate $t - norm$ to obtain the firing strength of each rule. The final output of a T2FLS-TSK is computed as follows [1]:

$$y = [y_L, y_R] = \int_{f_1 \in [\underline{f_1}, \bar{f_1}]} \cdots \int_{f_M \in [\underline{f_M}, \bar{f_M}]} 1 \Big/ \frac{\sum_{k=1}^{M} f_k y_k}{\sum_{k=1}^{M} f_k} \tag{2.15}$$

where $\underline{f_k}$ and $\bar{f_k}$ are the lower and upper firing strength of the rules which are computed as:

$$\underline{f_k}(x) = \underline{\mu}_{\tilde{A}_{1k}}(x_1) * \underline{\mu}_{\tilde{A}_{2k}}(x_2) * \cdots * \underline{\mu}_{\tilde{A}_{nk}}(x_n) \tag{2.16}$$

$$\overline{f_k}(x) = \overline{\mu}_{\tilde{A}_{1k}}(x_1) * \overline{\mu}_{\tilde{A}_{2k}}(x_2) * \cdots * \overline{\mu}_{\tilde{A}_{nk}}(x_n) \tag{2.17}$$

where $\underline{\mu}_{\tilde{A}}(x)$ is the lower membership function and $\bar{\mu}_{\tilde{A}}(x)$ is the upper membership function of element $x \in X$.

2.4.4 Type Reduction

The outputs from the inference engine of an IT2FLS are IT2FSs. These outputs are converted into T1FSs, otherwise called type-reduced sets (TRSs), by the type reducer. According to Tai *et al.* [32], the type reduction (TR) procedure; which involves determining the centroid of *"an extraordinary large"* number of type-1 fuzzy sets [20], poses a computational bottleneck in computing the output of an IT2FLSs. According to Greenfield and Chiclana [93], the most widely used TR methods is the Karnik-Mendel (K-M) [12,94] iterative procedure which reduces a T2FS into a T1FS by computing the two end-points $[y_L, y_R]$. This iterative procedure is computationally intensive especially with large number of rules [95]. Interested readers are referred to [96] for a critical review of TR strategies.

2.4.5 Defuzzification

During the defuzzification process, the TRSs are sent to the defuzzifier in order to obtain a crisp output from the IT2FLS. The TRSs are formed by taking their left and right end-points, the defuzzified value is computed by taking the average of these two end points. The TR procedure leading to defuzzification is very complex and challenging [97]. Based on this premise, different methods have been formulated for TR of an IT2FLSs in order to by-pass this computationally intensive step. Alternative methods proposed in the literature for computing the outputs of IT2FLSs include those reported in [98–105]. Some of these alternative algorithms have closed-form representations and with these closed-form representations, analysis become much faster [106] and more convenient [107] than the K-M algorithm. Examples of the closed form algorithms include Nie-Tan (NT) [103], Wu-Tan [101], Begian-Melek-Mendel (BMM) [98] methods and those reported in Ulu *et al.* [104, 105]. The iterative K-M procedure and the Wu-Mendel algorithms entail the computations of the centroids. The output of the framework proposed in this thesis adopts the BMM approach.

The output of an IT2FLS based on BMM closed form representation is expressed as [98]:

$$y = (1-\beta)\frac{\sum_{k=1}^{M} \underline{f}_k y_k}{\sum_{k=1}^{M} \underline{f}_k} \quad + \quad \beta\frac{\sum_{k=1}^{M} \bar{f}_k y_k}{\sum_{k=1}^{M} \bar{f}_k} \tag{2.18}$$

It can be observed from Equation 2.18 that the output of the IT2FLS is a combination of the outputs of two T1FLSs consisting of the lower membership and upper membership functions, where β is an adjustable coefficient to weigh the output of the two T1 FLS. With the BMM TR and defuzzification method, it is not a requirement that the rules' output be sorted as is the case of the K-M iterative method. However, BMM approach requires that $\underline{y_k} = \bar{y_k} \equiv y_k$. It is worth mentioning that the BMM algorithm is an off-shoot of NT method that by-passes the computationally complicated TR procedure to directly compute the outputs of an IT2FLSs [32].

2.5 Uncertainty Modelling

Type-1 FSs have been used extensively for uncertainty modelling in the last few decades and have been applied in many applications with great success. Despite the widespread use of FS and its connotation of uncertainty, FS handles uncertainty about the meaning of words by using membership functions that are precise [12]; which is not necessarily realistic [32]. Real world applications are fraught with higher order uncertainties that make it difficult to determine the exact membership functions for the antecedent and consequent parts of a fuzzy set [31]. With these levels of uncertainties, it becomes inappropriate to use type-1 FS in certain applications. According to Hagras [34], using type-1 fuzzy sets can cause degradation in the FLS's performance, which can lead to poor control and inefficiency; and time wastage due to attempts to frequently redesign or tune the type-1 FL system in order to cope with the different uncertainties. Because uncertainty modelling cannot be properly accomplished with type-1 FSs, Zadeh [7] introduced the idea of type-2 FS (T2FS) which is characterised by membership functions that are themselves fuzzy and defined in the interval [0,1]. Mendel [12] stated that T2FLSs control the effects of the uncertainties associated with the meaning of words by *modelling the uncertainties* and concluded in [26] that IT2FLS is a scientifically correct model for modelling uncertainties associated with words. John and Coupland [108] pointed out that the use of IT2FLS is a step in the right direction to computing with words.

This implies the existence of uncertainty in determining the membership function values, and therefore, the introduction of the notion of the footprint of uncertainty (FOU) in IT2FS to model uncertainties that invariably exist in the rule base of the system [12]. According to Klir [109, pp. xiii] *"uncertainty is viewed as a manifestation of some information deficiency."* Mendel [12, pp. 66] quoting Klir and Folger [110] states:

> *" When dealing with real world problems, we can rarely avoid uncertainty. At the empirical level, uncertainty is an inseparable companion of almost any measurement, resulting from the combination of inevitable measurement errors and resolution limits of measuing instruments. At the cognitive level, it emerges from the vagueness and ambiguity inherent in natural language. At the social level, uncertainty has even strategic uses and it is often created and maintained by people for different purposes (privacy, secrecy, propriety)".*

In Mendel [12], three groups of uncertainties are identified namely fuzziness, strife and nonspecificity. Mendel pointed out that fuzziness is the uncertainty about the meanings of the words that are used in the definition of the rules in the rule base, strife is synonymous to the uncertainty about the rule consequent while nonspecificity is associated with uncertainty about the measurements that activate the FLS. Other kinds of uncertainties and ways of handling them are mentioned elsewhere. For example, in Mendel [86] two classes of uncertainties are identified namely random uncertainties which are mainly handled by probability theory and linguistic uncertainties which are fully handled by FS and its variants. According to Mendel, FS can successfully be used to handle both kinds of uncertainties. Mendel [12] outlined the different sources of uncertainty that can occur in a FLS. These include:

- Linguistic uncertainties. These may arise from different opinions of experts about words that are used to define the antecedents and consequent of the rule base as words mean different things to different people [12].
- Uncertainty about the rule consequent as different experts do not all agree about the consequent of a rule.
- Uncertainty about the measurements that activate the FLS. For example, sensors have uncertainties associated with the measurements.
- Uncertainty about the data that are used to tune the parameters of a FLS. This could arise as a result of noise in the training data.

All these uncertainties translate into uncertainties about FS membership functions [15]. Mendel *et al.* [1], re-echo that using a FS with precise membership function to model these forms of uncertainties can degrade the overall performance of the system. A generally maintained view is that a T2FS with a third dimension and additional degrees of freedom provided by the FOUs can directly model and handle these forms of uncertainties in most applications as their membership functions are uncertain. John and Coupland [108] provided an excellent historical perspective of T2FLSs and their role in uncertainty modelling. In order for a T2FLS to be successful in uncertainty modelling, different approaches have been adopted in the literature for the optimal adaptation of its design parameters.

2.6 IT2FLSs Design Methodology

Many methods have been proposed in the literature for the design of IT2FLSs. The design consists of the structure and parameter optimisations where intelligent methods are adopted to determine the optimal antecedent and consequent parameters through a process of learning and tuning. While learning does not involve predefined parameters for the optimisation of FLSs, tuning begins the optimisation with some predefined parameters and attempts to find the best set of parameters [111]. Quite often, in FLSs, the two terms are used interchangeably as there is no tuning without a learning capability in a FLS. This research only considers parameter optimisation or parameter tuning. The parameters of a FLS consist of the antecedent and the consequent parameters. Whilst the input space is partitioned into different fuzzy regions in the antecedent parts, the behaviours of the system in those regions are described in the consequent parts.

Methods employed for the parameter optimisation of FLSs are often drawn from derivative-based otherwise known as gradient descent (GD) methods (algorithmic optimisation methods), non-derivative-based (heuristics methods) and hybrid approaches [111]. In Hassan *et al.* [111], the authors listed the derivative-based methods to include such algorithms as back-propagation algorithms, least square method, radial basis function, Levenberg−Marquardt algorithm, Kalman filter-based algorithms and simplex method while the derivative-free methods include genetic algorithms, simulated annealing, particle swarm optimisation, artificial bee colonies, ant colony optimisation and sliding mode theory. The hybrid approaches are either combinations of the derivative-based, or derivative and derivative-free approaches. Specific discussion of these varieties of algorithms is beyond the scope of this research, rather this research focusses on two well known derivative-based methods for FLSs parameter update namely, the GD backpropagation algorithm and the extended Kalman filter (EKF)-based method namely, the decoupled EKF (DEKF).

2.6.1 The Gradient Descent Methods

Different algorithms have been reported in the literature for the design of IT2FLSs. However, gradient-based methods (iterative optimisation algorithms) are probably the most widely used methods for the optimisation of the parameters of fuzzy logic systems [2, 12, 112–117]. The most popular GD learning algorithm is the back-propagation methods [118], where the first derivatives of the cost function is computed with respect to the design parameters. The back-propagation GDs consist of learning iterations where a single iteration is called an epoch. The back-propagation learning algorithm consists of two passes namely:

- The feed-forward pass where the external inputs are transmitted forward and the outputs of each training vector is computed. During this phase, the parameters of the consequent parts of the IF-THEN rules are updated.

- The backward propagation then propagates the difference between the model output and the actual output backward towards the input. During this phase, the antecedent parameters of the FLS are updated.

According to Wang [119], the GD method is simple, easy to use, and with a fast repetition each time of the iteration. Wang further pointed out that the GD algorithm is guaranteed to find the local minimum through numerous times of iterations as long as it exists. Generally, the GD method arises when an algorithm follows the negative of the gradient of the function to reach its minimum. The GD start at a point, for instance, θ_i and compute the gradient at that point $\nabla_{\theta_i} f(\theta)$ and then takes a step, γ, in the direction of the negative of the gradient to find a new point θ_{i+1}, where every θ_{i+1} is computed as $\theta_{i+1} = \theta_i$ - $\gamma \nabla_{\theta_i} f(\theta)$. The gradient is computed at this new point and another step, γ, is taken in the direction of the negative of the gradient to obtain a new point θ_{i+2}. The algorithm proceeds until a minimum (local or global) is reached. The update rule for the generic parameter θ using GD is as expressed in Equation 2.19:

$$\theta_{i+1} = \theta_i - \gamma \frac{\partial E}{\partial \theta_i} \tag{2.19}$$

The GD method has been used over the years to update either or both the antecedent or/and consequent parts of a FLS. For example, Wang *et al.* [120] proposed a dynamical design of IT2FLS through a combination of an artificial neural network (ANN) and IT2FLS for handling uncertainties. The GD is adopted to tune the antecedent and consequent parts of the rules while a genetic algorithm is utilised to determine the optimal spread and learning of the designed system.

Khanesar *et al.* [4] proposed a T2FLS using a novel elliptic type-2 membership function in order to investigate the noise reduction property of a T2FLS. The authors in [4] utilised GD approach to tune both the antecedent and consequent parameters of the proposed type-2 membership function. Lin *et al.* [95] proposed a simplified TSK-type IT2FNN with online structure and parameter learning. These authors used the GD algorithm to adapt the parameters of the proposed model.

Juang and Tsao [121] proposed a self evolving interval type-2 fuzzy neural network (SEIT2FNN) with online structure and parameter learning with GD method used to update the parameters in the antecedent parts and rule-ordered Kalman filter to adjust the consequent parts. Lin *et al.* [122] proposed a TSK-based self-evolving compensatory interval type-2 fuzzy neural network (TSCIT2FNN) for minimising the effect of uncertainty in the rule base of a FLS. Their proposed system adopted two derivative-based methods (a first-order and second-order derivative methods). The antecedent parameters of the rules are tuned using GD algorithm while the consequent parameters are tuned using variable expansive Kalman filter approach. The designed system utilised A2-C0 TSK-design model and is applied to system modelling and noise cancellation. Castillo *et al.* [123] and Hassan *et al.* [111] provided excellent surveys of the different approaches for optimal design of an IT2FLSs. Detailed parameter update rule based on GD and applications using the model proposed in this thesis are presented in Chapter 4.

2.6.2 The Kalman Filter-based Methods

Despite the extensive use of the GD method for fuzzy systems' parameter tuning, it still suffers some drawbacks. Gradient descent being a first-order derivative-based method has the disadvantage of slow convergence and the possibility of getting stuck in local minima, leading to poor solutions [124]. To tackle this problems, second-order GD-based methods have been adopted in the literature for the adaptation of the parameters of FLSs among them is the Kalman filter (KF)-based approaches [115, 125], as they can converge in few iterations and are less likely to get trapped in local minima [125].

For instance, Juang *et al.* [126] investigated some dynamic system identifications and chaotic signal predictions under both noise-free and noisy conditions using a recurrent self-evolving IT2FNN (RSEIT2FNN). The authors utilised a rule-ordered KF algorithm to tune the consequent part parameters and GD to tune the antecedent parts. Lin *et al.* [122] adopted a variable expansive Kalman filter approach to tune the consequent parameters of their proposed TSCIT2FNN.

However, the basic KF works well for linear dynamic systems with white process and measurement noise but real world and problems are non-linear. Hence, for nonlinear systems, the KF is extended (Extended Kalman Filter- (EKF)) through a process of linearisation where the nonlinear function is linearised around the current parameter estimates.

The EKF has been used to learn the parameters of some traditional FLSs with great success. For instance, Simon [115] used EKF to optimise the parameters of a fuzzy system and demonstrate the effectiveness of the approach using a motor current estimator. The results of evaluation was compared with the optimisation of fuzzy systems using GD approach and adaptive neuro-fuzzy inference system (ANFIS). It is interesting to note that, with the performance of KF-based method in Simon [115], the author concluded that the use of KF-approach for the optimisation of FLS should be given serious consideration. In the same vein, Slim [127] investigated the prediction and estimation of non-linear dynamic system using a neuro-fuzzy system trained with EKF algorithm. The developed approach is evaluated using a Mackey-Glass benchmark problem and a financial time series. Analysis of results show that neuro-fuzzy approach trained with EKF compared favourably with classical ANN trained with back-propagation and ANFIS scheme. The EKF has also been used to update the parameters of intuitionistic fuzzy systems of type-1. For instance, Yihong [128] applied adaptive IF neural network for air defence situation and threat assessment in battle grounds. Although no comparisons are made with other models, the simulated results reveal creditability enhancement of threat assessment and improved quality of assessment with precision.

However, because of the high dimensionality of the fuzzy system parameters, using the standard EKF can be more complicated [42,125] especially for larger problem domains. In order to alleviate this computational burden, the EKF is used in a decoupled form (DEKF) because it is faster and easier to implement [42] with the most useful properties of the EKF still preserved [129]. The DEKF algorithm has been used previously in Khanesar *et al.* [125] to train a T2FLS where the parameters of both the antecedent and consequent parts of the T2FLS are grouped into two separate vectors (antecedent and consequent parameter vectors). The proposed system in [125] is applied to different problem domains and comparison is made with type-1 FLS trained with DEKF and T2FLS trained with GD. The authors concluded that the T2FLS trained with DEKF outperforms type-1 FLS trained with DEKF and T2FLS trained with GD.

Two stages are involved in the parameter update using the DEKF namely: the time update and the measurement update. During the time update, the current state is projected forward in time in order to obtain a prior estimate that is used for the next step. During the measurement update, a new measurement is propagated in order to obtain the posteriori estimate. In using the DEKF to learn the parameters of a FLS, the antecedent and consequent parameters are grouped into two separate vectors - one for the antecedent and the other for the consequent parameters. The generic parameter update rule in the i_{th} group is as expressed in Equation (2.20) to (2.22):

$$\theta_t^i = \theta_{t-1}^i + K_t^i[y_t - h(\theta_{t-1})] \tag{2.20}$$

$$K_t^i = P_t^i H_t^i[(H_t^i)^T P_t^i H_t^i + R^i]^{-1} \tag{2.21}$$

$$P_{t+1}^i = P_t^i - K_t^i P_t^i (H_t^i)^T + Q^i \tag{2.22}$$

where K is the Kalman gain, P is the covariance matrix of the state estimation error, R is the measurement noise covariance and Q is the covariance of process noise. Detailed parameter update rule using DEKF for the model proposed in this thesis and applications are presented in Chapter 5.

The methods adopted in this thesis to optimise the parameters of the proposed model are derivative-based methods only namely: a first-order back-propagation GD method and a second-order extended Kalman filter (EKF)-based method, the so-called DEKF and their hybrid - DEKF and GD (see Chapter 6).

2.7 Application of IT2FLSs to Uncertainty Modelling

The IT2FLSs have proven effective in many practical applications. Thanks to the simplification of the computations of IT2FSs provided by Mendel *et al.* [24], many people can now implement T2FLSs on a far greater scale. Some excellent reviews of IT2FLSs and their applications can be found in [27, 28, 30, 32]. In this section, a review of applications of IT2FLSs for uncertainty modelling in some problem domains is presented.

2.7.1 Application to Classification and Prediction Problems

In Najafi *et al.* [130], a new method for the automatic classification of celiac disease using IT2FLS is presented. Fuzzy C-mean clustering is applied for the determination of membership functions. The model is evaluated using a dataset from Poursina Hakim Research Institute. Analysis is carried out using IT2FLS with fuzzy C-mean, IT2FLS without fuzzy C-mean and type-1 FLS. Results reveal that both IT2FLSs outperformed the type-1 FLS.

Due to insufficient reliability and robustness in brain-computer interface technology, the practical use of brain-computer interface is limited. The main problem being the extensive variability and inconsistency of brain signal patterns. To cope with this problem, Pawel *et al.* [131] presented a new T2FLS classifier within the framework of an electroencephalogram-based brain-computer-interface. Evaluation of results demonstrate the superior performance of T2FLS over conventional brain-computer interface approaches such as linear discriminant analysis and support vector machine in terms of maximum classification accuracy and information transfer. Study shows that different support vector machines may produce different hyperplanes for the same sample [132]. Using these hyperplanes for decision making often lead to different conclusions. To circumvent this problem, Zarei *et al.* [132] proposed an IT2 fuzzy fusion model where different support vector machines are combined in an ensemble for classification problems. Simulation results demonstrate that IT2 fuzzy fusion model generates the best classifications compared to other models such as ANFIS, type-1 FLS and single support vector machine. The authors re-echo that the best performance of IT2 fuzzy fusion model is due to the model's ability to overcome the uncertainties in the rule-base and the shape of the membership function.

Yao *et al.* [133] proposed a novel approach for human behaviour recognition and summarisation based on IT2FL classification system. The parameters of the proposed model are optimised using big-bang big-crunch evolutionary algorithm. The big-bang big-crunch-based IT2FLS with fuzzy classification technique is shown to outperform classical IT2FLS.

David Enke *et al.* [134] presented a three stage stock market prediction involving differential evolution-based T2 fuzzy clustering and fuzzy type-2 neural network (FT2NN). The differential evolution is an optimisation technique in evolutionary computation. The differential evolution-based fuzzy-type clustering method generated the type-2 fuzzy IF-THEN rules. The authors pointed out that the difficulty associated with the choice of the parameter, "m" in standard FCM is removed with the use of IT2FCM. It is shown that the use of IT2FCM leads to better location of the cluster centers with better fuzzy rule model. For training the model, FT2NN was employed. The proposed model was used to forecast stock prices. Analysis of results reveals better stock price prediction accuracies using T2-fuzzy approach compared to fuzzy type-1 approaches.

Sumati *et al.* [135] investigated the application of IT2 subsethood-based neural FIS (IT2SUNFIS) for pattern classification of iris flower. The proposed model is compared with existing approaches using number of parameters, number of rules and re-substitution accuracy. Results revealed that IT2SUNFIS performs better than other comparative models.

Bernado *et al.* [136] developed a genetic T2FLS for modelling and predicting financial applications. The authors opined that the proposed framework is able to generate a summarised optimised T2FLSs financial models that are easy to read and analyse by any user of the model for financial predictions. The proposed system was evaluated using two financial domains namely: to predict the good or bad customers in a credit card approval domain and to predict arbitrage opportunities in the stock market. Experimental evaluation shows that the IT2FLS-based genetic approach outperformed other financial models such as evolving decision rule with comparable performance to ANN.

Recently, there has been an increased interest in sport videos, and intelligent methods need to be developed to automatically classify sport videos for easy analysis and understanding by experts as well as providing entertainment opportunities. Because of the complicated and dynamic nature of video sequences, classification of these videos becomes difficult due to inherent uncertainties in the images. In order to address these uncertainties, Song and Hagras [137] proposed an IT2FL classification system (IT2FLCS) for sport videos. The parameters of the proposed model are tuned using big-bang big-crunch algorithm. The proposed IT2FLCS is evaluated using soccer video with three classes and found to outperform its type-1 version and back propagation neural network.

Load forecasting is one of the important aspects in the energy sector needed for efficient management and operations of the energy system. However, the process of load forecasting is very complex and challenging due to nonlinear and random characteristics of the load demands. In order to cope with these challenges, Khosravi *et al.* [138] presented an IT2FLS for short term load forecasting. The constructed system implement the TSK fuzzy inferencing using Gaussian membership function with fixed mean and uncertain standard deviation. The authors adopted the A2-C0 and A2-C1 TSK-models trained with genetic algorithm. To test the viability of their proposed model, the authors also implemented an ANN and T1 TSK FLS on the short term load forecasting problem. The conclusions drawn show that IT2FLS outperforms both the type-1 FL and ANN on the short term load forecasting problem.

2.7.2 Application to Pattern Recognition Problems

One of the popular research areas in computer vision is face recognition and many approaches have been adopted for face recognition analysis. In Mendoza *et al.* [139], a face recognition application is proposed based on IT2FLS and modular neural networks. Two IT2FLSs are adopted for the construction of the overall model. The first IT2FLS extracts useful features from the training samples while the second IT2FLS rates the relevance of each module in the network. The Sugeno fuzzy integral is used for response integration in the integration module of the modular network and IT2 fuzzy system to rank the relevance of each module. The authors concluded that IT2FLSs improved the overall performance results in image recognition.

Melin [140] proposed IT2FL for image processing and pattern recognition. The author applied a new T2FL approach for image edge detection and the model is compared with three traditional approaches namely Sobel operators, edge detection by gradient magnitude and detection with type-1 fuzzy logic. The IT2FLS was found to outperform these three traditional approaches.

Recently, Yadav and Vishwakarma [141] proposed an improved approach based on IT2FL-based information extraction for face recognition systems. The purpose of the proposed system according to the authors is to minimize the effect of uncertainty in face recognition systems arising from variations in light direction, facial expression, etc. The model is evaluated using data from American Telephone & Telegraph (AT&T) face database. The authors claim that the sensitivity variations between images is reduced with IT2 membership functions.

2.7.3 Application to Clustering Problems

Qin *et al.* [142] proposed the clustering of sea surface temperature using IT2 fuzzy C-mean (IT2FCM) in order to discover spatial temporal patterns for enhanced climate change. The authors pointed out that due to the level of uncertainty in sea surface temperature, the use of standard FCM approach becomes inappropriate as it does not take into consideration the uncertainty in membership grades, hence the use of IT2 variant of FCM. According to the authors, IT2FCM achieved improved performance compared to standard FCM.

Yu *et al.* [143] presented a fuzzy clustering approach called interval type-2 possibilistic C-mean (IT2PCM) with alternating cluster estimation. The authors suggested that with the proposed approach, users are able to construct IT2 fuzzy membership functions with the flexibility of building cluster prototypes. The authors also claimed that the proposed approach is robust to inliers and outliers.

Wireless sensor networks have been applied in the monitoring of surrounding environments and communication of information to disparate base stations. The challenge in wireless sensor networks is the network lifetime that must be prolonged while ensuring precision. In order to maintain this balance, Cuevas-Martinez *et al.* [144] proposed a new fully distributed IT2FL controller for clustering in wireless sensor networks. The authors claimed that their proposed method significantly improve the whole network lifetime without incurring any central computation or complex procedures in the network nodes.

The design of IT2FCM-based NN is proposed in Kim *et al.* [145]. The hidden layer of the proposed model utilised IT2FCM clustering to handle uncertainty in the input space.

The connection weights of the proposed architecture are adjusted using local least square estimation-based learning. The authors pointed out that with the application of IT2FCM in the hidden layer, the proposed model is able to efficiently handle uncertainty in the input space better than the type-1 FCM.

2.7.4 Application to Control Problems

In Wati [146], a multi-input-multi-output IT2FL controller (IT2FLC) is designed for the automatic control of bath system temperature and water flow rate. The input to the shower system are hot water and cold water and the output is water at a certain temperature. Experimental analysis revealed an improved performance of IT2FLC over type-1 FLC in terms of fast step response of the output temperature and output flow rate of the shower systems.

Linda and Manic [147] designed an IT2FLS by incorporating into the model two novel quantifiers namely: the antecedent uncertainty quantifier and consequent uncertainty quantifier. The new proposed model was used in the design of a wall following navigation of a controller for autonomous mobile robot. Analysis of results shows that the proposed model provides accurate interpretation of uncertainty in the output of the IT2FLS.

In Ri *et al.* [148], the control of a mobile wheeled inverted pendulum is designed using the notion of IT2FL controller (IT2FLC) in order to model uncertainties and external disturbances. In particular, the study focused on the velocity, balancing and yaw steering controllers of mobile wheeled inverted pendulum. The proposed system is simulated under two conditions namely, with measurement uncertainties and external disturbances. In both cases, IT2FLC outperforms the T1FLC.

Bai and Wang [149] proposed a model-free approach for robot calibration based on IT2 fuzzy error interpolation method. In this way, the robot calibration does not undergo kinematic modelling and identification steps as opposed to the model-based approach. The proposed system is evaluated and compared with other interpolation techniques such as tri-linear, cubic spline and type-1 fuzzy error interpolation. Analysis of results revealed that the IT2 fuzzy interpolation outperforms the type-1 fuzzy error interpolation and other interpolation approaches.

IT2FS have also been used extensively in decision making. For instance, Cheng *et al.* [150] proposed a novel autocratic group decision making strategy using recommendation system by ranking IT2FSs. Other works employing IT2FSs in group decision making include [151–153]. A comprehensive survey of IT2FSs in decision making is provided in [29] and there are still more grounds to be explored [154].

2.8 Drawbacks of IT2FLSs

Despite the literature being replete with several works revolving around IT2FLSs that utilises IT2FSs in the rule base, it is worth noting that with IT2FSs, the variations of the uncertainties within the FOU of IT2FSs are not captured because the uncertainty is evenly spread across the FOU which practically leads to loss of some information [155] as compared to the GT2FSs. Another issue IT2FLSs, in particular, encounter is the curse of dimensionality [156], that is, the number of rules is exponentially proportional to the number of inputs and this increases the computational complexity of the system compared to the type-1 counterpart. Moreover, the IT2FSs strong assumption that non-membership functions are complementary to membership functions can make them unsuitable in some situations. IT2FLSs lack the capability of handling a situation with the characteristic of neither belonging nor not-belonging (indeterminate), which is a common phenomenon in natural language context. All the models so far discussed concentrate on using only the membership function in their fuzzy set definitions with an implicit assertion that non-membership functions are complementary to membership functions. Nevertheless, few models have attempted to address this problem by exploiting both membership and non-membership functions with hesitation indices otherwise known as Atanassov IFLSs utilising the Atanassov intuitionistic fuzzy sets (AIFSs).

2.9 Atanassov Intuitionistic Fuzzy Set

Because the classical FS non-membership function (ν) is complementary to the membership function (μ), that is, $\nu = 1 - \mu$ with no form of uncertainty whatsoever, Atanassov [8] extended the concept of Zadeh's fuzzy sets to intuitionistic fuzzy sets, hereafter referred to as AIFSs, which handle uncertainty by taking into account both the membership and non-membership degrees of an element x to a fuzzy set A together with extra degree of indeterminacy (hesitation).

An Atanassov intuitionistic fuzzy set (AIFS), characterized by a membership and non-membership functions, is a generalization of FS. Whereas the FS focuses on assigning mem-

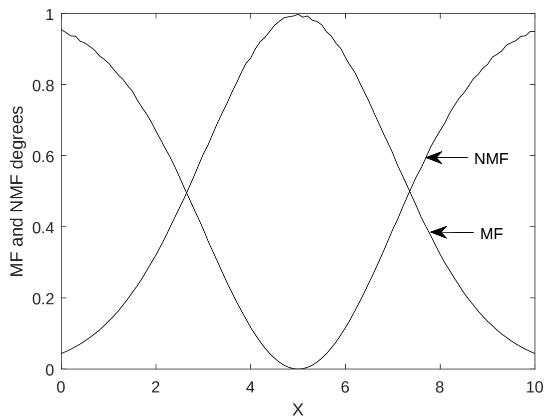

Figure 2.5: A-intuitionistic Gaussian membership and non-membership functions - AIFS

bership grades to elements in a UoD, AIFS assigns both membership and non-membership to each element of a set, and allows explicit representation of non-belongingness.

According to Ejegwa *et al.* [157], the degree of non-membership of an element in a fuzzy set may not always be 1 minus the degree of membership, that is, ($v(x) \neq 1 - \mu(x)$), an assertion implicit in classical FS, because there may be some degree of hesitation of that element to the set. Thus the semantic representation of AIFS, A^* includes the degree of membership, degree of non-membership and the hesitation margin $\{(\mu_{A^*}(x), \nu_{A^*}(x), \pi_{A^*}(x)) \mid x \in X\}$ respectively. Given the background of AIFS, it can formally be defined as follows:

2.9.1 Type-1 AIFS: Definition

Definition 2.9.1 Given a finite, non-empty set X, an AIFS A^* in X is an object having the form: $A^* = \{(x, \mu_{A^*}(x), \nu_{A^*}(x)) : x \in X)\}$, where the function $\mu_{A^*}(x) : X \to [0,1]$ defines the degree of membership and $\nu_{A^*}(x) : X \to [0,1]$ defines the degree of non-membership of element $x \in X$ and for every element $x \in X$, $0 \leq \mu_{A^*}(x) + \nu_{A^*}(x) \leq 1$ [8].

When $\nu_{A^*}(x) = 1 - \mu_{A^*}(x)$ for every $x \in X$, then the AIFS A^* collapses to ordinary fuzzy set A. Thus, given an AIFS, the degree of hesitancy of x to A^* is given by:

$$\pi_{A^*}(x) = 1 - (\mu_{A^*}(x) + \nu_{A^*}(x)).$$

This is called the A-intuitionistic fuzzy (IF) index of x in A^*. Barrenechea *et al.* [158] pointed out that the IF-index is an important attribute of AIFS as valuable information of each element can be obtained. The authors also noted that the IF-index plays very important role in algorithms performance.

For ordinary fuzzy set A, $\pi_A(x) = 0 \ \forall x \in X$.

Given an instance of a FS:

$$A = \{(x, \mu_A(x)) \mid \ \forall x \in X\} \tag{2.23}$$

The FS A can be represented as AIFS:

$$A = \{(x, \mu_A(x), 1 - \mu_A(x)) \mid \ \forall x \in X\} \tag{2.24}$$

where $1 - \mu_A(x)$ represent the non-membership function of a FS.

Conversely, given an AIFS:

$$A^* = \{(x, \mu_{A^*}(x), \nu_{A^*}(x)) \mid \ \forall x \in X\} \tag{2.25}$$

If all the elements of the AIFS satisfy the condition:
$\mu_{A^*}(x) + \nu_{A^*}(x) = 1$, a classical FS is recovered. Then A^* can also be expressed as:

$$A^* = \{(x, \mu_{A^*}(x), 1 - \mu_{A^*}(x)) \mid \ \forall x \in X\} \tag{2.26}$$

A FLS that utilises AIFSs in their rule base is referred to as Atanassov intuitionistic fuzzy logic system (AIFLS).

2.9.2 Atanassov Intuitionistic Fuzzy Logic Systems

The AIFLS consists of four basic modules similar to classical type-1 FLS. However for AIFLS, these modules utilise intuitionistic fuzzy sets and are therefore referred to as the intuitionistic fuzzifier, the intuitionistic fuzzy rulebase, the intuitionistic inference engine and the intuitionistic defuzzifier [128]. The intuitionistic fuzzifier maps the external crisp inputs into AIFS for which the corresponding membership and non-membership degrees are obtained. The rules may be constructed using experts knowledge or from numerical data. The inference engine combines the rule using a $t - norm$ to produce a type-1 IFS which are then defuzzified to obtain the final crisp output.

Apart from the T2FSs, one of the the most accepted generalisations of fuzzy sets is the Atanassov IFSs [76] as it has received greater attention since its appearance [159] and increasingly new concepts are linked to the notion of AIFSs [160]. According to [161–163], AIFSs are found to be useful for dealing with vagueness. With AIFS, the fuzzy characteristic of “neither this nor that” (indeterminate state) can be effectively described, thus

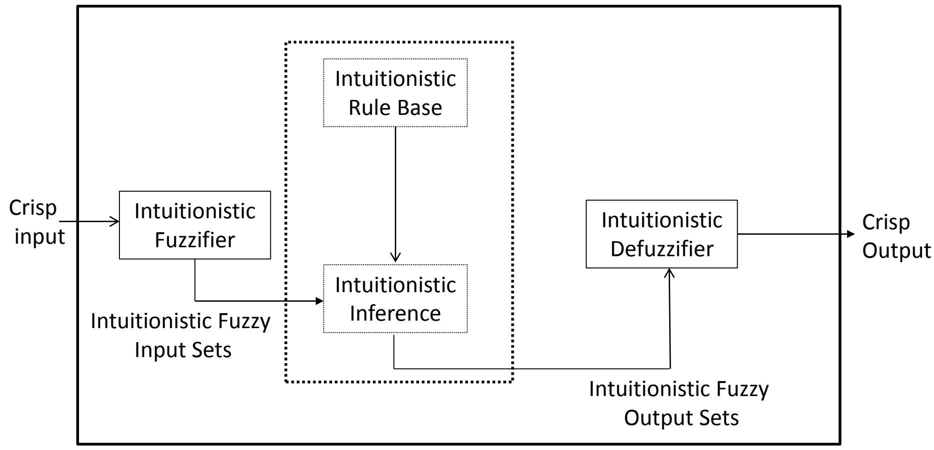

Figure 2.6: Structure of AIFLS

providing AIFS the flexibility and the ability to capture more information than FS [39]. Szmidt and Kacprzyk [164] stated that AIFSs are useful in problem domains where the use of linguistic variable to describe the problem in terms of membership functions only seems too restrictive. According to Olej and Hajek [165], the representation of attributes by means of membership and non-membership functions provides a better way to express uncertainty. Castillo *et al.* [41] pointed out that the non-membership degrees and intuitionistic fuzzy indices enable the representation of imperfect knowledge and also allow adequate description of many real world problems. In particular, Cornelis *et al.* [166] pointed out that the real essence of AIFS is that often when people are assessing a degree, they are often reluctant to pinpoint the degree decisively, which could mean strong commitment because to some extent they are hesitant about such assessments. People prefer to fix a certain threshold to indicate the positive and negative evidence with some hesitation. According to Shinoj and Sunil [167], AIFS can be used as a more appropriate tool in this context for simultaneously representing both membership and non-membership of an element to a set and not insisting that the assessment be complementary [38]. The AIFS-based models are appropriate in a wide variety of situations in which human opinions are elicited. A typical example is the scenario painted in Chapter 1 about people's perception of Zadeh's seminal paper on fuzzy set. Another good example is a voting scenario given in [38, 39], for example, there are people who will vote for, vote against, abstain or cast invalid votes [41]. In this voting contexts, the description by a linguistic variable using the membership function only is inadequate. This voting example will be recalled whenever it becomes relevant.

According to Own [81], when dealing with the problem of vagueness where there is insufficient information leading to an inability to satisfactorily specify the membership function, the AIFS theory becomes more suitable than fuzzy sets to deal with such problems. In Szmidt and Kacprzyk [168], it is argued that AIFS is a tool for a more human consistent reasoning under imperfectly defined facts and imprecise knowledge.

2.10 Practical Applications of AIFSs

Studies involving AIFSs have drawn much attention in recent times and have been successfully applied in different problem domains. In medical domain, for instance, Chaira [169] proposed a novel global medical image thresholding approach using AIFS. Due to uncertainties in real time medical images with noisy, vague and indistinguishable characteristics, conventional approaches become inadequate. In particular, Chaira pointed out that thresholding approach with two uncertainty specifications of membership and non-membership helps to appropriately segment the image for thresholding. The author adopted the Sugeno-type intuitionistic fuzzy generator for computing the non-membership values of the medical images. It is concluded that the thresholded images with the proposed IF approach provides better medical images than the conventional fuzzy approaches employing fuzzy cluster-based thresholding, measures of fuzziness and fuzzy compactness.

Khatibi and Montazer [170] proposed a medical classification of bacteria using five similarity measures. The method of IFSs are compared with classical FSs to examine their capabilities in handling uncertainty in the medical pattern recognition. The authors noted that utilizing the AIFS framework not only provides more accurate classification results, but also, the related errors generated by the AIFS are smaller than those of traditional FSs. Other works employing AIFS in medical domains are reported in [164, 171].

Hajek and Olej [172] proposed a TSK-based intuitionistic fuzzy neural network (IFNN) with application to credit scoring using text information. The proposed IFNN is trained using two approaches namely GD and KF. The authors pointed out that the introduction of non-membership function into a fuzzy inference system improve the performance of the system. Particularly, the proposed methods of IFNN-GD and IFNN-KF are compared with ANFIS and found to significantly outperform ANFIS. Hajek and Olej concluded that KF approach is one of the viable approaches for the adaptations of the consequent parameters of IFNN. In a different application context, Yihong *et al.* [128], demonstrated how IFNN based on TSK-inference can be applied to threat assessment in battle grounds. The authors re-echo that the use of AIFS is more appropriate for analysing the situation and threat assessment information than the traditional FS. The proposed method employs a neural network learning based on EKF. Results show improvement of assessment quality with enhanced threat assessment creditability.

In Hajek and Olej [173], an adaptive IF- inference system (IFIS) based on TSK fuzzy inferencing for regression problems is presented. Several optimisation approaches such as subtractive clustering algorithm, Moore-Penrose pseudo-inverse, KF, Kaczmarz algorithm and gradient descent are utilised to tune the parameters of the IFIS. The performance evaluations in terms of the root mean squared error (RMSE) show that IFIS outperforms the classical FIS in all the problem instances. In Olej [165, 174], a novel TSK based IFIS for time series prediction is also proposed.

Castillo *et al.* [41] presented an intuitionistic fuzzy system for time series analysis in plant monitoring and diagnosis. The output of the proposed intuitionistic fuzzy system is a combination of two traditional fuzzy systems. The authors in [41] used their proposed approach for plant monitoring and claimed that intuitionistic fuzzy logic has the potential of modelling uncertainty in a dynamic process. The authors concluded that the new method of fuzzy inferencing with intuitionistic fuzzy systems can be applied to control problems and time series predictions. Olej and Hajek [165] presented a TSK-type intuitionistic fuzzy inference systems for ozone time series prediction. The authors opined that the use of AIFS *"present a strong possibility to express uncertainty".* Intarapaiboon [175] applies AIFS to text classification using similarity measures. Szmidt and Kacprzyk [176] presented IFSs as an efficient and effective tool for feature selection in text categorisation.

Other application domains using AIFS worth mentioning include: control [177, 178], bankcruptcy forecasting [179], decision making [162, 180–182] and e-learning to evaluate student knowledge of Mathematics in university courses [183]. As more number of neurons tend to slow down the learning process of a modular neural network, Sotirov *et al.* [184], proposed an intuitionistic fuzzy intercriteria analysis approach for reducing the number of neurons/parameters in a modular neural network thereby speeding up the learning process.

2.11 Drawbacks of AIFSs

These studies adopting AIFS have focussed on type-1 AIFSs and AIFLSs. AIFSs and AIFLSs have their limitations. With uncertainties arising from different sources, it becomes more appropriate to map such uncertainties into membership and non-membership function uncertainties. Using AIFS with single membership and non-membership functions in handling such uncertainties is not realistic, as the determination of the exact membership and non-membership functions is difficult to pinpoint. Hence, similar to the notion of a classical T1FS, the type-1 AIFS may not handle or minimize the plethora of uncertainties that are inherent in many real world applications as their membership and non-membership degrees are exactly defined. For AIFS, the uncertainty disappears once the membership and non-membership parameters are specified.

To tackle this problem, Atanassov and Gargov [73] extended the concept of AIFS to interval valued AIFSs (IVAIFS) which is a generalization of the notion of AIFS in the sense of IVFS (a special case of IT2FS). The IVAIFS are characterised by membership and non-membership functions that are intervals and defined in the referential set of $[0, 1]$. In Chapter 3 the semantic differences between IVAIFS and the new framework proposed in this thesis are highlighted.

2.12 Studies Involving Combination of AIFSs and IT2FSs

Few studies have been conducted involving the combination of AIFS and IT2FS. Naim and Hagras [185] argued that the combination of AIFS and T2FLS are well suited for handling imprecision and vagueness. Some research has shown interest in the arithmetic operations of T2AIFS. In Cuong *et al.* [186] some set theoretic operations for T2AIFS and their properties are discussed. The authors concluded that many applications will benefit from the use of such sets. Similarly, Jana [187] has proposed some novel arithmetic operations on GT2AIFS on the basis of (α, β)-cut methods with application to transportation problems. Recently, Singh and Garg [188] proposed some distance measures for T2AIFS and applied the proposed measures to multi-criteria decision making. A few FL structure have attempted to solve the problem of uncertainty modelling by exploiting both AIFS and IT2FS. An example application in this category (of the combination of AIFS and IT2FS) is found in the work of Nguyen *et al.* [189], where IT2 fuzzy C-mean (IT2FCM) and AIFS are applied for clustering of different types of images especially those corrupted with noise. Experimental results reveal improvement in the clustering quality of images using IT2FCM and AIFS compared to representative algorithms like FCM and IT2FCM.

The combination of AIFS and IT2FS have also been applied to image thresholding. For instance, Nghiem *et al.* [190] applied intuitionistic T2FS to image thresholding using Sugeno intuitionistic fuzzy generator. The authors claim that their proposed method exhibits higher thresholding quality with noisy images compared to typical algorithms such as image segmentation using type-1 fuzzy set and AIFS alone.

In the literature, numerous approaches based on type-1 AIFS for decision-making have also been proposed. However, few examples are recorded in the literature involving the combined approach of AIFS and IT2FL. For instance, Naim and Hagras [185], presented a hybrid approach where IT2 and AIFS are utilised in multi-criteria group decision making (MCGDM). The proposed system employs IT2FS to handle the linguistic uncertainty while utilising intuitionistic evaluation in the design of the non-membership function degrees. The authors applied the proposed method to the evaluation of postgraduate study involving ten candidates. Analysis of results shows that variations in the group decision making using the proposed method of IT2FS and IF evaluation provided better agreement with the human experts decision than AIFS, FS and IT2 fuzzy systems.

In Naim *et al.* [191], fuzzy logic-MCGDM (FL-MCGDM) is proposed for selecting appropriate and convenient lighting level for reading to meet each individual needs as this varies among users. The proposed hybrid system was developed using the concepts of IT2FS and the hesitation indices provided by the AIFS. The membership function of the IT2FS for the left and right end-points were represented in intuitionistic values. Experimental evaluation revealed a significant correlation between the user's linguistic appraisal and the result provided by the proposed FL-MCGDM system. The authors concluded that the combination of T2FS and AIFS provides FL-MCGDM with enhanced capability for decision making. Another FL-MCGDM is proposed in Naim and Hagras [35] for intelligent shared environment. The proposed model also utilised IT2FS and hesitation indices of AIFS in the design of the decision making model. In order to evaluate the effectiveness of the designed approach, the authors applied the model to an intelligent apartment and concluded that the results are consistent with the human decision as compared to classical fuzzy MCGDM. In a study by Own [81], a switching between T2FSs and AIFSs is proposed. In [81], the switching relation between T2FSs and AIFSs is defined axiomatically. The advantages of T2FSs are exploited and the switching results are applied in pattern recognition and medical diagnosis reasoning to show the usefulness of the proposed method.

The research reported here adopts a similar idea, as discussed above, of using both AIFS and IT2FS in the design of the proposed framework. However, the motive and approach for the framework proposed in this thesis are quite distinct from those advanced in the above models. Among other things, no framework listed above obviously shows the benefit of explicitly using membership and non-membership functions that are intervals together with IF-indices for uncertainty modelling. Moreover, these existing models do not have any learning or parameter optimisation mechanism whatsoever. Hence, this work is an attempt in this direction to develop a framework that fuses both concepts and models uncertainty using separately defined membership and non-membership functions that are intervals with ANN learning capability. A more careful treatment of the proposed framework is provided in Chapter 3.

2.13 Summary

In this chapter, a survey of related works in uncertainty modelling is provided. In particular, the discussion focussed on approaches involving IT2FSs and AIFSs and their fusion. How these existing approaches address the issues of uncertainty modelling are discussed. In the context of minimising the effects of uncertainties in applications, existing relevant works separately adopting the notion of IT2FSs and AIFs and systems are reviewed. The major barriers to the effective application of IT2FSs and AIFSs and systems, which this thesis aim to investigate have been highlighted. Whilst the classical IT2FLSs have made significant waves in modelling large amounts of uncertainties, they are not able to manage indeterminate (hesitant) states well. For IT2FLSs, it is the assignments of only the membership grades (lower and upper) to every element in a UoD with no hesitation, thus enforcing duality that non-membership degrees are complementary to the membership degrees. AIFLSs, on the other hand, may not handle the amount of uncertainty inherent in many real world applications as their single membership and non-membership functions cannot incorporate information from diverse sources simulteneously. In other words, AIFLSs are useful for defining an uncertain term from a single point of view. Any change in perception of the same linguistic term will entail frequent re-tuning of the AIFS membership and non-membership functions so that it can deal with the various uncertainties. This may lead to a sub-optimal system performance under certain operation and environment conditions.

Although, different approaches for dealing with these challenges in terms of the combination of AIFS and IT2FS have been addressed, most of these approaches are less relevant to the problem domain investigated in this thesis. In addition, some of these approaches utilise only a single IT2FS and evaluate the hesitation on the primary membership function of the IT2FS. For instance, the works of [35,185,191] while effective in handling MCGDM, do not consider the specification of non-membership function as a separate region but rather IT2FS is employed with intuitionistic evaluation (hesitation) on the membership function FOUs. They do not explicitly apply membership and non-membership functions that are intervals. Moreover, no learning or optimisation whatsoever, has been carried out on these sets. The model presented in Chapter 3 provides a point of departure of the model proposed in this thesis from existing approaches in the literature. As mentioned earlier, the focus of this research is to systematically integrate Atanassov's notion of IFS into IT2FLS with the aim of modelling linguistic uncertainties using membership and non-membership functions that are intervals in [0,1] with the hesitation degrees defined for both membership and non-membership functions.

Chapter 3

Model Formulation

> As complexity rises, precise statements lose meaning and meaningful statements lose precision.
>
> Lotfi A. Zadeh

3.1 Introduction

In Chapter 2, some concepts which underpin the contributions of this research are introduced. The main purpose of this chapter is to introduce a new TSK-based interval type-2 Atanassov-intuitionistic fuzzy logic system (IT2AIFLS-TSK) that utilises fuzzy membership and non-membership functions together with intuitionistic fuzzy indices (IF-indices) for uncertainty modelling. The general framework is based on TSK-fuzzy inference. It is argued that the fuzzy non-membership and intuitionistic fuzzy indices can be incorporated into an IT2FLS in order to handle uncertainty well and mitigate their effects. As pointed out in Eyoh *et al.* [3], the integration of these two concepts can bring about a synergistic effect in uncertainty modelling with the capacity for improved system performance.

3.2 Generalised Type-2 A-Intuitionistic Fuzzy Set

Here, a new definition for a generalised T2AIFS (GT2AIFS) is provided, for the first time. A GT2AIFS $\tilde{A}^*$ in the universe of discourse, X consists of type-2 membership and non-membership grades of $x \in X$ defined as $\mu_{\tilde{A}^*}(x,u) : u \in J_x^{\mu} \subseteq [0,1]$ and $\nu_{\tilde{A}^*}(x,u) : u \in J_x^{\nu} \subseteq [0,1]$ respectively [2]. The primary membership (J_x^{μ}) and primary non-membership (J_x^{ν}) of element $x \in \tilde{A}^*$ are elements in the domain (x,u) which form supports of a GT2AIFS in the third dimension for membership and non-membership functions respectively and are defined as follows [2]:

$$J_x^{\mu} = \left\{(x,u) : u \in \left[\underline{\mu}_{\tilde{A}^*}(x), \overline{\mu}_{\tilde{A}^*}(x)\right]\right\}$$

$$J_x^{\nu} = \left\{(x,u) : u \in \left[\underline{\nu}_{\tilde{A}^*}(x), \overline{\nu}_{\tilde{A}^*}(x)\right]\right\}$$

Definition 3.2.1 A generalised T2AIFS denoted by $\tilde{A}^*$ is characterised by a type-2 membership function $\mu_{\tilde{A}^*}(x,u)$, and a type-2 non-membership function $\nu_{\tilde{A}^*}(x,u)$ [2], i.e.,

$$\tilde{A}^* = \{(x,u), \mu_{\tilde{A}^*}(x,u), \nu_{\tilde{A}^*}(x,u) \mid \forall x \in X, \forall u \in J_x^{\mu}, \ \forall u \in J_x^{\nu}\} \tag{3.1}$$

in which $0 \leq \mu_{\tilde{A}^*}(x,u) \leq 1$ and $0 \leq \nu_{\tilde{A}^*}(x,u) \leq 1$
where $\forall u \in J_x^{\mu}$ *and* $\forall u \in J_x^{\nu}$ conform to the T1 constraint that $0 \leq \mu_{A^*}(x) + \nu_{A^*}(x) \leq 1$.

That is, when uncertainties disappear, a T2 membership and non-membership functions must reduce to a T1 membership and non-membership functions respectively. Also the amplitudes of both membership and non-membership functions must lie in the closed interval of 0 and 1. That is, $0 \leq \mu_{\tilde{A}^*}(x,u) \leq 1$ and $0 \leq \nu_{\tilde{A}^*}(x,u) \leq 1$. Alternatively, a GT2AIFS, $\tilde{A}^*$, may be represented as [2]:

$$\tilde{A}^* = \int_{x \in X} \left[\int_{u \in J_x^{\mu}} \int_{u \in J_x^{\nu}} \left\{\mu_{\tilde{A}^*}(x,u), \nu_{\tilde{A}^*}(x,u)\right\}\right] / (x,u) \tag{3.2}$$

where $\int\int\int$ represents union over all admissible values of x and u for the membership and non-membership over a continuous UoD, and $\int$ is replaced by $\sum$ for discrete UoD. When the secondary membership functions $\mu_{\tilde{A}^*}(x,u) = 1$, and secondary non-membership functions $\nu_{\tilde{A}^*}(x,u) = 1$, a GT2AIFS translates to an interval type-2 Atanassov intuitionistic fuzzy set (IT2AIFS) (see Figure 3.1)

3.3 Interval Type-2 Atanassov Intuitionistic Fuzzy Set

Many real life problems involve dealing with multiple assessments. Returning to the voting scenario mentioned in Chapter 2, Section 2.9, where some people will:

- vote for
- vote against
- cast invalid vote or abstain from the poll

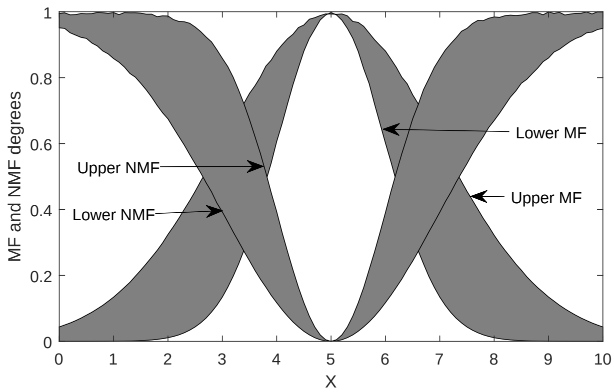

Figure 3.1: An IT2 A-intuitionistic Gaussian membership and non-membership functions - IT2AIFS [2]

Suppose further that the task is to classify the different classes of voters according to their ages. Employing a single membership and non-membership functions, while useful in many situations, might not be sufficient in this context. Ideally, people, when making any assessments are reluctant to decisively pin-point a single numerical value, be it membership or non-membership as doing so entails strong commitment [166] and no individual wants to be overly involved. Rather, people prefer to specify a certain range because they are hesitant to some degree about such assessment. Incorporating the notion of IF-indices into IT2FSs gives room for more flexibility in fuzzy set descriptions. This allows for more ease and ability to handle uncertainty and non-linearity. With the capacity to deal with uncertain membership and non-membership functions, IT2AIFSs allow for better modelling of real life situations than classical IT2FSs. What follows is a formal definition of IT2AIFS.

Definition 3.3.1 An IT2AIFS, $\tilde{A}^{*}$, is characterised by membership bounding functions and non-membership bounding functions defined as $\bar{\mu}_{\tilde{A}^{*}}(x)$, $\underline{\mu}_{\tilde{A}^{*}}(x)$ and $\bar{\nu}_{\tilde{A}^{*}}(x)$, $\underline{\nu}_{\tilde{A}^{*}}(x)$ respectively for all $x \in X$ with constraints: $0 \leq \overline{\mu}_{\tilde{A}^{*}}(x) + \underline{\nu}_{\tilde{A}^{*}}(x) \leq 1 \, and \; 0 \leq \underline{\mu}_{\tilde{A}^{*}}(x) + \overline{\nu}_{\tilde{A}^{*}}(x) \leq 1$ [189].

For instance, the interpretation of IT2AIFS for 'vote' is similar to the classical IT2FS expressed as:

$$vote = Medium/Day1 \ + \ High/Day2 \ + \ Low/Day3 \tag{3.3}$$

The difference is that the linguistic terms: $Medium$, $High$, and Low are now fuzzy-fuzzy (membership and non-membership) sets. That is, two membership type-1 AIFSs and two non-membership type-1 AIFSs. For each $x \in X$, there exist a third parameter $\pi(x)$ called the IF-index or hesitancy degree which comes as a result of an expert not being certain of the degree of membership and non-membership of element $x \in X$, and may cater to either membership, non-membership values or both. In the framework proposed in this thesis, the IF-index caters for both the membership and the non-membership functions of a set.

Two IF-indices used in this thesis are the IF-index of centre and IF-index of variance[1]. These indices were previously used in Hajek and Olej [172] and defined in this work as:

$$\pi_c(x) = max\left(0, \left(1 - \left(\mu_{\tilde{A}^*}(x) + \nu_{\tilde{A}^*}(x)\right)\right)\right)$$
$$\overline{\pi}_{var}(x) = max\left(0, \left(1 - \left(\overline{\mu}_{\tilde{A}^*}(x) + \underline{\nu}_{\tilde{A}^*}(x)\right)\right)\right)$$
$$\underline{\pi}_{var}(x) = max\left(0, \left(1 - \left(\underline{\mu}_{\tilde{A}^*}(x) + \overline{\nu}_{\tilde{A}^*}(x)\right)\right)\right)$$

such that: $0 \leq \pi_c(x) \leq 1$ *and* $0 \leq \pi_{var}(x) \leq 1$.

The IF-indices for this study are $m - by - n$ matrices randomly generated in the interval [0,1], where m is the number of linguistic terms and n is the number of inputs. These IF-indices are then incorporated into the FOUs of the IT2AIFS. The capability of taking the contribution of IF-index into account, aside from the non-membership degree, in the partitioning of the input space gives this approach an advantage over some conventional IT2 fuzzy approaches [3].

As defined above, an IT2AIFS $\tilde{A}^*$ is characterised by interval type-2 membership function, $\mu_{\tilde{A}^*}(x,u)$ and interval type-2 non-membership function, $\nu_{\tilde{A}^*}(x,u)$ for all $x \in X$ expressed as:

$$\begin{aligned}\tilde{A}^* &= \int_{x \epsilon X} \int_{u \epsilon J_x^\mu} \int_{u \epsilon J_x^\nu} 1/\left(x,u\right) \\ &= \int_{x \epsilon X} \left[\int_{u \epsilon J_x^\mu} \int_{u \epsilon J_x^\nu} 1/\left(u\right)\right]/x\end{aligned} \tag{3.4}$$

[1]Petr Hajek in an email conversation pointed out that *"IF-index of centre is used to express the hesitancy on the centre of the membership function while the IF-index of variance represents the hesitancy on the radius"* and these values are small numbers in the interval [0,1]

where x is the primary variable, and u is the secondary variable. The uncertainty about an IT2AIFS is completely described by the FOUs that are bounded by two T1 membership functions - an upper membership function given as $\bar{\mu}_{\tilde{A}^*}(x)$ and a lower membership function expressed as $\underline{\mu}_{\tilde{A}^*}(x)$ and two T1 non-membership functions which are - an upper non-membership function, $\bar{\nu}_{\tilde{A}^*}(x)$ and a lower non-membership function, $\underline{\nu}_{\tilde{A}^*}(x)$ as shown in Figure 3.1 and expressed as:

$$\begin{aligned} \bar{\mu}_{\tilde{A}}(x) &\equiv \overline{FOU_\mu(\tilde{A}^*)} \ \ \forall x \in X \\ \underline{\mu}_{\tilde{A}}(x) &\equiv \underline{FOU_\mu(\tilde{A}^*)} \ \ \forall x \in X \\ \bar{\nu}_{\tilde{A}}(x) &\equiv \overline{FOU_\nu(\tilde{A}^*)} \ \ \forall x \in X \\ \underline{\nu}_{\tilde{A}}(x) &\equiv \underline{FOU_\nu(\tilde{A}^*)} \ \ \forall x \in X \end{aligned} \tag{3.5}$$

Thus, two FOUs are defined for IT2AIFS namely: FOU_μ regarding the uncertainty of the membership function and FOU_ν defined with respect to the non-membership function of IT2AIFS $\tilde{A}^*$ (see Figure 3.1) as follows [3, 40, 46]:

$$FOU_\mu\left(\tilde{A}^*\right) = \bigcup_{\forall x \in X} \left[\underline{\mu}_{\tilde{A}^*}(x), \bar{\mu}_{\tilde{A}^*}(x)\right] \tag{3.6}$$

$$FOU_\nu\left(\tilde{A}^*\right) = \bigcup_{\forall x \in X} \left[\underline{\nu}_{\tilde{A}^*}(x), \bar{\nu}_{\tilde{A}^*}(x)\right] \tag{3.7}$$

The background on which the proposed fuzzy framework is based is provided. The different variants of fuzzy sets that have motivated this research are highlighted. However, because of the associated complexities of the GT2FS as discussed in Section 1.1 and Subsection 2.2.3, the IT2 Atanassov intuitionistic fuzzy logic framework is constructed based on the notion of IT2FS.

3.4 A Comparison Between Interval Valued Atanassov Intuitionistic Fuzzy Set and Interval Type-2 Atanassov Intuitionistic Fuzzy Set

In Subsection 2.3.1, the differences between IVFS and IT2FS [84] are discussed. similar to the discussion on IT2FS and its representations, it is argued that IT2AIFS can also be used in a more general perspective to represent concepts that are not possible with IVAIFSs namely as intuitionistic type-1 fuzzy sets, as interval-valued intuitionistic fuzzy sets, as intuitionistic-fuzzy multi sets and as multi-interval intuitionistic fuzzy sets, hence the adoption of IT2AIFS instead of IVAIFSs. Secondly, for IVAIFS, the general constraints is that the summation of the upper-bound membership and upper-bound non-membership degrees is less than or equal to 1. i.e.

$$0 \leq \bar{\mu}_{\tilde{A}} + \bar{\nu}_{\tilde{A}} \leq 1 \tag{3.8}$$

The point of departure is that, for IT2AIFS proposed here, the summation of the upper-bound membership and lower-bound non-membership is less than or equal to 1 and the summation of the lower-bound membership and upper-bound non-membership degrees is less than or equal to 1, i.e. for IT2AIFS, the constraints are [189]:

$$0 \leq \overline{\mu}_{\tilde{A}^*}(x) + \underline{\nu}_{\tilde{A}^*}(x) \leq 1 \tag{3.9}$$

and

$$0 \leq \underline{\mu}_{\tilde{A}^*}(x) + \overline{\nu}_{\tilde{A}^*}(x) \leq 1 \;\; \forall x \in X. \tag{3.10}$$

These two constraints present IT2AIFS as a new and distinct concept completely different from IVAIFS. It is useful to make these distinctions in the context of this research as it serves to distinguish the much broader concept of IT2AIFS from the more specific concept of IVAIFS. A FLS that utilises at least one IT2AIFS in the rule base is referred to as interval type-2 Atanassov intuitionistic fuzzy logic system (IT2AIFLS) - a new and sound framework proposed in this thesis.

3.5 TSK-based Interval Type-2 Atanassov-Intuitionistic Fuzzy Logic System Framework

The proposed IT2AIFLS-TSK takes the best of two worlds - AIFLS and IT2FLS. This way, IT2AIFLS-TSK:

- Assigns to each element of a set both membership and non-membership grades that are intervals.
- Enables hesitation and thus relaxes the complementarity assessments of classical IT2FS such that:

$$\tilde{A}^*{}_{\mu^c} = 1/FOU\left(\tilde{A}^*{}_{\mu^c}\right) \not\equiv 1/\left[1 - \bar{\mu}_A(x), 1 - \underline{\mu}_A(x)\right] \tag{3.11}$$

and

$$\tilde{A}^{*}{}_{\nu^c} = 1/FOU\left(\tilde{A}^{*}{}_{\nu^c}\right) \not\equiv 1/\left[1-\bar{\nu}_A(x), 1-\underline{\nu}_A(x)\right] \tag{3.12}$$

- Relaxes the inequality of IVAIFS in Equation 3.8 (See Equations 3.9 and 3.10)
- Incorporates more uncertainties (fuzziness) and captures more information, thus relaxes the single membership and non-membership functions of AIFSs.

The reference to TSK-type inference for IT2AIFLS is hereafter omitted for notational convenience. According to Hisdal [192, pp. 385], *"increased fuzziness in a description means increased ability to handle inexact information in a logically correct manner."* The structure of IT2AIFLS (see Figure 3.2) is similar to the AIFLS except in the defuzzification module. For IT2AIFLS, there is a type-reducer before the actual defuzzification. The type-reducer converts the IT2AIFS from the IF-inference engine into an AIFS. The type-reduced set (AIFS) is then defuzzified into crisp number as the final output. Another difference between AIFLS and IT2AIFLS, is that the fuzzy sets are IT2AIFS.

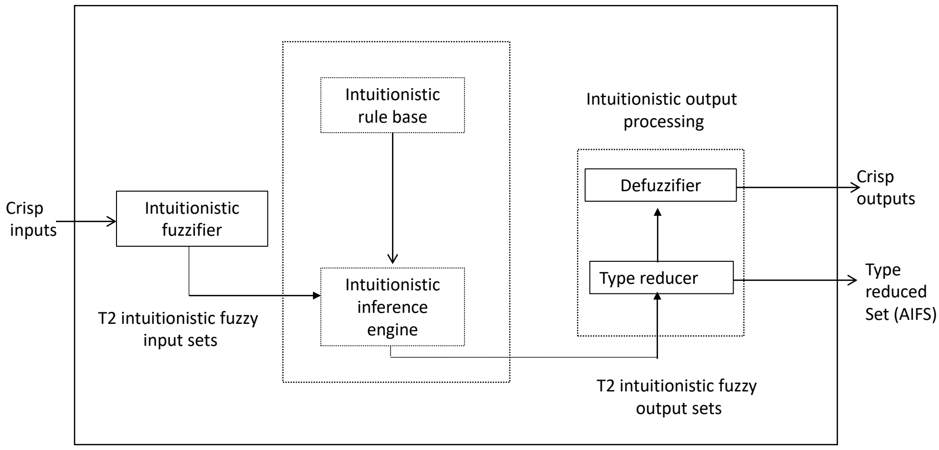

Figure 3.2: Type-2 A-Intuitionistic Fuzzy Logic System [3]

3.5.1 Fuzzification

There are two fuzzification procedures namely: singleton and non-singleton. In this thesis, the focus is on singleton fuzzification because it is faster to compute [154] and therefore suitable for the proposed model of IT2AIFLS. The fuzzification process involves the mapping of a numeric input vector $x \in X$ into an IT2AIFS $\tilde{A}^*$ in X which activates the inference engine. For each crisp input $x \in X$, interval type-2 A-intuitionistic fuzzy values for membership and non-membership are generated. Here, interval singleton type-2 fuzzification is used to obtain membership and non-membership values, because with this, obtaining a closed-form computation for the inference mechanism is possible [118].
For membership:

$$\mu_{\tilde{A}^*}(x) = \begin{cases} 1/1, & \text{if } x = x' \\ 1/0, & \text{if } x \neq x' \end{cases}$$

For non-membership:

$$\nu_{\tilde{A}^*}(x) = \begin{cases} 1/1, & \text{if } x = x' \\ 1/0, & \text{if } x \neq x' \end{cases}$$

The firing strength for membership and non-membership functions are intervals $[\underline{f^{\mu}}, \overline{f^{\mu}}]$ and $[\underline{f^{\nu}}, \overline{f^{\nu}}]$ respectively. A number of membership functions exists which are employed in the computation of type-2 fuzzy grades (fuzzification). These include triangular, trapezoidal, Gaussian, sigmoidal and others. In the literature, many applications benefit from the use of Gaussian functions for the design of FLSs [12, 16]. In this thesis, the Gaussian function is also adopted for the representation of both the membership and non-membership functions of the IT2AIFS, because according to Wu [193, pp. 7], *"Gaussian IT2 FLCs are simpler in design because they are easier to represent and optimize, always continuous, and faster for small rulebases."* Moreover, computing the derivatives of membership function parameters is straightforward with the Gaussian function especially in gradient descent-based optimisation algorithms [16] adopted in this research. For classical Gaussian IT2FLS, uncertainties can be associated to the standard deviation or mean of the fuzzy set. Mathematically, the classical Gaussian membership function is defined as follows:

$$\mu_{ik}(x_i) = exp\left(-\frac{(x_i - c_{ik})^2}{2\sigma_{ik}^2}\right) \tag{3.13}$$

where each membership function in the antecedents of the rule can be represented as an upper and lower membership functions with c and σ representing the centre and standard deviation respectively assigned to the i_{th} input and k_{th} rule of the fuzzy system.

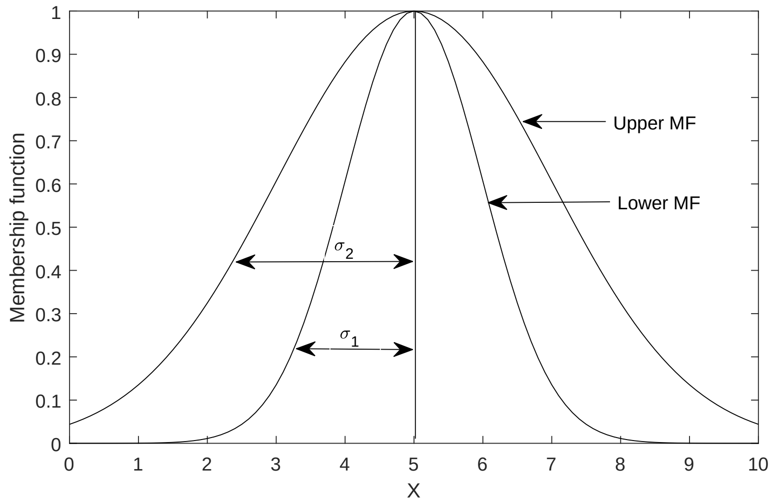

Figure 3.3: Gaussian MF with uncertain standard deviation

In this thesis, the classical Gaussian function is modified with the inclusion of hesitation indices. Thus, for IT2AIFS, Atanassov intuitionistic Gaussian membership (Equations 3.14 and 3.15) and non-membership functions (Equations 3.16 and 3.17) with uncertain standard deviation are utilised which are defined as follows:

$$\overline{\mu_{ik}}(x_i) = exp\left(-\frac{(x_i - c_{ik})^2}{2\sigma_{2,ik}^2}\right) * (1 - \pi_{c,ik}(x_i)) \tag{3.14}$$

$$\underline{\mu_{ik}}(x_i) = exp\left(-\frac{(x_i - c_{ik})^2}{2\sigma_{1,ik}^2}\right) * (1 - \pi_{c,ik}(x_i)) \tag{3.15}$$

$$\overline{\nu_{ik}}(x_i) = (1 - \overline{\pi}_{var,ik}(x_i)) - \left[exp\left(-\frac{(x_i - c_{ik})^2}{2\sigma_{1,ik}^2}\right) * (1 - \pi_{c,ik}(x_i))\right] \tag{3.16}$$

$$\underline{\nu_{ik}}(x_i) = \left(1 - \underline{\pi}_{var,ik}(x_i)\right) - \left[exp\left(-\frac{(x_i - c_{ik})^2}{2\sigma_{2,ik}^2}\right) * (1 - \pi_{c,ik}(x_i))\right] \tag{3.17}$$

where $\pi_{c,ik}$ is the IF-index of centre and $\pi_{var,ik}$ is the IF-index of variance [172]. The parameters $\bar{\sigma}_{2,ik}$, $\underline{\sigma}_{1,ik}$, $\pi_{c,ik}$, $\pi_{var,ik}$ and c are premise parameters that define the degree of membership and non-membership of each element to the fuzzy set $\tilde{A}^*$. Shown in Figure

3.1 is an IT2 A-intuitionistic Gaussian membership and non-membership functions which characterise IT2AIFS. The FOU for the membership is bounded by lower membership and upper membership functions while the FOU of the non-membership is bounded by lower non-membership and upper non-membership functions respectively. The FOUs of the model are as shown in Figure 3.1. The bounds of the FOUs are somewhat wavy (ripples). A concept which incorporates the hesitations in the definition of the FOUs of IT2AIFS. The scaling in Equations 3.14 and 3.15 captures the hesitation of the expert in the definition of the membership function FOU while Equations 3.16 and 3.17 include some shifting which captures the hesitation in the FOU of the non-membership function of the IT2AIFS. This representation satisfies the constraint in Definition 3.3.1. For instance, the membership and non-membership grades of $x = 4.0$ in Figure 3.4 are approximately $\{0.60, 0.88, 0.11, 0.39\}$, which satisfies the constraints: $0 \leq \overline{\mu}_{\tilde{A}^*}(x) + \underline{\nu}_{\tilde{A}^*}(x) \leq 1 \, and \; 0 \leq \underline{\mu}_{\tilde{A}^*}(x) + \overline{\nu}_{\tilde{A}^*}(x) \leq 1$ as shown below:

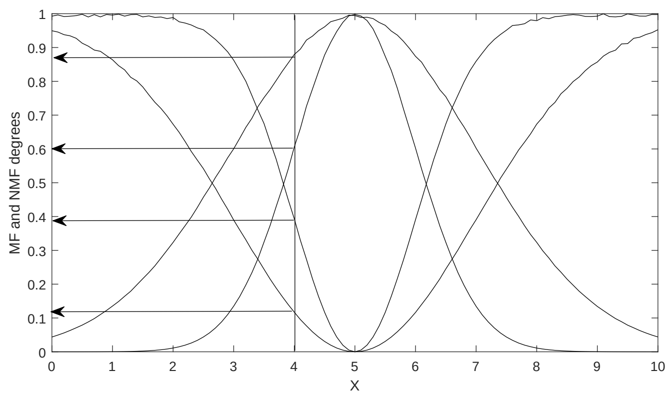

Figure 3.4: IT2AIFS

$$\begin{aligned}
\overline{\mu}_{\tilde{A}^*}(x) + \underline{\nu}_{\tilde{A}^*}(x) &= 0.88 + 0.11 \\
&= 0.99 \in [0, 1] \\
\pi_{\tilde{A}^*}(x) &= 1 - 0.99 \\
&= 0.01 \in [0, 1]
\end{aligned}$$

$$\begin{aligned}
\underline{\mu}_{\tilde{A}^*}(x) + \overline{\nu}_{\tilde{A}^*}(x) &= 0.60 + 0.39 \\
&= 0.99 \in [0, 1] \\
\pi_{\tilde{A}^*}(x) &= 1 - 0.99 \\
&= 0.01 \in [0, 1]
\end{aligned}$$

3.5.2 Rules

The rule representation of IT2AIFLS is similar to the classical IT2FLS, the only exception is that both membership and non-membership functions are involved in the inputs of the IT2AIFLS, that is, the fuzzy sets are IT2AIFSs. The IF-THEN rule of an IT2AIFLS can thus be expressed as follows:

$$\begin{aligned} R_k : \text{IF } x_1 \text{ is } \tilde{A}^*{}_{1k} \text{ and } \cdots \text{and } x_n \text{ is } \tilde{A}^*{}_{nk} \\ \text{THEN } y_k \text{ is } f\left(x_1, x_2, \cdots, x_n\right) \\ = w_{1k}x_1 + w_{2k}x_2 + \cdots + w_{nk}x_n + b_k \end{aligned} \quad (3.18)$$

where $\tilde{A}^*{}_{1k}$, $\tilde{A}^*{}_{2k}$, $\cdots$,$\tilde{A}^*{}_{ik}$, $\cdots$, $\tilde{A}^*{}_{nk}$ are IT2AIFS and y_k is the output of the kth rule formed by linear combination of the input vector: $(x_1, x_2, \cdots, x_n)$. The above general rule for IT2AIFLS can be decomposed into both membership and non-membership functions as follows:
For the function indicating membership in an IT2AIFS, the rule in Equation 5.1 translates to:

$$R_k^{\mu} : \text{IF } x_1 \text{ is } \tilde{A}^*{}^{\mu}_{1k} \text{ and } \cdots \text{and } x_n \text{ is } \tilde{A}^*{}^{\mu}_{nk} \text{ THEN } y_k^{\mu} = w_{1k}^{\mu}x_1 + w_{2k}^{\mu}x_2 + \cdots + w_{nk}^{\mu}x_n + b_k^{\mu} \quad (3.19)$$

For the function indicating non-membership in an IT2AIFS, the rule becomes:

$$R_k^{\nu} : \text{IF } x_1 \text{ is } \tilde{A}^*{}^{\nu}_{1k} \text{ and } \cdots \text{and } x_n \text{ is } \tilde{A}^*{}^{\nu}_{nk} \text{ THEN } y_k^{\nu} = w_{1k}^{\nu}x_1 + w_{2k}^{\nu}x_2 + \cdots + w_{nk}^{\nu}x_n + b_k^{\nu} \quad (3.20)$$

where y_k^{μ} and y_k^{ν} are the membership and non-membership outputs of the kth rule, $w's$ are the function parameters (coefficients of the independent variables) plus a constant term b known as the bias.

3.5.3 Inference

There are generally two main types of fuzzy inferencing namely: Mamdani and TSK which differ in their representation and output evaluation. In this work, a TSK fuzzy inferencing where the output of each IF-THEN rule is a linear function is assumed. What necessitate this assumption is the intention to optimise the parameters of the proposed model with the aim of obtaining more accurate input-output relationships between pairs of data as possible under uncertainty.

In this research, model II, otherwise known as A2-C0, discussed in Chapter 2, Subsection 2.4.3, is adopted to investigate the reasoning behind IT2AIFLS with learning ability similar to adaptive-neuro fuzzy inference system (ANFIS) [194] and T2-ANFIS [195] approaches. The antecedent parts of the IT2AIFLS are IT2AIFS while the consequent parts are linear functions of the inputs. An IT2AIFLS structure with two inputs, three membership and non-membership functions and nine rules is as shown in Figure 3.5.

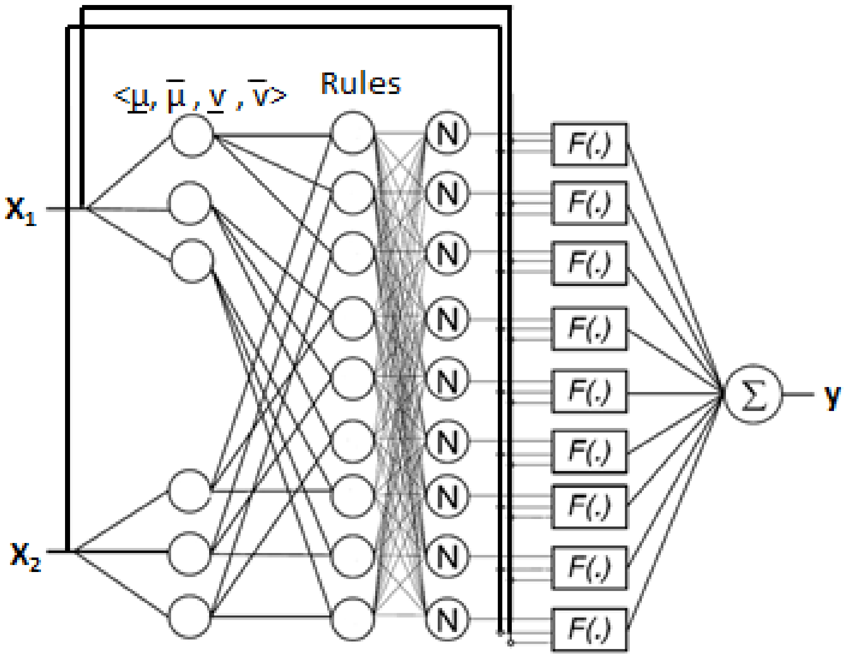

Figure 3.5: An IT2AIFLS Structure - adapted from [4]

3.5.4 Output Processing

According to Hajek and Olej [179], the output of AIFLS-TSK can be computed using two approaches: (i) by the composition of membership function output, y^{μ}, and non-membership function output, y^{ν} [173] and (ii) by defuzzification methods based on weighted average and weighted sum [179]. This research adopts the direct defuzzification based on weighted average of the membership and non-membership functions to compute the output of the proposed IT2AIFLS. In particular, to align with closed-form representation, the proposed IT2AIFLS adopts the inference mechanism proposed in Begian *et al.* [98] for the classical IT2FLS which is expressed as:

$$y = (1-\beta)\frac{\sum_{k=1}^{M} \underline{f}_k y_k}{\sum_{k=1}^{M} \underline{f}_k} + \beta\frac{\sum_{k=1}^{M} \bar{f}_k y_k}{\sum_{k=1}^{M} \bar{f}_k} \tag{3.21}$$

with the condition that:

$$0 \leq (1-\beta)\frac{\underline{f}_k}{\sum_{k=1}^{M}\underline{f}_k} + \beta\frac{\bar{f}_k}{\sum_{k=1}^{M}\bar{f}_k} \leq 1 \tag{3.22}$$

The output of IT2AIFLS in closed-form is an offshoot of Equation 3.21 and represented as [2, 46]:

$$y = \frac{(1-\beta)\sum_{k=1}^{M}\left(\underline{f_k^{\mu}} + \overline{f_k^{\mu}}\right)y_k^{\mu}}{\sum_{k=1}^{M}\underline{f_k^{\mu}} + \sum_{k=1}^{M}\overline{f_k^{\mu}}} + \frac{\beta\sum_{k=1}^{M}\left(\underline{f_k^{\nu}} + \overline{f_k^{\nu}}\right)y_k^{\nu}}{\sum_{k=1}^{M}\underline{f_k^{\nu}} + \sum_{k=1}^{M}\overline{f_k^{\nu}}} \tag{3.23}$$

With this representation, the contribution of each rule to the final output becomes:

$$r_k = \frac{(1-\beta)\left(\underline{f_k^{\mu}} + \overline{f_k^{\mu}}\right)}{\sum_{k=1}^{M}\underline{f_k^{\mu}} + \sum_{k=1}^{M}\overline{f_k^{\mu}}} + \frac{\beta\left(\underline{f_k^{\nu}} + \overline{f_k^{\nu}}\right)}{\sum_{k=1}^{M}\underline{f_k^{\nu}} + \sum_{k=1}^{M}\overline{f_k^{\nu}}} \tag{3.24}$$

where $\underline{f}_k^{\mu}$, $\overline{f}_k^{\mu}$, $\underline{f}_k^{\nu}$ and $\overline{f}_k^{\nu}$ are the lower membership, upper membership, lower non-membership and upper non-membership firing strengths respectively. This is a modification of a novel inference method proposed in [98] for IT2-TSK fuzzy systems and motivated by the Nie-Tan [103] closed form type-reduction method for IT2FLSs where iterations are not required in the computation of the defuzzified crisp value but depends only on the lower and upper bounds of the membership function FOUs. As shown in Equation 3.23, the final output of IT2AIFLS apart from also utilising the bounds of the membership function FOUs, also utilises the upper and lower bounds of the non-membership function FOUs with an additional design factor β [98] to weigh their contributions in the final output, similar to Equation 3.21. In this thesis, the implication operator employed is the "prod" $t-norm$ such that:

$$\underline{f_k^{\mu}}(x) = \underline{\mu}_{\tilde{A}^*{}_{1k}}(x_1) * \underline{\mu}_{\tilde{A}^*{}_{2k}}(x_2) * \cdots * \underline{\mu}_{\tilde{A}^*{}_{nk}}(x_n)$$

$$\overline{f_k^{\mu}}(x) = \overline{\mu}_{\tilde{A}^*{}_{1k}}(x_1) * \overline{\mu}_{\tilde{A}^*{}_{2k}}(x_2) * \cdots * \overline{\mu}_{\tilde{A}^*{}_{nk}}(x_n)$$

$$\underline{f_k^{\nu}}(x) = \underline{\nu}_{\tilde{A}^*{}_{1k}}(x_1) * \underline{\nu}_{\tilde{A}^*{}_{2k}}(x_2) * \cdots * \underline{\nu}_{\tilde{A}^*{}_{nk}}(x_n)$$

$$\overline{f_k^{\nu}}(x) = \overline{\nu}_{\tilde{A}^*{}_{1k}}(x_1) * \overline{\nu}_{\tilde{A}^*{}_{2k}}(x_2) * \cdots * \overline{\nu}_{\tilde{A}^*{}_{nk}}(x_n)$$

where $*$ is the "prod" operator, y_k^{μ} and y_k^{ν} are the outputs of the kth rule corresponding to membership and non-membership functions respectively. In IT2AIFLS, the final output is a weighted average of each IF-THEN rule's output and as such do not require any defuzzification procedure [173]. The parameter β is a user defined parameter, $0 \leq \beta \leq 1$;

specifying the contribution of the membership and non-membership values in the final output. Obviously, if $\beta = 0$, the outputs of the IT2AIFLS is determined by the membership function only and if $\beta = 1$, then only the non-membership will contribute to the system's outputs. If $0 \leq \beta \leq 1$, then both the membership and non-membership functions contribute to the final output. The parameter β in Equation 3.23 is initially specified and then tuned to allow for adaptive adjustment of the membership and non-membership functions in the final output. With the neural network learning ability, the parameters of the IT2AIFS are tuned with learning algorithms as discussed in Chapters 4, 5 and 6.

To support the argument in this thesis, the proposed model of IT2AIFLS is presented and evaluated on well known benchmark datasets from diverse domains and characteristics. Details of these are presented in Chapters 4, 5 and 6.

3.6 Summary

In this chapter, the detailed design of a new class of IT2FLS, otherwise known as IT2AIFLS utilising the new IT2AIFS is presented. In addition, the IT2AIFS and existing IVAIFS are clearly distinguished and a new definition for a T2AIFS is given. The proposed model merges the capabilities of AIFS and IT2FS in a synergistic manner coupled with the ANN learning capability. The embedded ANN architecture in the proposed model allows for the optimisation of its parameters. The proposed framework is intended to be more robust with the capacity to capture more information and enable hesitation. In this way, important limiting assumptions underlying existing approaches of AIFSs, IVAIFSs and IT2FSs are relaxed. The next chapters demonstrate how effective this framework can be used to achieve an improved system performance.

Chapter 4

Gradient Descent Learning of IT2AIFLS with Application to Time Series and Regression Problems

Knowledge is an unending adventure
at the edge of uncertainty

Jacob Bronowski

4.1 Introduction

In Section 2.6, some existing approaches for tuning the parameters of a FLS for improved performance in uncertain environments are listed. In this chapter, a novel application of the GD method to tune the parameters of the developed IT2AIFLS framework is presented. The empirical evaluation is carried out in the context of simulations of benchmark time series and regression problems to aid comparison with existing studies in the literature.

4.2 IT2AIFLS Rule Structure

For ease of reference, the IF-THEN rule of an IT2AIFLS discussed in Subsection 3.5.2 is recalled. The generic TSK rule representation is expressed in Equation 4.1:

$$R_k : \text{IF } x_1 \text{ is } \tilde{A}^*{}_{1k} \text{ and } x_2 \text{ is } \tilde{A}^*{}_{2k} \text{ and } \cdots \text{and } x_n \text{ is } \tilde{A}^*{}_{nk}$$
$$\text{THEN } y_k = \sum_{i=1}^{n} w_{ik} x_i + b_k \quad (4.1)$$

where $\tilde{A}^*{}_{1k}$, $\tilde{A}^*{}_{2k}$, $\cdots$, $\tilde{A}^*{}_{ik}$, $\cdots$, $\tilde{A}^*{}_{nk}$ are IT2AIFS and y_k is the output of the kth rule. For the function indicating membership in an IT2AIFS, the rule in Eqn (4.1) is decomposed to:

$$R_k^{\mu} : \text{IF } x_1 \text{ is } \tilde{A}^{*\mu}_{1k} \text{ and } x_2 \text{ is } \tilde{A}^{*\mu}_{2k} \text{ and } \cdots \text{and } x_n \text{ is } \tilde{A}^{*\mu}_{nk}$$
$$\text{THEN } y_k^{\mu} = \sum_{i=1}^{n} w_{ik}^{\mu} x_i + b_k^{\mu} \quad (4.2)$$

For the function indicating non-membership in an IT2AIFS, the rule becomes:

$$R_k^{\nu} : \text{IF } x_1 \text{ is } \tilde{A}^{*\nu}_{1k} \text{ and } x_2 \text{ is } \tilde{A}^{*\nu}_{2k} \text{ and } \cdots \text{and } x_n \text{ is } \tilde{A}^{*\nu}_{nk}$$
$$\text{THEN } y_k^{\nu} = \sum_{i=1}^{n} w_{ik}^{\nu} x_i + b_k^{\nu} \quad (4.3)$$

where y_k^{μ} is the membership function output and y_k^{ν} is the non-membership function output of the kth rule, w and b are the consequent parameters. This research utilises the GD algorithm for the update of both the antecedent and the consequent parts of the rules. The cost function for a single output is defined as:

$$E = \frac{1}{2} \left(y^a - y \right)^2 \quad (4.4)$$

where y^a is the actual output and y is the proposed model output.

4.3 Parameter Update Rule

In this section, the antecedent and consequent parameters of the developed framework of IT2AIFLS are tuned using the GD method because it is simple, easy to use and is guaranteed to find a minimum (local) [119], for convex optimisation problems.

4.3.1 Consequent Parameter Update

For ease of reference, the generic GD parameter update rule in Subsection 2.6.1 is recalled in Equation 4.5:

$$\theta_{i+1} = \theta_i - \gamma \frac{\partial E}{\partial \theta_i} \tag{4.5}$$

The generic updates for the consequent parameters are as expressed in Equations 4.6 and 4.7.

$$w_{ik}(t+1) = w_{ik}(t) - \gamma \frac{\partial E}{\partial w_{ik}} \tag{4.6}$$

$$b_k(t+1) = b_k(t) - \gamma \frac{\partial E}{\partial b_k} \tag{4.7}$$

From Equations 4.6 and 4.7, the consequent parameters (w and b) update for membership functions are as expressed in Equations 4.8 and 4.9.

$$w_{ik}^{\mu}(t+1) = w_{ik}^{\mu}(t) - \gamma \frac{\partial E}{\partial w_{ik}^{\mu}} \tag{4.8}$$

$$b_k^{\mu}(t+1) = b_k^{\mu}(t) - \gamma \frac{\partial E}{\partial b_k^{\mu}} \tag{4.9}$$

and for the non-membership function, the consequent parameters update are as defined in Equations 4.10 and 4.11.

$$w_{ik}^{\nu}(t+1) = w_{ik}^{\nu}(t) - \gamma \frac{\partial E}{\partial w_{ik}^{\nu}} \tag{4.10}$$

$$b_k^{\nu}(t+1) = b_k^{\nu}(t) - \gamma \frac{\partial E}{\partial b_k^{\nu}} \tag{4.11}$$

where γ is the learning rate (step size) that must be carefully chosen as a large value may lead to instability, and small value on the other hand may lead to a slow learning process. The learning rates were chosen using trial and error approach. The learning rate and IF-indices for this research are assumed to be fixed, that is, they are not tuned. The derivatives in Equations 4.6 and 4.7 are computed as follows:

$$\begin{aligned}\frac{\partial E}{\partial w_{ik}} &= \frac{\partial E}{\partial y}\frac{\partial y}{\partial y_k}\frac{\partial y_k}{\partial w_{ik}} = \sum_k \frac{\partial E}{\partial y}\left[\frac{\partial y}{\partial y_k^{\mu}}\frac{\partial y_k^{\mu}}{\partial w_{ik}^{\mu}} + \frac{\partial y}{\partial y_k^{\nu}}\frac{\partial y_k^{\nu}}{\partial w_{ik}^{\nu}}\right] \\ &= \left(y\left(t\right) - y^a\left(t\right)\right) * \left[(1-\beta)\left(\frac{\underline{f}_k^{\mu}}{\sum_{k=1}^{M}\underline{f}_k^{\mu} + \sum_{k=1}^{M}\overline{f}_k^{\mu}} + \frac{\overline{f}_k^{\mu}}{\sum_{k=1}^{M}\underline{f}_k^{\mu} + \sum_{k=1}^{M}\overline{f}_k^{\mu}}\right)\right. \\ &\quad \left. + \beta\left(\frac{\underline{f}_k^{v}}{\sum_{k=1}^{M}\underline{f}_k^{v} + \sum_{k=1}^{M}\overline{f}_k^{v}} + \frac{\overline{f}_k^{v}}{\sum_{k=1}^{M}\underline{f}_k^{v} + \sum_{k=1}^{M}\overline{f}_k^{v}}\right)\right] * x_i\end{aligned} \tag{4.12}$$

$$\frac{\partial E}{\partial b_k} = \frac{\partial E}{\partial y}\frac{\partial y}{\partial y_k}\frac{\partial y_k}{\partial b_k} = \sum_k \frac{\partial E}{\partial y}\left[\frac{\partial y}{\partial y_k^\mu}\frac{\partial y_k^\mu}{\partial b_k^\mu} + \frac{\partial y}{\partial y_k^\nu}\frac{\partial y_k^\nu}{\partial b_k^\nu}\right]$$
$$= (y(t) - y^a(t)) * \left[(1-\beta)\left(\frac{\underline{f}_k^\mu}{\sum_{k=1}^M \underline{f}_k^\mu + \sum_{k=1}^M \overline{f}_k^\mu} + \frac{\overline{f}_k^\mu}{\sum_{k=1}^M \underline{f}_k^\mu + \sum_{k=1}^M \overline{f}_k^\mu}\right)\right.$$
$$\left. + \beta\left(\frac{\underline{f}_k^v}{\sum_{k=1}^M \underline{f}_k^v + \sum_{k=1}^M \overline{f}_k^v} + \frac{\overline{f}_k^v}{\sum_{k=1}^M \underline{f}_k^v + \sum_{k=1}^M \overline{f}_k^v}\right)\right] * 1 \quad (4.13)$$

where y_k is defined as in Equation 4.1.

4.3.2 Antecedent Parameter Update

The antecedent parts of the rules accept crisp external values which are fuzzified using the membership and non-membership functions of the IT2AIFLS. For optimal performance of the developed framework, the antecedent parameters are also tuned to obtain values that are good enough for estimation close to the actual output values.

In this study, the classical Gaussian function is modified with the inclusion of hesitation indices. Thus, for IT2AIFS, intuitionistic Gaussian membership functions (Equations 3.14 and 3.15) and non-membership functions (Equations 3.16 and 3.17) with uncertain standard deviation are utilised which are defined as follows [2, 46, 47]:

$$\overline{\mu_{ik}}(x_i) = exp\left(-\frac{(x_i - c_{ik})^2}{2\sigma_{2,ik}^2}\right) * (1 - \pi_{c,ik}(x_i)) \quad (4.14)$$

$$\underline{\mu_{ik}}(x_i) = exp\left(-\frac{(x_i - c_{ik})^2}{2\sigma_{1,ik}^2}\right) * (1 - \pi_{c,ik}(x_i)) \quad (4.15)$$

$$\overline{\nu_{ik}}(x_i) = (1 - \overline{\pi}_{var,ik}(x_i)) - \left[exp\left(-\frac{(x_i - c_{ik})^2}{2\sigma_{1,ik}^2}\right) * (1 - \pi_{c,ik}(x_i))\right] \quad (4.16)$$

$$\underline{\nu_{ik}}(x_i) = \left(1 - \underline{\pi}_{var,ik}(x_i)\right) - \left[exp\left(-\frac{(x_i - c_{ik})^2}{2\sigma_{2,ik}^2}\right) * (1 - \pi_{c,ik}(x_i))\right] \quad (4.17)$$

The antecedent parameters in Equation 4.14 to 4.17 are c, σ_1 and σ_2. The centre, c, for both membership and non-membership functions, is the same. For the membership function, the standard deviation (σ) for lower membership function is σ_1 and the standard deviation of the upper membership function is σ_2. For the non-membership functions, the reverse is the case (i.e. $\underline{\nu} \leftarrow \sigma_2$ and $\overline{\nu} \leftarrow \sigma_1$).

$$c_{ik}(t+1) = c_{ik}(t) - \gamma\frac{\partial E}{\partial c_{ik}} \quad (4.18)$$

$$\sigma_{1,ik}(t+1) = \sigma_{1,ik}(t) - \gamma \frac{\partial E}{\partial \sigma_{1,ik}} \tag{4.19}$$

$$\sigma_{2,ik}(t+1) = \sigma_{2,ik}(t) - \gamma \frac{\partial E}{\partial \sigma_{2,ik}} \tag{4.20}$$

The derivatives in Equation 4.18 to 4.20 are computed as follows:

$$\frac{\partial E}{c_{ik}} = \sum_k \frac{\partial E}{\partial y} \left[\frac{\partial y}{\partial \underline{f}_k^\mu} \frac{\partial \underline{f}_k^\mu}{\partial \underline{\mu}_{ik}} \frac{\partial \underline{\mu}_{ik}}{\partial c_{ik}} + \frac{\partial y}{\partial \overline{f}_k^\mu} \frac{\partial \overline{f}_k^\mu}{\partial \overline{\mu}_{ik}} \frac{\partial \overline{\mu}_{ik}}{\partial c_{ik}} + \frac{\partial y}{\partial \underline{f}_k^\nu} \frac{\partial \underline{f}_k^\nu}{\partial \underline{\nu}_{ik}} \frac{\partial \underline{\nu}_{ik}}{\partial c_{ik}} + \frac{\partial y}{\partial \overline{f}_k^\nu} \frac{\partial \overline{f}_k^\nu}{\partial \overline{\nu}_{ik}} \frac{\partial \overline{\nu}_{ik}}{\partial c_{ik}} \right] \tag{4.21}$$

$$\frac{\partial E}{\sigma_{1,ik}} = \sum_k \frac{\partial E}{\partial y} \left[\frac{\partial y}{\partial \underline{f}_k^\mu} \frac{\partial \underline{f}_k^\mu}{\partial \underline{\mu}_{ik}} \frac{\partial \underline{\mu}_{ik}}{\partial \sigma_{1,ik}} + \frac{\partial y}{\partial \underline{f}_k^\nu} \frac{\partial \underline{f}_k^\nu}{\partial \underline{\nu}_{ik}} \frac{\partial \underline{\nu}_{ik}}{\partial \sigma_{2,ik}} \right] \tag{4.22}$$

$$\frac{\partial E}{\sigma_{2,ik}} = \sum_k \frac{\partial E}{\partial y} \left[\frac{\partial y}{\partial \overline{f}_k^\mu} \frac{\partial \overline{f}_k^\mu}{\partial \overline{\mu}_{ik}} \frac{\partial \overline{\mu}_{ik}}{\partial \sigma_{2,ik}} + \frac{\partial y}{\partial \overline{f}_k^\nu} \frac{\partial \overline{f}_k^\nu}{\partial \overline{\nu}_{ik}} \frac{\partial \overline{\nu}_{ik}}{\partial \sigma_{1,ik}} \right] \tag{4.23}$$

where:

$$\frac{\partial y}{\partial \underline{f}_k^\mu} = \frac{\partial y}{\partial \overline{f}_k^\mu} = (1-\beta) \left[\frac{y_k^\mu}{\sum_{k=1}^M \underline{f_k} + \sum_{k=1}^M \overline{f}_k^\mu} - \frac{y^\mu}{\sum_{k=1}^M \underline{f}_k^\mu + \sum_{k=1}^M \overline{f}_k^\mu} \right]$$
$$y^\mu = \frac{\sum_{k=1}^M \left(\underline{f}_k^\mu + \overline{f}_k^\mu \right) * y_k^\mu}{\sum_{k=1}^M \underline{f}_k^\mu + \sum_{k=1}^M \overline{f}_k^\mu} \tag{4.24}$$

$$\frac{\partial y}{\partial \underline{f}_k^\nu} = \frac{\partial y}{\partial \overline{f}_k^\nu} = \beta \left[\frac{y_k^\nu}{\sum_{k=1}^M \underline{f}_k^\nu + \sum_{k=1}^M \overline{f}_k^\nu} - \frac{y^\nu}{\sum_{k=1}^M \underline{f}_k^\nu + \sum_{k=1}^M \overline{f}_k^\nu} \right]$$
$$y^\nu = \frac{\sum_{k=1}^M \left(\underline{f}_k^\nu + \overline{f}_k^\nu \right) * y_k^\nu}{\sum_{k=1}^M \underline{f}_k^\nu + \sum_{k=1}^M \overline{f}_k^\nu} \tag{4.25}$$

$$\frac{\partial \underline{\mu}_k(x_i)}{\partial c_{ik}} = (1-\pi_{c,ik}) * (x_i - c_{ik}) * exp\left(-\frac{(x_i - c_{ik})^2}{2 * \sigma_{1,ik}^2} \right) / \sigma_{1,ik}^2 \tag{4.26}$$

$$\frac{\partial \overline{\mu}_k(x_i)}{\partial c_{ik}} = (1-\pi_{c,ik}) * (x_i - c_{ik}) * exp\left(-\frac{(x_i - c_{ik})^2}{2 * \sigma_{2,ik}^2} \right) / \sigma_{2,ik}^2 \tag{4.27}$$

$$\frac{\partial \underline{\mu}_k(x_i)}{\partial \sigma_{1,ik}} = (1-\pi_{c,ik}) * (x_i - c_{ik})^2 * exp\left(-\frac{(x_i - c_{ik})^2}{2 * \sigma_{1,ik}^2} \right) / \sigma_{1,ik}^3 \tag{4.28}$$

$$\frac{\partial \overline{\mu}_k(x_i)}{\partial \sigma_{2,ik}} = (1-\pi_{c,ik}) * (x_i - c_{ik})^2 * exp\left(-\frac{(x_i - c_{ik})^2}{2 * \sigma_{2,ik}^2} \right) / \sigma_{2,ik}^3 \tag{4.29}$$

The derivatives:

$$\frac{\partial \underline{\nu}_k(x_i)}{\partial c_{ik}} = -(1-\pi_{c,ik}) * (x_i - c_{ik}) * exp\left(-\frac{(x_i - c_{ik})^2}{2 * \sigma_{1,ik}^2} \right) / \sigma_{1,ik}^2 \tag{4.30}$$

$$\frac{\partial \overline{\nu}_k(x_i)}{\partial c_{ik}} = -(1-\pi_{c,ik}) * (x_i - c_{ik}) * exp\left(-\frac{(x_i-c_{ik})^2}{2*\sigma_{2,ik}^2}\right)/\sigma_{2,ik}^2 \tag{4.31}$$

$$\frac{\partial \underline{\nu}_k(x_i)}{\partial \sigma_{1,ik}} = -(1-\pi_{c,ik}) * (x_i - c_{ik})^2 * exp\left(-\frac{(x_i-c_{ik})^2}{2*\sigma_{1,ik}^2}\right)/\sigma_{1,ik}^3 \tag{4.32}$$

$$\frac{\partial \overline{\nu}_k(x_i)}{\partial \sigma_{2,ik}} = -(1-\pi_{c,ik}) * (x_i - c_{ik})^2 * exp\left(-\frac{(x_i-c_{ik})^2}{2*\sigma_{2,ik}^2}\right)/\sigma_{2,ik}^3 \tag{4.33}$$

With the use of a t-norm "prod" operator,

$$\frac{\partial \underline{f}_k^{\mu}}{\partial \underline{\mu}_{ik}} = \prod_{j=1,j\neq i}^{M1} \underline{\mu}_{jk}, \qquad \frac{\partial \overline{f}_k^{\mu}}{\partial \overline{\mu}_{ik}} = \prod_{j=1,j\neq i}^{M1} \overline{\mu}_{jk} \tag{4.34}$$

$$\frac{\partial \underline{f}_k^{v}}{\partial \underline{\nu}_{ik}} = \prod_{j=1,j\neq i}^{M1} \underline{\nu}_{jk}, \qquad \frac{\partial \overline{f}_k^{v}}{\partial \overline{\nu}_{ik}} = \prod_{j=1,j\neq i}^{M1} \overline{\nu}_{jk} \tag{4.35}$$

The output of IT2AIFLS in closed-form is represented as [2, 46, 47]:

$$y = \frac{(1-\beta)\sum_{k=1}^{M}\left(\underline{f}_k^{\mu} + \overline{f}_k^{\mu}\right) y_k^{\mu}}{\sum_{k=1}^{M}\underline{f}_k^{\mu} + \sum_{k=1}^{M}\overline{f}_k^{\mu}} + \frac{\beta\sum_{k=1}^{M}\left(\underline{f}_k^{\nu} + \overline{f}_k^{\nu}\right) y_k^{\nu}}{\sum_{k=1}^{M}\underline{f}_k^{\nu} + \sum_{k=1}^{M}\overline{f}_k^{\nu}} \tag{4.36}$$

where $\underline{f}_k^{\mu}$, $\overline{f}_k^{\mu}$, $\underline{f}_k^{\nu}$ and $\overline{f}_k^{\nu}$ are the lower membership, upper membership, lower non-membership and upper non-membership firing strengths respectively.

The parameter β in Equation 4.36 is tuned for adaptive adjustment of the membership and non-membership function values in the final output. The value of β is adjusted as follows:

$$\begin{aligned}\beta(t+1) = \beta(t) - \gamma\frac{\partial E}{\partial \beta}, \ \frac{\partial E}{\partial \beta} &= \frac{\partial E}{\partial y}\frac{\partial y}{\partial \beta} \\ &= (y - y^a)\left(\frac{\sum_{k=1}^{M}\left(\underline{f}_k^{v} + \overline{f}_k^{v}\right) * y_k^{v}}{\sum_{k=1}^{M}\underline{f}_k^{v} + \sum_{k=1}^{M}\overline{f}_k^{v}} - \frac{\sum_{k=1}^{M}\left(\underline{f}_k^{\mu} + \overline{f}_k^{\mu}\right) * y_k^{\mu}}{\sum_{k=1}^{M}\underline{f}_k^{\mu} + \sum_{k=1}^{M}\overline{f}_k^{\mu}}\right)\end{aligned} \tag{4.37}$$

In the next section, the experimental analysis and discussion of simulation results are presented.

4.4 Experiments and Results

In this section, the experimental analyses on publicly available benchmark time series and regression problems are presented. The datasets and the criteria used in the evaluation were carefully selected to facilitate comparison of the approach introduced here with existing methods in the literature. Each of the datasets are arranged as closely as possible to those reported previously in the literature. The robustness of the approach is measured by evaluation in the presence of some noise in the data such as the Friedman problem [196]. The performance metrics are evaluated on the test dataset. Using the test dataset to evaluate model performance gives an unbiased estimate of the model errors [1]. The following performance metrics are adopted to aid comparison with existing studies in the literature.

- the mean squared error (MSE)

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(y_i^a - y_i)^2 \tag{4.38}$$

- the root mean squared error (RMSE)

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i^a - y_i)^2} \tag{4.39}$$

- the mean absolute error (MAE)

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|y^a - y| \tag{4.40}$$

- the non-dimensional error index (NDEI)

$$NDEI = \frac{RMSE}{std(y^a)} \tag{4.41}$$

- the normalised mean squared error (NMSE)

$$NMSE = \frac{1}{N}\sum_{i=1}^{N}(y - y^a)^2 / \frac{1}{N}\sum_{i=1}^{N} y * \frac{1}{N}\sum_{i=1}^{N} y^a \tag{4.42}$$

where y^a is the actual output, y is the output of the model and N is the number of testing data points. The MSE is a quadratic expression which has only one minimum. According to Picton [197], GD method is a faster approach at arriving at this minimum where the parameter adjustments are proportional to the derivative of the error functions, but in opposite direction.

The number of parameters of the proposed framework for all datasets is $8n + 2M(n + 1)$, where n is the number of inputs, and M is the number of rules. For each input in this study, the number of linguistic terms are arbitrarily set to two in order to reduce the

[1] https://uk.mathworks.com

computational burden of the system. The β value for all experiments is initialised to 0.5 to ensure equal initial membership and non-membership contributions. The initial values of membership and non-membership function consequent parameters are randomly generated from unit interval [0,1]. For all experiments, it is assumed that there are uncertainties in only the antecedent part of each rule. The entire experiments were conducted using $MATLAB^{©}$ 2016 running on a 64-bit Intel core i3-4130 CPU@3.40GHz /8GB RAM configuration computer.

4.4.1 A Comparison of IT2AIFLS, FIS, IFIS and IT2FLS on Regression Problems

This section compares the performance of IT2AIFLS with FIS, IFIS and IT2FLS on regression problems. The regression datasets used for the analysis are energy, stock and autoMPG6 which are obtained from [198] and Friedman from [196]. The same computational protocol in Hajek and Olej [173] are adopted for Friedman, energy, stock and autoMPG6 dataset to aid comparison with FIS and IFIS.

Datasets Description

Friedman [196]: The Friedman prediction problem uses a synthetic dataset with the following data generation formula:

$$y = 10sin\left(\pi x_1 x_2\right) + 20\left(x_3 - 0.5\right)^2 + 10x_4 + 5x_5 + \hat{n} \tag{4.43}$$

where $x_i{}'s$ are the input variables and $\hat{n}$ is white Gaussian noise with a mean of zero and standard deviation of 1. Shown in Figure 4.1 is the histogram of additive white Gaussian noise distribution for Friedman problem. The Friedman data consists of five input variables x_1, x_2, x_3, x_4, x_5 independently and uniformly distributed over [0, 1] and one target variable, y, generated using equation (4.43). For the Friedman dataset, 1200 data samples are randomly generated which are then split equally into 600 samples for training and 600 samples for testing (this is referred to as Friedman#1). There are a total of 32 rules for Friedman dataset with 8(5) + 2*32(5+1) = 424 parameters.

Energy [198]: The daily electric energy problem involves the prediction of the daily average price of TkWhe electricity energy in Spain. The data set contains real values from 2003 about the daily consumption in Spain of energy from hydroelectric, nuclear electric, carbon, fuel, natural gas and other special sources of energy. There are a total of 365

Table 4.1: Dataset characteristics

Dataset	No. of input	No. of samples	Train set	Test set
Friedman#1	5	1200	600	600
Friedman#2	5	1200	200	1000
Energy	6	365	183	182
Stock	9	950	475	475
AutoMPG6	5	392	196	196
Elect. volt. line	2	495	396	99
Elect.Maint	4	1059	847	212
Mackey-Glass	4	1000	500	500
Annual sunspot	4	280	221	(35)(24)
Tree ring	8	1533	1150	383
Canadian lynx	7	114	100	14
Abalone	8	4177	3342	835
House sales	15	21613	15129	6484

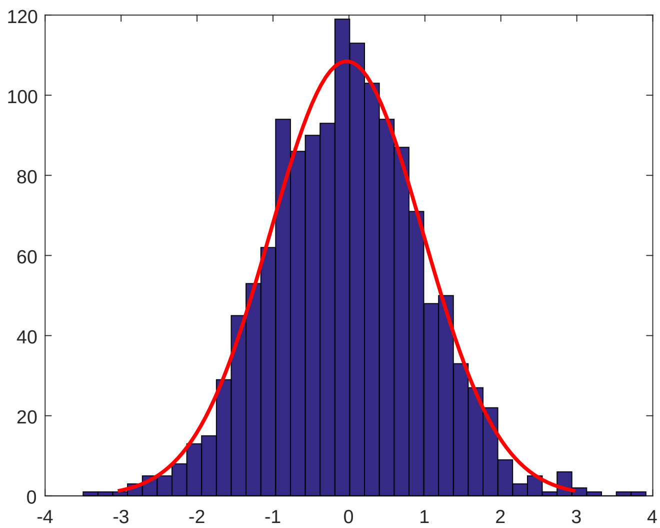

Figure 4.1: Histogram of additive white Gaussian noise for noisy Friedman problem

data instances. For the energy dataset, IT2AIFLS generated 64 rules with a total of 944 parameters. Shown in Table 4.2 is the excerpt from the energy data.

Stock [198]: The stock dataset is a highly non-stationary dataset and consists of daily

Table 4.2: Excerpt from energy dataset

	hydroelectric	nuclear	coal	fuel	gas	special	consume
1	179183	175973	78429.1	4680.73	8117.13	8023	1.7328
2	206035	186774	79129.5	4342.43	5715.18	8159	1.5835
3	198435	180633	64465.2	4566.84	0	8215	1.50531
4	187029	171382	51913.4	5342.54	0	8346	0.955205

stock prices from January 1988 to October 1991 for ten aerospace companies. The task is to predict the price of the 10th company based on the prices of the other nine companies. The dataset consists of 950 samples. Stock data is a high-dimensional dataset with a total number of 512 rules and 10312 parameters for the IT2AIFLS. Table 4.3 shows the input samples excerpted from stock data.

Table 4.3: Excerpt from stock dataset

	x1	x2	x3	x4	x5	x6	x7	x8	x9	y
1	17.219	50.5	18.75	43	60.875	26.375	67.75	19	48.75	34.875
2	17.891	51.375	19.625	44	62	26.125	68.125	19.125	48.75	35.625
3	18.438	50.875	19.875	43.875	61.875	27.25	68.5	18.25	49	36.375
4	18.672	51.5	20	44	62.625	27.875	69.375	18.375	49.625	36.25

AutoMPG6 [198]: The task here is to predict the city-cycle fuel consumption in miles per gallon (mpg) in terms of 1 multi-valued discrete and 5 continuous attributes (where two multi-valued discrete attributes - Cylinders and Origin - from the original dataset are removed). For autoMPG6, 392 data samples are available for analysis. The total number of parameters for AutoMPG6 are 424 with 32 rules. Table 4.4 contains the first four samples excerpted from autoMPG6 data.

Table 4.4: Excerpt from autoMPG6 dataset

	displacement	horse-power	weight	acceleration	model_year	mpg
1	91	70	1955	20.5	71	26
2	232	100	2789	15	73	18
3	350	145	4055	12	76	13
4	318	140	4080	13.7	78	17.5

The analysis of the above datasets was previously conducted by Hajek and Olej [173] using type-1 intuitionistic fuzzy inference system (IFIS) and fuzzy inference system (FIS). The study in [173] is extended by employing IT2AIFLS to the same datasets. For ease of comparison, the above datasets are arranged as closely as possible to those reported in Hajek and Olej [173]. The datasets (Friedman#1, energy, stock and autompg6) are randomly sampled 5 times and sequentially split into two equal parts as in Table 4.1 for each run, with 500 training epochs. The results presented in Table 4.5 show the average RMSE and standard deviation over 25 simulations for each dataset. The initial values of the consequent parts of the rule (w and b) for membership and non-membership, are generated randomly from the interval [0, 1] and updated using equations in Subsection 4.3.1. The learning rate is chosen as 0.1. The RMSE defined in Equation 4.39 is used as a performance criterion. Table 4.5 shows the comparison of the RMSE on the test data using IT2AIFLS, FIS, IFIS and IT2FLS (which also use the design parameter β to weigh the lower and upper membership contributions to the final output). From Table 4.5, IT2AIFLS outperforms both FIS and IFIS on the selected test samples. This is consistent with the reports in the literature that T2FLSs (IT2AIFLSs in this case) model uncertainty in certain applications better than T1FLSs [4, 11]. The proposed model with both membership and non-membership functions, in the overall, also outperforms the classical IT2FLS defined with only the fuzzy membership functions.

Table 4.5: RMSE and std of IT2AIFLS vs FIS/IFIS/IT2FLS on regression problems

Models	Friedman#1	Energy	Stock	AutoMPG6
FIS [173]	1.353 ± 0.026	7.443 ± 1.579	1.423± 0.227	3.702 ± 0.211
IFIS [173]	1.332 ± 0.032	4.776 ± 2.776	1.402 ± 0.219	3.684 ± 0.195
IT2FLS	1.095 ± 0.046	0.567 ± 0.125	0.750 ± 0.026	1.792 ± 0.048
IT2AIFLS	**1.026 ± 0.011**	**0.558 ± 0.005**	**0.611 ± 0.006**	**1.700 ± 0.064**

Due to additive noise in the Friedman dataset, 30 Monte-Carlo simulations are also realised and the average RMSE and standard deviation are 1.0865 and 0.058 respectively.

4.4.2 Friedman#2

This example studies the Friedman problem as reported in Juang *et al.* [199]. In this example, further experiments are performed using the Friedman dataset to evaluate the performance of the proposed model on non-fuzzy and fuzzy approaches, particularly with other type-2 fuzzy approaches. For comparison purpose, the experimental set-up as reported in [199] are adopted. Similar to Juang *et al.* [199], 1400 samples are randomly generated using Equation (4.43), 200 samples are used for training, 200 for validation while the remaining 1000 samples are used for testing (referred to as test set 1) and this is repeated 20 times with the average RMSE and standard deviation reported in Table 4.6. The learning rate, $\gamma = 0.1$. The plot of the actual and predicted output is as shown in Figure 4.2. This problem was also analysed in [200] and [201]. While Carney and Cunningham [200] employed neural bootstrap aggregation (NBAG), benchmark and simple bagged ensemble; Lee *et al.* [201] on the other hand proposed a general regression neural network with fuzzy adaptive resonance theory (GRNNFA) for the analysis of this first set of data. Similar to Juang *et al.* [199], noise free nonlinear Friedman equation is also investigated. In this second case, 1000 test samples are generated with $\hat{n} = 0$ (no noise added - referred to as test set 2). Similar studies using this dataset are reported in Juang *et al.* [199] namely, self-constructing neural fuzzy inference network (SONFIN) and support vector based fuzzy model (SVR-FM) are reported for type-1 fuzzy models. The parameters of SONFIN are learned using training-error minimisation through the combination of Kalman filtering and a GD algorithm. For type-2 systems, approaches such as T2FLS, self-evolving interval type-2 fuzzy neural network (SEIT2FNN) and interval type-2 fuzzy neural network with support vector regression (IT2FNN-SVR) are reported. The T2FLS employs GD for parameter learning and referred to as T2FLS-G. The SEIT2FNN is designed with structure learning and utilised rule-ordered Kalman filter together with GD for parameter learning. The SEIT2FNN has IT2FS in the antecedents trained with GD with TSK interval type-1 sets in the consequent. Two flavours of IT2FNN-SVR are proposed in Juang *et al.* [199] namely IT2FNN-SVR(N) and IT2FNN-SVR(F). The difference between these two is in the representation of the input nodes. The former consists of input nodes with numerical values and interval output nodes while the latter consists of input nodes with fuzzy numbers and interval output nodes. The SONFIN and SEIT2FNN are previous studies involving Juang in [199]. The results are compared with these models already reported in the literature as shown in Table 4.6. The results in Table 4.6 indicate the RMSE and standard deviation for AIFLS, IT2AIFLS and similar works in the literature. It is shown that IT2AIFLS exhibits lower RMSE compared to its type-1 counterpart, the non-fuzzy, the two T1FLSs and the T2FLSs. For 30 Monte-Carlo realisations, the average RMSE and standard deviation for IT2AIFLS on Friedman#2 with additive noise are 1.5057 and 0.1022 respectively.

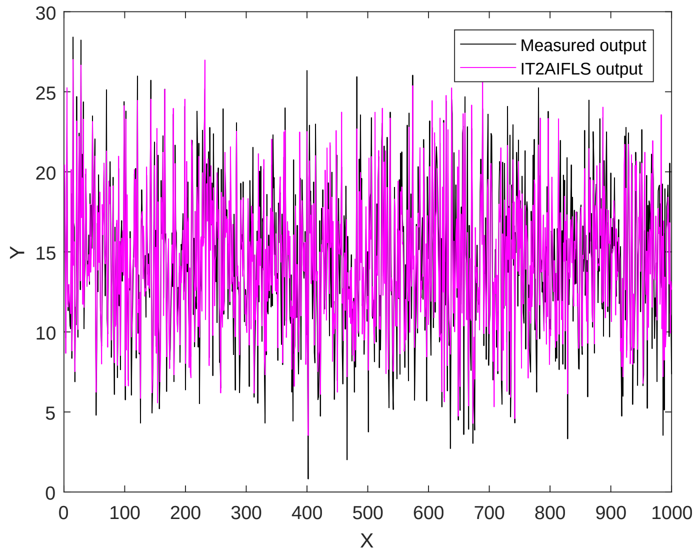

Figure 4.2: Actual and predicted outputs of Friedman with Gaussian white noise

Table 4.6: Performance comparison of IT2AIFLS with existing models on Friedman#2

Models	RMSE(tst1) (N)	Test1(std)	RMSE(tst2) (NF)	Test2 (std)
NBAG [200]	2.1218	-	-	-
Bench [200]	2.3178	-	-	-
Simple [200]	2.2244	-	-	-
GRNNFA [201]	2.136	-	-	-
SONFIN [202]	2.531	0.138	2.398	0.131
T2FLS-G [118]	2.597	0.137	2.479	0.145
SEIT2FNN [121]	1.941	0.170	1.598	0.216
IT2FNN-SVR(N) [199]	1.788	0.145	1.537	0.201
IT2FNN-SVR(F) [199]	1.597	0.120	1.291	0.151
IT2FLS	1.778	0.152	1.419	0.210
AIFLS	2.375	0.129	2.227	0.186
IT2AIFLS	**1.494**	**0.111**	**1.116**	**0.104**

N = Noisy, NF = Noise free

4.4.3 Electrical Engineering Distribution Problems

In Cordón *et al.* [203], two problems involving electrical distribution in rural towns in Spain are reported and have become real-world benchmark problems in fuzzy logic fields.

The task here is to relate some characteristics of certain village with actual low voltage line it contains and also relate the maintenance cost of the network in certain towns with some of their characteristics.

- **Computing the Length of Low Voltage Lines**

The first problem proposed in [203] is to estimate the length of low voltage lines in rural towns using some available inputs. The dataset consist of 495 instances with actual values measured by a company. Table 4.7 contains the first four input samples for computing the length of low voltage lines. The dataset is randomly sampled and divided into 396 instances for training set and 99 instances for testing set with each consisting of three attributes namely:

- Number of clients in the population (inhabitants).
- Radius of i population in the sample (distance).
- Line length, population i (length).

Table 4.7: Excerpt from low voltage line lengths

inhabitants	distance	length
15	605	2146
13	696.669983	2148
25	443.329987	2178
22	373.329987	1322

There are a total of 4 rules generated for low voltage line estimation with 40 parameters. The results presented in Table 4.8 are averaged over 10 simulations (similar to previous studies) with 100 epochs and learning rate set to 0.1. It can be observed in Table 4.8 that IT2AIFLS has superior performance compared to the classical non-linear regression models, neural networks, the evolutionary approaches and other fuzzy approaches.

Further experiments were conducted to ascertain if extra number of parameters leads to improved system performance. To achieve this, the same Gaussian consequents are applied to both the membership and non-membership outputs of the IT2AIFLS, however, with the additional 4 parameters of IF-indices, these translate to 4 rules and 28 parameters for IT2AIFLS with 4 rules and 24 parameters for the classical IT2FLS. The results in Table 4.9 are averaged over 10 simulation runs. As shown in Table 4.9, IT2AIFLS outperforms the classical IT2FLS because of the additional number of parameters in terms of IF-indices.

Table 4.8: Performance comparison of IT2AIFLS with existing models on low voltage line length estimation problem

Models	RMSE(tst)
Linear [203]	457.8821
Exponential [203]	443.8513
Second order Polynomial [203]	450.8126
Third-order polynomial [203]	450.5452
Three layer Perceptron [203]	408.7689
GA-P [203]	399.7962
Interval GA-P [203]	398.4181
WM Fuzzy model [203]	424.384
Mamdani Fuzzy model [203]	408.2511
TSK Fuzzy model [203]	385.3751
Gr + MF [204]	390.7979
Genetic Learning Process [205]	383.4866
HSLR(WM,3,5) [206]	409.04523
GT2FLS-sampling [207]	594.02365
GT2FLS-VSCTR [207]	590.90565
AIFLS	262.2775
IT2AIFLS	**255.3325**

Table 4.9: Performance comparison of IT2AIFLS with classical IT2FLS on voltage length estimation problem

Models	Parameter	RMSE(std)	Run-time (s)
IT2FLS	24	260.7010	12.35
IT2AIFLS	28	**260.1041**	**24.76**

- **Computing the Maintenance Costs of Medium Voltage Lines**

The second problem is to estimate the maintenance cost (not based on real data). The dataset consists of 1059 samples with 5 attributes namely:

- Sum of the length of all street in the town ($x1$).

- Total area of the town ($x2$).

- Area that is occupied by buildings ($x3$).

- Energy supply to the town ($x4$).

- Maintenance costs of medium voltage line (y).

Shown in Table 4.10 are a few samples of the inputs for electricity maintenance cost estimation problem. Similar to previous studies, the 1059 samples are randomly sampled and divided into two sets: 847 instances for training and 212 instances for testing as reported in [203–206]. The IT2AIFLS model is executed for 100 epochs with learning rate set to 0.1. There are 16 rules generated for maintenance cost estimation with a total of 192 parameters. In order to relate the dependent variable (maintenance cost) with the independent variables, the IT2AIFLS is applied to both the training and test sets and results are compared with those in the literature. Figure 4.3 shows the adaptation of β values for electrical maintenance cost estimation problem while Figure 4.4, shows the correlation between the actual and predicted outputs for electrical maintenance cost estimation. This result is significant because it means that IT2AIFLS has a high predictive

Table 4.10: Excerpt from electrical maintenance cost dataset

	x1	x2	x3	x4	y
1	11	3.3	54.96	55	4329.33
2	4	1.2	19.98	40	2016.44
3	0.9	0.27	4.5	1.8	249.42
4	2	1.2	19.98	10	1044.22
5	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]

capability and can be useful in modelling natural attributes of physical phenomena. Table 4.11 shows the performance of IT2AIFLS with other models in the literature in terms of their RMSEs. The results in Table 4.11 show a significant performance improvement of IT2AIFLS over other works in the literature.

4.4.4 Time Series Prediction

In this subsection, the IT2AIFLS is evaluated using well known publicly available benchmark time series problems.

- **Mackey-Glass Time Series**

Mackey-Glass is a well known time series dataset defined by the following differential delay equation:

$$\frac{dx(t)}{dt} = \frac{a * x(t-\tau)}{1 + x(t-\tau)^n} - b * x(t)$$

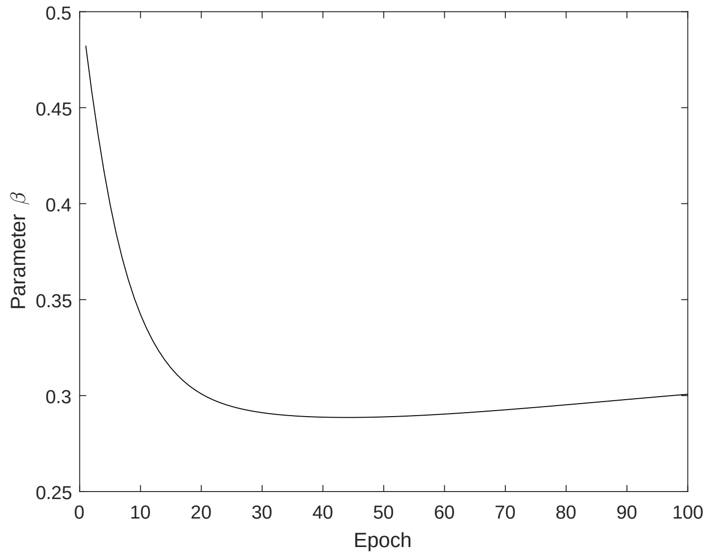

Figure 4.3: Adaptation of β values for electrical maintenance cost

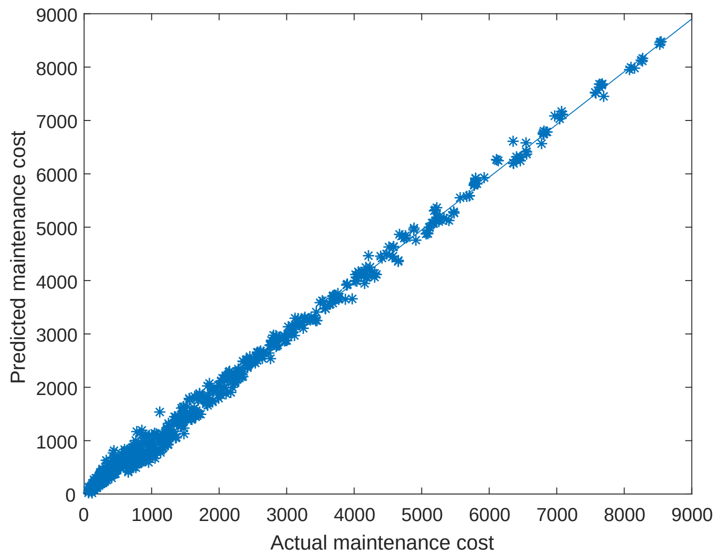

Figure 4.4: Correlation analysis between the actual and predicted outputs for electrical maintenance cost

where a, b and n are constant real numbers, t is the current time and τ is a non-negative time delay constant. The system tends to display a deterministic/periodic behaviour at $\tau \leq 17$ which turns chaotic when $\tau > 17$. For comparison with other works in the literature such as [42, 210–213], the target output, y, is chosen as $x(t+6)$, with input vector

Table 4.11: Performance comparison of IT2AIFLS with other models on electrical maintenance cost estimation problem

Models	RMSE(tst)
Linear [203]	191.8828
Second order Polynomial [203]	212.9131
Three layer Perceptron [203]	181.9478
GA-P [203]	147.9324
Interval GA-P [203]	135.3699
WM Fuzzy model [203]	166.1776
Mamdani Fuzzy model [203]	150.3030
TSK Fuzzy model [203]	108.7934
Gr + MF [204]	102.0490
Genetic LP [205]	102.3034
HSLR(WM,3,5) [206]	154.3276
SA-IT2FLS [208]	75.2400
AT1-SCRATCH [209]	89.6619
AT2-SCRATCH [209]	101.7790
AT2-OPT [209]	88.4542
AT2-BLUR [209]	82.8083
AIFLS	61.1401
IT2AIFLS	**53.7200**

$(x(t-18), x(t-12), x(t-6), x(t))$ and $\tau = 17$. For each input in this study, two linguistic terms are used. Similar to previous studies, 1000 data instances are generated with the first 500 data points used for training and the remaining 500 for testing. The results of applying different approaches to the prediction of Mackey-Glass are listed in Table 4.12 in terms of their RMSE. As shown in Table 4.12, IT2AIFLS outperforms the modified differential evolution radial basis function neural network (MDE-RBF NN) and other fuzzy approaches.

For a fair comparison with existing IT2FLS (TSK), another experiment is conducted with the same computational settings (1000 number of training and 200 testing instances) as reported in Kayacan and Khanesar [42]. The motive for this separate experiment is to compare the performance of the model proposed in this thesis with an existing study utilising classical IT2FLS. Similar to IT2AIFLS, the antecedent and consequent parameters of the IT2FLS in [42] are updated using GD and equally used the same parameter β to adjust the upper and lower values of the membership grades in the final output. As shown in Table 4.13, after training and testing, IT2AIFLS outperforms IT2FLS with the RMSE of 0.0168. An AIFLS for Mackey-Glass prediction is also implemented in order to evaluate the performance of the IT2AIFLS over its T1 model. From Tables 4.12 and 4.13, IT2AIFLS outperforms AIFLS because of the extra degrees of freedom offered by the FOUs of the IT2AIFLSs.

Table 4.12: Performance comparison of IT2AIFLS with other appraches on Mackey-Glass time series forecasting

Models	Train/Test set	RMSE (tst)
ANFIS Ensemble with IT2 FLS [214]	400/400	0.04933
ANFIS Ensemble with T1 FLS [214]	400/400	0.12043
Fuzzy-Singular Value Decomposition [212]	500/500	0.012
MDE-RBF NN [211]	500/500	0.013
Genetic Fuzzy Ensemble [213]	500/500	0.0264
Fuzzy Genetic Algorithm [213]	500/500	0.049
Radial Basis Function AFS [210]	500/500	0.0114
AIFLS	500/500	0.0236
IT2AIFLS	500/500	**0.0079**

Table 4.13: Comparison of IT2FLS-TSK and IT2AIFLS on Mackey-Glass time series

Models	Train/Test set	RMSE(tst)
IT2FLS-TSK [42]	1000/200	0.0250
AIFLS	1000/200	0.0234
IT2AIFLS	1000/200	**0.0168**

- **Annual Sunspot Time Series**

In this example, the annual sunspot time series, a highly complex and non-stationary real world time series is considered to evaluate the effectiveness of the proposed model. The annual sunspot series for the years 1700 to 1979 is investigated. The series reflects the yearly average relative number of sunspots observed and the dataset is available at the

National Geographical Data Center website [215]. For a fair comparison with previous studies as reported in [194, 216–223], the whole dataset is divided into three sets: the data from 1700 to 1920 are used for training, data from 1921 to 1955 form the first test set while data from 1956 to 1979 form the second test set. The input generation vector is: $[x(t-4), x(t-3), x(t-2), x(t-1)]$ with $x(t)$ as the output variable. Table 4.14 contains the first four samples of the sunspot data. For sunspot dataset, 16 rules are generated

Table 4.14: Excerpt from sunspot time series data

	x1	x2	x3	x4	y
1	8.3	18.3	26.7	38.3	60
2	18.3	26.7	38.3	60	96.7
3	26.7	38.3	60	96.7	48.3
4	38.3	60	96.7	48.3	33.3

with 192 parameters. The model is trained for 200 epochs. The performance measure utilised for this time series is the NMSE defined in Equation 4.42.

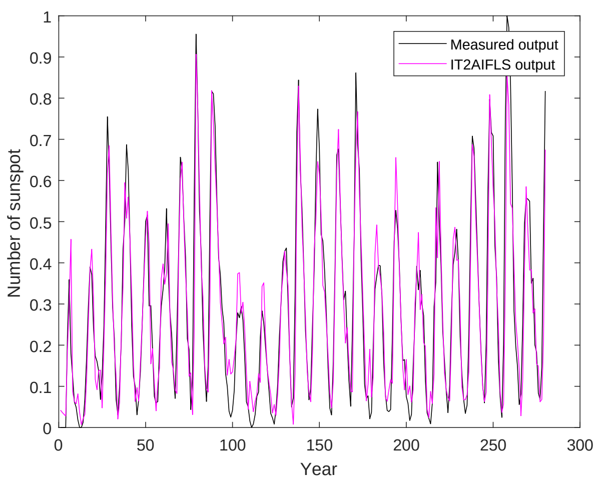

Figure 4.5: Actual and predicted values of sunspot time series using IT2AIFLS

Figure 4.5 shows the actual and predicted output of sunspot time series while Figure 4.6 shows the adaptation of its β values. From Table 4.15, it can be seen that IT2AIFLS has lower NMSE on both test sets compared to its type-1 variant and similar works in the literature. This indicates a good generalisation capability of IT2AIFLS on data not used during the training of the model. Using the same procedure for the results presented in Table 4.9, the performance of the traditional IT2FLS is compared with the proposed intuitionistic version using sunspot time series dataset. The results are averaged over 10

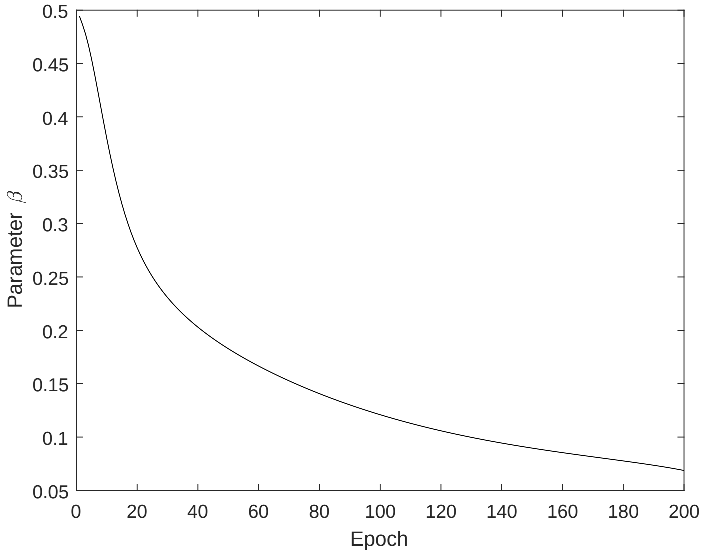

Figure 4.6: Adaptation of β values for sunspot time series

Table 4.15: Performance comparison of IT2AIFLS with other models on sunspot time series

Models	NMSE(trn)	NMSE (tst1)	NMSE (tst2)
Tong and Lim [216]	0.097	0.097	0.28
Weigend [217]	0.082	0.086	0.35
Svarer [218]	0.090	0.082	0.35
Transversal Net [219]	0.0987	0.0971	0.3724
Recurrent Net [219]	0.1006	0.0972	0.4361
RFNN [220]	-	0.074	0.21
ANFIS [194]	0.0550	0.1915	0.4068
FENN [221]	-	-	0.18
FWNN-S [222]	0.0895	0.1093	0.1510
FWNN-R [222]	0.0796	0.1099	0.2549
FWNN-M [222]	0.0828	0.0973	0.1988
LLNF [223]	-	0.085	0.1219
OSSA-LLNF [223]	-	0.0602	0.0846
AIFLS	0.1250	0.0136	0.0174
IT2AIFLS	**0.1207**	**0.0105**	**0.0148**

Table 4.16: Performance comparison of IT2AIFLS with classical IT2FLS using sunspot time series dataset

Models	Parameter	NMSE (trn)	NMSE (tst1)	NMSE (tst2)	Run-time(s)
IT2FLS	184	0.1352	0.0129	0.0189	17.51
IT2AIFLS	192	**0.1258**	**0.0120**	**0.0172**	**32.58**

simulations and presented in Table 4.16. As shown in Table 4.16, IT2AIFLS outperforms classical IT2FLS because of the aditional number of parameters provided by the non-membership functions and IF-indices. It is conjectured that IT2AIFLS is computationally more efficient in terms of prediction accuracies because of the non-membership and IF-indices (hesitations) embedded in the FOUs of both the membership and non-membership functions of IT2AIFSs which increase the fuzziness of the system leading to improved performance. However, in both problem instances, low voltage line estimation and sunspot datasets, the IT2AIFLS suffers some drawback in terms of the runtime. This poses a challenge to this algorithm. Nevertheless, if the main goal of analysis is to predict the model's output as closely as possible to the actual values, and if the computational time is not an issue, then the use of IT2AIFLS for prediction is justified.

- **Tree Ring Time Series**

The tree ring time series obtained from [224] contains annual measures of tree rings width measured in Argentina for the period 441-1974. Similar to Pouzols and Lendasse [225], the data generating format is $[x(t-9), x(t-8), x(t-7), x(t-5), x(t-3), x(t-2), x(t-1), x(t)]$, except $x(t-4)$ and $x(t-6)$. The task is to predict $x(t+1)$ and this represents the width of the tree ring for the next year. Shown in Table 4.17 are the first few samples of the tree ring time series input data. Shown in Figure 4.7 is the plot of the tree ring dataset. The dataset

Table 4.17: Excerpt from tree ring time series data

	x1	x2	x3	x4	x5	x6	x7	x8	y
1	0.731	1.061	1.03	0.904	1.134	0.945	1.043	1.027	0.783
2	1.061	1.03	1.104	1.167	0.945	1.043	1.027	0.783	1.058
3	1.03	1.104	0.904	1.134	1.043	1.027	0.783	1.058	1.029
4	1.104	0.904	1.167	0.945	1.027	0.783	1.058	1.029	0.886
5	0.904	1.167	1.134	1.043	0.783	1.058	1.029	0.886	0.979

is randomly split into 75% training and 25% testing. In Pouzols and Lendasse [225], the evolving fuzzy optimally pruned extreme learning machine (eF-OP-ELM), is reported for analysing this time series. The dynamic evolving neuro-fuzzy inference system (DENFIS), evolving Takagi-Sugeno (eTS) model and online sequential method for fuzzy systems based on online sequential ELM (OS-fuzzy-ELM) are also reported in [225] for tree ring time series analysis. All computational protocols in this study are arranged as close as possible to those reported in [225] to ease comparison with existing studies in the literature. The performance criterion adopted for this analysis is the NDEI defined in Equation 4.41. Table 4.18 shows the average cross validation NDEI and standard deviation of the tree ring dataset for 25 trials. Figure 4.8 shows the actual and predicted output of the tree ring time series. From Table 4.18, IT2AIFLS outperforms other fuzzy models with reduced NDEI.

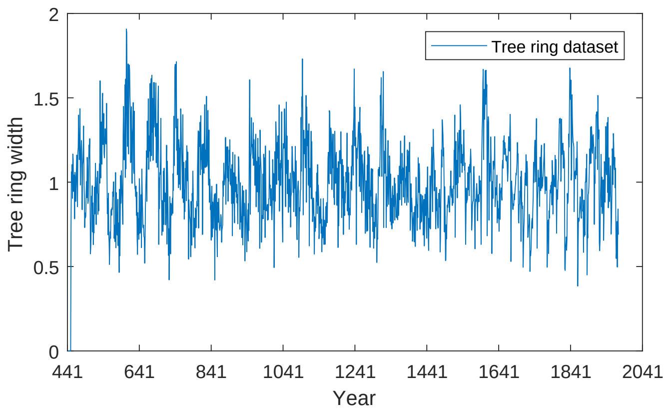

Figure 4.7: Tree ring dataset

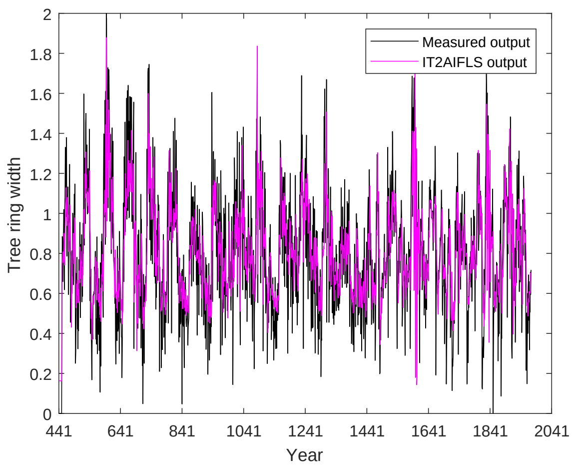

Figure 4.8: Actual and predicted outputs of tree ring time series

Table 4.18: Performance comparison of IT2AIFLS with other models using tree ring time series forecasting

Models	NDEI	NDEI (std)
DENFIS [226]	0.959	0.624
eTS [227]	0.714	0.457
OS-Fuzzy-ELM [228]	0.794	0.511
eF-OP-ELM [225]	0.841	0.536
IT2AIFLS	**0.395**	**0.157**

- **Canadian Lynx Time Series**

The Canadian lynx time series is selected in order to compare the performance of IT2AIFLS against non-fuzzy approaches. Canadian lynx dataset is a time series that shows the number of lynx trapped in the Mckenzie river district per year in northern Canada and corresponds to the period 1821-1934. Similar to previous studies such as [229–231], the logarithms to the base 10 of the data are used in the analysis. Figures 4.9 and 4.10 show the original and the logarithmic transformed data of the Canadian lynx series respectively, with a periodicity of approximately 10 years. Table 4.19 contains the first four samples of the Canadian lynx time series input data. The time series consists of 114 observations of which 100 samples are used for training and the remaining 14 are used for testing. Similar

Table 4.19: Excerpt from Canadian lynx time series data

	x1	x2	x3	x4	x5	x6	x7	y
1	3.5942	3.4504	3.1688	2.94	2.7672	2.5065	2.4298	3.774
2	3.774	3.5942	3.4504	3.1688	2.94	2.7672	2.5065	3.6946
3	3.6946	3.774	3.5942	3.4504	3.1688	2.94	2.7672	3.4111
4	3.4111	3.6946	3.774	3.5942	3.4504	3.1688	2.94	2.7185
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]

to Wang *et al.* [231], the maximum training epoch adopted is 2000. For this time series, the MSE and MAE defined in Equations 4.38 and 4.40 are utilised as the performance evaluation metrics. As shown in Table 4.20, IT2IFLS outperforms the listed non-fuzzy approaches on the Canadian lynx dataset.

- **Santa Fe A Time Series**

This example considers the Santa Fe Laser dataset of the Santa Fe A time series competition obtained from [241]. The data were measured from a far-infra-red laser in a chaotic

Table 4.20: Performance comparison of IT2AIFLS with non-fuzzy models on Canadian lynx time series

Models	MSE(tst)	MAE(tst)
Zhang's ARIMA [229]	0.020486	0.112255
ANN [229]	0.020466	0.112109
ANN (p,d,q) [232]	0.013609	0.089625
Zhang's Hybrid ARIMA/ANNs model [229]	0.017233	0.103972
Hybrid ARIMA/ERNN model [233]	0.009	-
SETAR [234]	0.014	-
FNN [234]	0.009	-
Generalised Hybrid ARIMA/ANNs model [235]	0.00999	0.085055
ANN/PNN model [230]	0.014872	0.079628
ARIMA/PNN model [230]	0.011461	0.084381
MNM-ANN-DEA [236]	0.00663	-
GA-BPNN [231]	0.013599	0.081477
DE-BPNN [231]	0.012899	0.080542
ANN Ensemble [237]	0.00715	-
RBF-AR [238]	0.0073	-
ADE-BPNN [231]	0.010392	0.070723
GMDH [239]	0.0082	0.0634
LSSVM [239]	0.0074	0.0657
GLSSVM [239]	0.0056	0.0552
L&NL-ANN [240]	0.006	-
IT2AIFLS	**0.00463**	**0.0205**

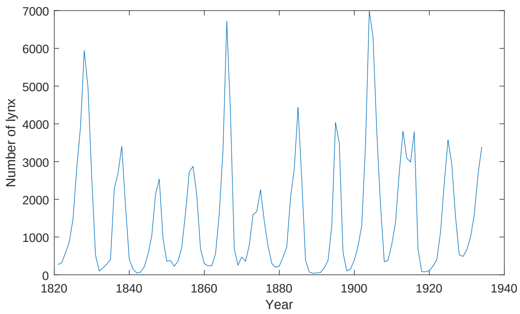

Figure 4.9: Original Canadian lynx time series (1821-1934)

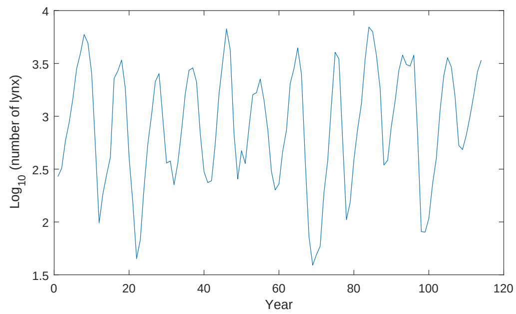

Figure 4.10: Transformed Canadian lynx time series (log10)

state. This series had been analysed in [242], and a model called pattern modelling and recognition system (PMRS) are proposed. The performance comparison using neural network (NN) and a statistical exponential-smoothing (ES) are also reported. The Santa Fe A time series is a univariate time series measured from a physical system in the laboratory. Shown in Table 4.21 are the first four samples of the Santa Fe A time series data. To aid comparison with previous studies, the experimental set-ups are arranged as closely as possible to those reported in [199, 242]. From the Santa Fe A time series, 1000 input-output data pairs are generated using the format: $[x(t-1), x(t-2), x(t-3), x(t-4), x(t-5); x(t)]$ giving five inputs and one output, $y = x(t)$. All samples are scaled to be within the range

Table 4.21: Excerpt from Santa Fe A time series data

	x1	x2	x3	x4	x5	y
1	22	41	95	141	86	21
2	21	22	41	95	141	32
3	32	21	22	41	95	72
4	72	32	21	22	41	138

$[0, 1]$ by dividing each by the maximum value of the dataset [199]. A reverse of this scaling procedure is performed before comparing with the actual output values. Similar to [199], 90% of the samples are used for training while 10% are used for testing. The learning rate is set to 0.5 with 100 training epochs. Table 4.22 shows the performance of IT2AIFLS and other models (fuzzy and non-fuzzy) on both the training and test sets. The results show that IT2AIFLS reduces the RMSE of the test set compared to the non-fuzzy and other fuzzy approaches except SVR-FM with $\epsilon = 0.001$. The reason for this could be in the large number of parameters [243] (4484 parameters) which could lead to the possible improvement in the approximation capability of SVR-FM ($\epsilon = 0.001$). The performance of IT2AIFLS on the test set of Santa-Fe time series is an indication of a good generalisation capability of the model. The Santa Fe A dataset is shown in Figure 4.11.

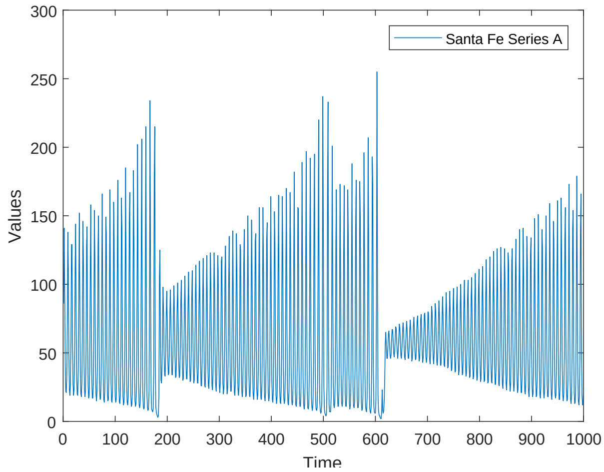

Figure 4.11: Plot of Santa Fe A time series

Table 4.22: Performance comparison of IT2AIFLS with other models on Santa Fe A time series dataset

Models	Rule	Parameter	RMSE(trn)	RMSE(tst)
ES [242]	-	-	-	56.20
NN [242]	-	-	-	24.6
PMRS [242]	-	-	-	14.23
SONFIN [202]	9	144	6.956	5.983
T2FLS-G	5	135	8.50	7.16
SEIT2FNN [121]	5	135	7.677	5.766
IT2FNN-SVR(N) [199]	5	106	13.565	4.337
IT2FNN-SVR(F) [199]	5	106	9.094	3.474
SVR-FM ($\epsilon = 0.1$) [244]	31	188	14.370	9.707
SVR-FM ($\epsilon = 0.001$) [244]	747	4484	7.069	1.650
IT2AIFLS	32	**424**	**8.355**	**2.261**

4.4.5 Complex High Dimensional Regression Problems

The effectiveness of the proposed model is demonstrated using a real world high dimensional regression datasets namely the abalone dataset. The abalone dataset is a highly noisy dataset that contains physical measurements of abalone (large edible sea snails). The dataset consists of 4177 samples with 8 input attributes. The goal is to predict the age of abalone by counting the number of rings on the abalone through a microscope [198]. Shown in Figures 4.12 and 4.13 are the inputs and measured outputs of abalone data respectively. Table 4.23 contains a few samples of the abalone data.

Table 4.23: Excerpt from abalone dataset

	sex	length	diameter	height	whole_wt	shucked_wt	viscera_wt	shell_wt	rings
1	3	0.4	0.305	0.1	0.3415	0.176	0.0625	0.0865	7
2	2	0.635	0.5	0.15	1.376	0.6495	0.361	0.31	10
3	3	0.37	0.27	0.09	0.1855	0.07	0.0425	0.065	7
4	1	0.68	0.54	0.155	1.534	0.671	0.379	0.384	10
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]

Similar to [245–248], 5-fold cross validation is adopted where the dataset is randomly split into five folds with each set containing 20% of the dataset. For each run, four folds are used for training and one for testing. Each fold is executed 5 times and the average cross validation error for 25 trials is computed. Each trial was executed for 100 epochs with learning rate set to 0.1. For the abalone dataset, 256 rules are generated while 4672 parameters are tuned. Figure 4.14 shows the actual and predicted output of abalone data using IT2AIFLS trained with GD. The result of evaluation of the abalone dataset using IT2AIFLS is compared with IT2FLS, AIFLS and similar works in the literature. As shown in Table 4.24, IT2AIFLS exhibits MSE that is lower than other models in this problem domain. The reason for this improved performance may be due to the fact that other models such as those reported in [245–248] all make use of type-1 FLSs. The proposed model also outperforms the classical IT2FLS and the AIFLS because of the additional parameters provided by the non-membership function FOU and IF-indices of IT2AIFLS. These additional parameters provide IT2AIFLS the extra design degrees of freedom with the potential to outperform type-1 FLS, AIFLS and classical IT2FLS in this problem domain.

- **House Sales in King County, USA [249]**

The house sales dataset is one of the large-scale high dimensional regression problems obtained from [249]. The purpose of this analysis is to demonstrate the prediction per-

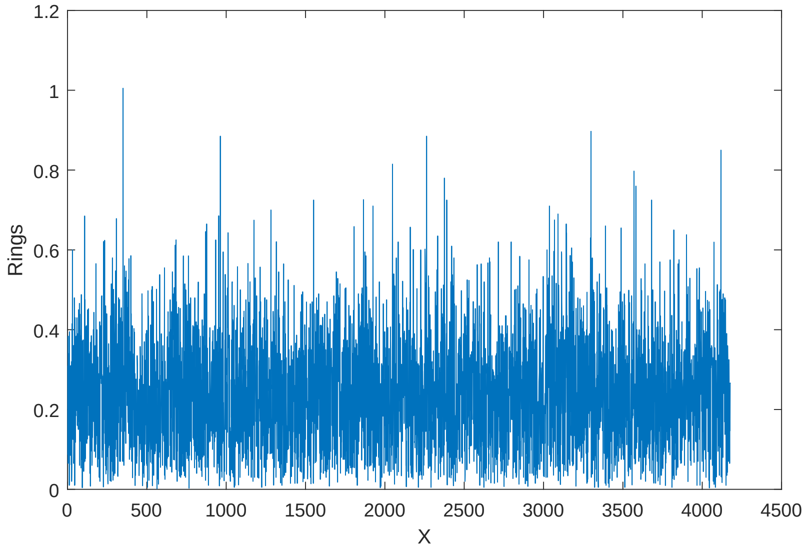

Figure 4.12: Plot of abalone data inputs set

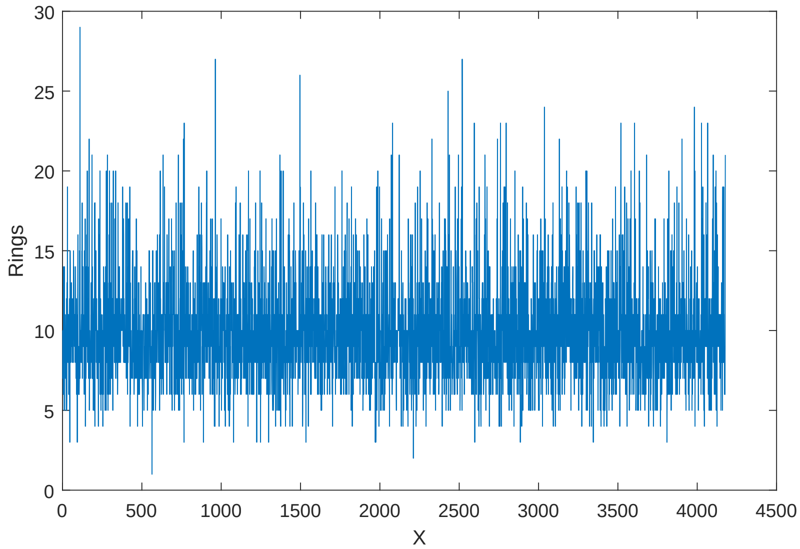

Figure 4.13: Plot of actual outputs of abalone dataset

Table 4.24: Comparison of IT2AIFLS with other models using abalone dataset

Models	MSE(tst)	MSE(std)
TS-NSGA-II [248]	2.526	0.242
TS-NSGA-SPEA2$_{Acc}$ [248]	2.511	0.263
TS-NSGA-II$_A$ [248]	2.535	0.265
TS-NSGA-II$_U$ [248]	2.520	0.237
TS-NSGA-SPEA2 [248]	2.518	0.246
TS-NSGA-SPEA2$_{Acc^2}$ [248]	2.517	0.230
Multiobjective GFS [247]	2.423	0.173
FSMOGFS [246]	2.697	0.204
FSMOGFS^e [246]	2.708	0.216
FSMOGFS+TUN [246]	2.454	0.163
FSMOGFS^e +TUN^e [246]	2.509	0.184
ANFIS-SUB [245]	2.733	-
TSK-IRL [245]	2.642	-
Linear-LMS [245]	2.472	-
LEL-TSK [245]	2.412	-
METSK-HD^e [245]	2.392	-
IT2FLS	2.798	0.045
AIFLS	2.763	0.074
IT2AIFLS	**1.042**	**0.034**

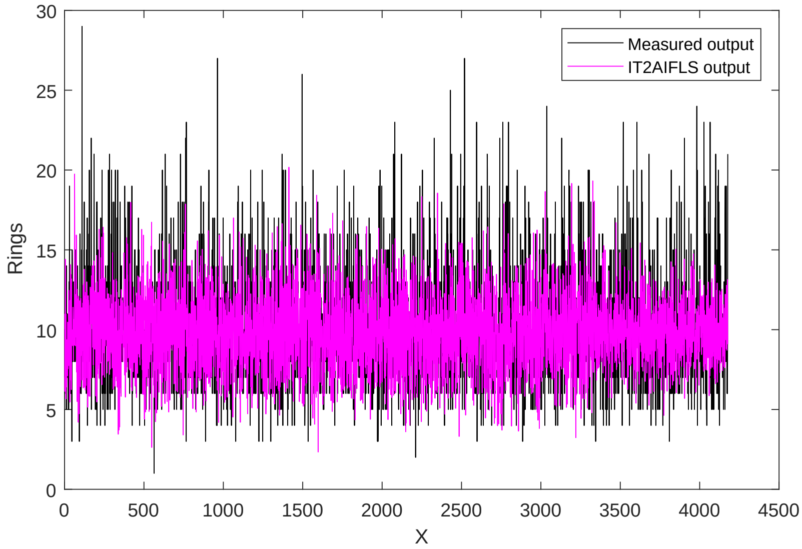

Figure 4.14: Actual and predicted outputs of abalone dataset

formance between IT2AIFLS and classical IT2FLS. The house sales dataset consists of 18 features and 21,613 samples and the task is to predict the house price as closely as possible to the actual price. Figure 4.15 shows the house sales feature ranking. All the features below the mean ranking of 0.2 are regarded as negligible and a total of 15 input features are used in the analysis in order to reduce the computational burden of the system. The entire dataset is split into 70% training and 30% testing with 10 simulation runs and 100 epochs for each run.

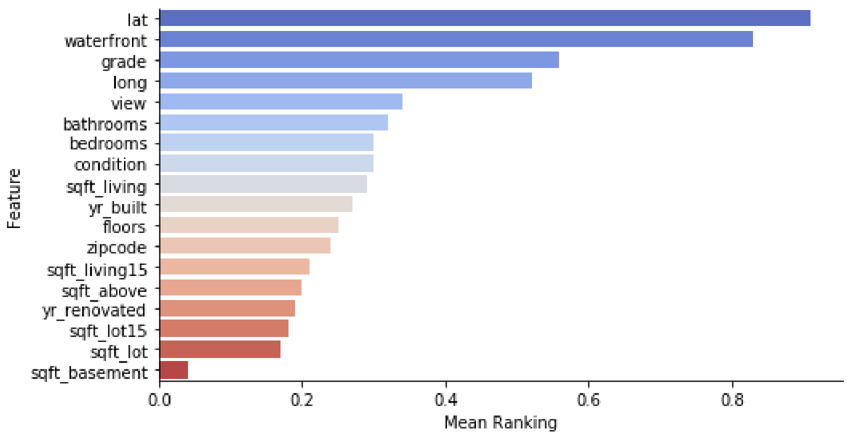

Figure 4.15: Feature ranking of house sales data [5]

Table 4.25: Comparison of IT2AIFLS with classical IT2FLS using large and high dimensional house sales data

Models	RMSE (trn)	RMSE(tst)
IT2FLS	3.2348e-05	1.5337e-05
IT2AIFLS	2.9157e-05	1.4159e-05

Table 4.25 shows the performance of IT2AIFLS and the classical IT2FLS. As shown in the table, IT2AIFLS performs better than the classical IT2FLS with reduced RMSE on this problem domain. It can be concluded that the proposed model of IT2AIFLS is a more viable method for regression problems. Thus, IT2AIFLS with fuzzy membership and non-membership functions tend to be more consistent with human or natural language description than the classical IT2FLS with only the interval membership function. Presented in Table 4.26 are the results of all the datasets analysed in this chapter. The best results are shown in bold face.

Table 4.26: Summary of results

Dataset	Measure	Proposed Model	Best other
Friedman#1	RMSE	**1.0260**	IT2FLS - 1.0950
Friedman#2 (tst1)	RMSE	**1.4940**	IT2FNN-SVR(F) - 1.5950
Friedman#2 (tst2)	RMSE	**1.1160**	IT2FNN-SVR(F) - 1.2910
Energy	RMSE	**0.5580**	IT2FLS - 0.5670
Stock	RMSE	**0.6110**	IT2FLS - 0.7500
AutoMPG6	RMSE	**1.7000**	IT2FLS - 1.7920
Low voltage line	RMSE	**255.3325**	Genetic LP - 383.4866
Maintenance cost	RMSE	**53.7200**	SA-IT2FLS - 75.2400
Mackey-Glass	RMSE	**0.0079**	RBF AFS - 0.0114
Sunspot	NMSE	**0.0105**	OSSA-LLNF - 0.0602
Tree ring	NDEI	**0.3950**	RBF AFS - 0.7140
Canadian lynx	MSE	**0.0046**	GLSSVM - 0.0056
Santa Fe	RMSE	2.2610	**SVR-FM - 1.6500**
Abalone	MSE	**1.0420**	METSK-HD^{e} - 2.3920
House sales	RMSE	**1.4159e-05**	IT2FLS - 1.5337e-05

4.5 Summary

In this chapter, a novel application of GD learning technique for the adjustment of the antecedent and consequent parameters of the proposed IT2AIFLS IF-THEN rules is investigated for the first time. The GD-based IT2AIFLS is evaluated on benchmark time series and regression problems. The performance of the proposed framework is compared with its type-1 variant, classical IT2FLS and other similar studies in the literature (fuzzy and non-fuzzy). Analysis of results reveal that with the integration of non-membership functions and IF-indices into IT2FLS, the new IT2AIFLS outperforms its type-1 variant with precise membership and non-membership functions and the classical IT2FLS with only membership functions in many applications investigated in this chapter. The performance of the IT2AIFLS is also better than most of the similar works in the literature. The IT2AIFLS accommodate more imprecision from the IF-indices and non-membership functions. Whilst the non-membership functions allow IT2AIFLS to capture more information, the IF-indices allow evaluation of concepts to be more meaningful and consistent with human reasoning and natural language representation than other representative FLSs such as classical IT2FLSs. These lead to increased level of fuzziness in IT2AIFS with increase in the prediction accuracy. In the next chapter, the effectiveness of the new framework - IT2AIFLS - is demonstrated by exploiting a second-order learning strategy.

Chapter 5

Extended Kalman Filter-based Learning of IT2AIFLS for System Identification and Time Series Predictions

> Believe in uncertainty, because by it anything is possible
>
> ---
>
> Imo Eyoh

5.1 Introduction

In Chapter 4, the parameters of the IT2AIFLS are optimised using a first order derivative based method - GD. In this chapter, the parameters of the proposed framework is tuned using decoupled extended Kalman filter (DEKF), a second-order derivative based method. The resulting system is evaluated using one synthetic dataset and two real world datasets. To aid comparison with alternative approaches, the classical IT2FLS and AIFLS are also implemented. Statistical comparison between the pairs of FLS models investigated here is conducted.

5.1.1 Rules

The generic IT2AIFLS IF-THEN rule structure is rewritten in Equation 5.1 for ease of reference:

$$R_k : \text{IF } x_1 \text{ is } \tilde{A}^*{}_{1k} \text{ and } x_2 \text{ is } \tilde{A}^*{}_{2k} \text{ and } \cdots \text{ and } x_n \text{ is } \tilde{A}^*{}_{nk} \text{THEN } y_k = \sum_{i=1}^{n} w_{ik}x_i + b_k \tag{5.1}$$

where $\tilde{A}^*{}_{1k}$, $\tilde{A}^*{}_{2k}$, $\cdots$, $\tilde{A}^*{}_{ik}$, $\cdots$, $\tilde{A}^*{}_{nk}$ are IT2AIFS applied to the kth rule and y_k is the output, w_{ik}'s and b_k's ($k = 1 \cdots M$) are the consequent parameters.

5.1.2 Inference

The inference mechanism for IT2AIFLS is expressed in Equation 5.2 [2, 46, 47].

$$y = \frac{(1-\beta)\sum_{k=1}^{M}\left(\underline{f_k^{\mu}} + \overline{f_k^{\mu}}\right) y_k^{\mu}}{\sum_{k=1}^{M}\underline{f_k^{\mu}} + \sum_{k=1}^{M}\overline{f_k^{\mu}}} + \frac{\beta\sum_{k=1}^{M}\left(\underline{f_k^{\nu}} + \overline{f_k^{\nu}}\right) y_k^{\nu}}{\sum_{k=1}^{M}\underline{f_k^{\nu}} + \sum_{k=1}^{M}\overline{f_k^{\nu}}} \tag{5.2}$$

where $\underline{f}_k^{\mu}$, $\overline{f}_k^{\mu}$ and $\underline{f}_k^{\nu}$, $\overline{f}_k^{\nu}$ are the lower, upper membership and the lower, upper non-membership firing strength respectively, y_k^{μ} and y_k^{ν} are the corresponding outputs of the k_{th} rule. The design parameter β is defined in Equation 5.3 such that, $0 \leq \beta \leq 1$.

$$y = \begin{cases} MF\ \ only & \text{if } \beta = 0 \\ MF\ \ and\ \ NMF & \text{if } 0 < \beta < 1 \\ NMF\ \ only, & \text{if } \beta = 1 \end{cases} \tag{5.3}$$

Hence, the parameter β in the unit interval [0,1] determines the magnitude of membership and non-membership functions in the final output. The "prod" t-norm is used as the implication operator and are defined for membership function, Equations 5.4 and 5.5 and non-membership function, Equations 5.6 and 5.7 as follows:

$$\underline{f_k^{\mu}}(x) = \underline{\mu}_{\tilde{A}^*{}_{1k}}(x_1) * \underline{\mu}_{\tilde{A}^*{}_{2k}}(x_2) * \cdots * \underline{\mu}_{\tilde{A}^*{}_{nk}}(x_n) \tag{5.4}$$

$$\overline{f_k^{\mu}}(x) = \overline{\mu}_{\tilde{A}^*{}_{1k}}(x_1) * \overline{\mu}_{\tilde{A}^*{}_{2k}}(x_2) * \cdots * \overline{\mu}_{\tilde{A}^*{}_{nk}}(x_n) \tag{5.5}$$

$$\underline{f_k^{\nu}}(x) = \underline{\nu}_{\tilde{A}^*{}_{1k}}(x_1) * \underline{\nu}_{\tilde{A}^*{}_{2k}}(x_2) * \cdots * \underline{\nu}_{\tilde{A}^*{}_{nk}}(x_n) \tag{5.6}$$

$$\overline{f_k^{\nu}}(x) = \overline{\nu}_{\tilde{A}^*{}_{1k}}(x_1) * \overline{\nu}_{\tilde{A}^*{}_{2k}}(x_2) * \cdots * \overline{\nu}_{\tilde{A}^*{}_{nk}}(x_n) \tag{5.7}$$

where $*$ is the "prod" operator.

5.2 Parameter Updates

In Chapter 4, the GD is used to optimise the parameters of the proposed model. Whilst GD is guaranteed to reach a minimum [119], they are known disadvantages [117]. Particularly, GD with a learning rate parameter may lead to slow convergence and the possibility of getting trapped in local minima. Exploiting second-order derivative-based method such as the EKF-based methods for the parameter update of the T2FLSs may help to speed up the convergence with smaller possibility of getting stuck in local minima [42, 125]. The authors in [125] pointed out that Kalman filter-based approaches can be a *powerful tool* for the optimisation of T2FLSs. Hence, in this section, the antecedent and consequent parameter updates for IT2AIFLS using EKF-based approach are exploited.

5.2.1 Extended Kalman Filter Parameter Update Rule

The basic idea behind the IT2AIFLS prediction method is to approximate the relationship between inputs and outputs of a system as closely as possible. Assuming that the IT2AIFLS model is trained by adjusting the parameters using sets of input-output pairs, then the output of a fuzzy logic system may be represented as $y = f(X, \theta)$. The parameter X denotes the inputs into the system with θ representing the unknown parameters of the model. For IT2AIFLS, these will include both the membership and non-membership functions parameters. The generic non-linear dynamic state equation can be expressed as:

$$\theta_{t+1} = f(\theta_t) + \omega_t \tag{5.8}$$

$$y_t = h(\theta_t) + \upsilon_t \tag{5.9}$$

where θ is the system's state, ω is the process noise with zero mean and covariance Q while υ is the measurement noise with zero mean and covariance R. For Kalman filter, the process and measurement noise are assumed to be Gaussian and uncorrelated and:

$$E(\theta_0) = \overline{\theta}_0 \tag{5.10}$$

$$E[(\theta_0 - \overline{\theta}_0)(\theta_0 - \overline{\theta}_0)^T] = P_0 \tag{5.11}$$

$$E(\omega_t) = 0 \tag{5.12}$$

$$E(\omega_t \omega_l^T) = Q\delta_{tl} \tag{5.13}$$

$$E(\upsilon_t) = 0 \tag{5.14}$$

$$E(\upsilon_t \upsilon_l^T) = R\delta_{tl} \tag{5.15}$$

where $E(.)$ is the expectation operator and δ_{tl} is the Kronecker delta. The state can be estimated using Taylor expansion as:

$$\begin{aligned} f(\theta_t) &= f(\hat{\theta}_t) + F_t(\theta_t - \hat{\theta}_t) + H.O.T \\ h(\theta_t) &= h(\hat{\theta}_t) + H_t(\theta_t - \hat{\theta}_t) + H.O.T \end{aligned} \tag{5.16}$$

where:

$$F_t = \left.\frac{\partial f(\theta)}{\partial \theta}\right|_{\theta=\hat{\theta}_t} \qquad \text{and} \qquad H_t^T = \left.\frac{\partial h(\theta)}{\partial \theta}\right|_{\theta=\hat{\theta}_t}$$

and H.O.T is the higher order term. The system in Eqn (5.16) can be approximated as in Eqn (5.17) when the higher order terms are neglected.

$$\begin{aligned} \theta_{t+1} &= F_t\theta_t + \omega_t + \phi_t \\ y_{t+1} &= H_t^T\theta_t + \upsilon_t + \varphi_t \end{aligned} \tag{5.17}$$

where ϕ_t and φ_t are random error terms for state and observation equations respectively and expressed as:

$$\begin{aligned} \phi_t &= f(\hat{\theta}_t) - F_t\hat{\theta}_t \\ \varphi_t &= h(\hat{\theta}_t) - H_t\hat{\theta}_t \end{aligned} \tag{5.18}$$

The desired estimation of the parameters in Equation 5.17 can therefore be obtained using the recursive Kalman procedures in Equation 5.19 to 5.21 [115, 250, 251].

$$K_t = P_tH_t[(H_t)^TP_tH_t + R]^{-1} \tag{5.19}$$

$$\hat{\theta}_t = f(\hat{\theta}_{t-1}) + K_t[y_t - h(\hat{\theta}_{t-1})] \tag{5.20}$$

$$P_{t+1} = F_t(P_t - K_tP_t(H_t)^T)F_t^T + Q \tag{5.21}$$

The vector F_t is taken as an identity matrix (I) and Equation 5.22 to 5.24 are obtained [125].

$$K_t = P_tH_t[(H_t)^TP_tH_t + R]^{-1} \tag{5.22}$$

$$\hat{\theta}_t = \hat{\theta}_{t-1} + K_t[y_t - h(\hat{\theta}_{t-1})] \tag{5.23}$$

$$P_{t+1} = P_t - K_tP_t(H_t)^T + Q \tag{5.24}$$

where K is the Kalman gain and P is the covariance matrix of the state estimation error.

In applying the EKF to IT2AIFLS, all the unknown parameters are gathered in a single vector. The computational cost of EKF is in the order of DB^2 where D is the dimension of the output of the system and B is the total number of the parameters. Thus, for an IT2AIFLS with n inputs, M number of rules and a single output, the total number of parameters to be tuned is $8n + 2M(n+1)$. The computational cost of the standard EKF for IT2AIFLS is therefore $64n^2 + 4M^2(n^2 + 2n + 1) + 32nM(n+1)$ which is a very large number in most applications. To reduce the computational cost of EKF, this research adopt the decoupled EKF.

5.2.2 Decoupled Extended Kalman Filter

As discussed in Subsection 2.6.2, using the standard EKF is computationally burdensome because of the high dimensionality of the parameters. In order to reduce the computational burden, a simplified version of the EKF called the decoupled extended Kalman filter (DEKF) proposed in [124, 252] is used as suggested in [115] and being a second order derivative-based method, convergence is expected to be faster [115, 253]. The assumption for DEKF is that the intra-correlation among parameters of the model is high while the inter-correlation is low [253]. Hence, by decoupling the parameters and ignoring these inter-correlation [253], parameter interactions are made to occur only at the second-order level [125]. Thus, instead of having one large vector of parameters, smaller groups (vectors) of parameters are utilised with small interactions between groups, thereby increasing the computational efficiency of the DEKF.

By using the DEKF to learn the parameters of IT2AIFLS, the antecedent and the consequent parameters are grouped into two vectors - one for the antecedent and the other for the consequent parameters.

5.3 Antecedent Update Rule

In the antecedent, the state space is partitioned into sets of intuitionistic fuzzy regions which determines the set of rules generated from a piece of training data. As discussed in Section 2.6, the design of a FLS includes the determination of the unknown parameters of the FLS in the antecedent parts. During the antecedent parameter update, all the unknown parameters in the antecedent parts of IT2AIFLS are gathered into a single vector and represented as:

$$\theta^1 = [c_{11}, c_{21}, \cdots, c_{nm}, \sigma_{11}, \sigma_{21}, \cdots, \sigma_{nm}]^T \tag{5.25}$$

where n is the number of inputs and m is the number of interval type-2 intuitionistic fuzzy partitions.

The Equation 5.25 is further decomposed into membership function antecedent parameters as in Equation 5.26

$$\theta^{1\mu} = [c_{11}, c_{12}, \cdots, c_{nm}, \sigma^{\mu}_{11}, \sigma^{\mu}_{12}, \cdots, \sigma_{nm}]^T \tag{5.26}$$

and non-membership antecedent parameters as in Equation 5.27

$$\theta^{1\nu} = [c_{11}, c_{12}, \cdots, c_{nm}, \sigma^{\nu}_{11}, \sigma^{\nu}_{12}, \cdots, \sigma_{nm}]^T \tag{5.27}$$

where c is the center and is the same for both membership functions and non-membership functions of the IT2AIFLS, $\sigma^{\mu}_1 = \sigma^{\nu}_1$ and $\sigma^{\mu}_2 = \sigma^{\nu}_2$. The generic parameter update rule in the i_{th} group is as in Equation 5.28 to 5.30:

$$\theta^i_t = \theta^i_{t-1} + K^i_t[y_t - h(\theta_{t-1})] \tag{5.28}$$

$$K^i_t = P^i_t H^i_t [(H^i_t)^T P^i_t H^i_t + R^i]^{-1} \tag{5.29}$$

$$P^i_{t+1} = P^i_t - K^i_t P^i_t (H^i_t)^T + Q^i \tag{5.30}$$

For the IT2AIFLS, the unknown parameters in the antecedent are gathered into the first vector and represented as:

$$\theta^1 = [c_{11}, c_{21}, \cdots, c_{nm}, \sigma_{11}, \sigma_{21}, \cdots, \sigma_{nm}]^T \tag{5.31}$$

The Equation 5.31 is further decomposed into membership function antecedent parameters in Equation 5.32

$$\theta^{1\mu} = [c_{11}, c_{12}, \cdots, c_{nm}, \sigma^{\mu}_{11}, \sigma^{\mu}_{12}, \cdots, \sigma_{nm}]^T \tag{5.32}$$

and non-membership antecedent parameters in Equation 5.33

$$\theta^{1\nu} = [c_{11}, c_{12}, \cdots, c_{nm}, \sigma^{\nu}_{11}, \sigma^{\nu}_{12}, \cdots, \sigma_{nm}]^T \tag{5.33}$$

The derivative matrix, H is defined in Equation 5.34 for membership function,

$$H^{\mu} = \frac{\partial y}{\partial \theta^{\mu}} \tag{5.34}$$

and Equation 5.35 for non membership function.

$$H^{\nu} = \frac{\partial y}{\partial \theta^{\nu}} \tag{5.35}$$

The update rule for the parameters in θ^1 then follow the Kalman filtering recursive procedures as in Equation 5.28 to 5.30 with membership and non-membership functions having separate Kalman filter parameters as shown in Equation 5.36 to 5.41:

$$K_t^\mu = P_t^\mu H_t^\mu [(H_t^\mu)^T P_t^\mu H_t^\mu + R^\mu]^{-1} \tag{5.36}$$

$$\hat{\theta}_t^\mu = \hat{\theta}_{t-1}^\mu + K_t^\mu [y_t - h(\hat{\theta}_{t-1})] \tag{5.37}$$

$$P_{t+1}^\mu = P_t^\mu - K_t^\mu P_t^\mu (H_t^\mu)^T + Q^\mu \tag{5.38}$$

$$K_t^\nu = P_t^\nu H_t^\nu [(H_t^\nu)^T P_t^\nu H_t^\nu + R^\nu]^{-1} \tag{5.39}$$

$$\hat{\theta}_t^\nu = \hat{\theta}_{t-1}^\nu + K_t^\nu [y_t - h(\hat{\theta}_{t-1})] \tag{5.40}$$

$$P_{t+1}^\nu = P_t^\nu - K_t^\nu P_t^\nu (H_t^\nu)^T + Q^\nu \tag{5.41}$$

With the DEKF, the reduction in the computational cost is in the order $64n^2 + 4M^2(n^2 + 2n + 1)$ and the computational complexity of DEKF to EKF is in the ratio:

$$\frac{64n^2 + 4M^2(n^2 + 2n + 1)}{64n^2 + 4M^2(n^2 + 2n + 1) + 32nM(n + 1)}$$

This is a significant improvement compared to the standard EKF for training IT2AIFLS. The DEKF therefore has an advantage over the conventional EKF in terms of resource utilisation and coupled with the complexity of IT2AIFLS, DEKF becomes the preferred learning approach in this research.

5.4 Consequent Parameter Update

The parameters of the consequent are grouped into the second vector and represented as:

$$\theta^2 = [w_{11}, w_{21}, \cdots, w_{Mn}, b_1, b_2, \cdots, b_M]^T \tag{5.42}$$

where M is the number of rules, The Equation 5.42 is also decomposed into Equations 5.43 and 5.44 for membership and non-membership consequent parameters respectively

$$\theta^{2\mu} = [w_{11}^\mu, w_{12}^\mu, \cdots, w_{Mn}^\mu, b_1^\mu, b_2^\mu, \cdots, b_M^\mu]^T \tag{5.43}$$

$$\theta^{2\nu} = [w_{11}^\nu, w_{12}^\nu, \cdots, w_{Mn}^\nu, b_1^\nu, b_2^\nu, \cdots, b_M^\nu]^T \tag{5.44}$$

with the membership and non-membership functions having separate Kalman filter parameters. The derivative matrix, H, is defined as:

$$H^{\mu} = \frac{\partial y}{\partial \theta^1} \qquad \text{and} \qquad H^{\nu} = \frac{\partial y}{\partial \theta^2} \tag{5.45}$$

$$\tag{5.46}$$

for membership and non-membership function parameters respectively.

The update rule for the parameters in θ^2 then follow the same recursive procedures as in Equation 5.28 to 5.30

5.5 Experiments and Results

In this section, the evaluation of the proposed learning algorithm of IT2AIFLS is conducted using one synthetic and two real world datasets namely Australia's New South Wales (NSW) electricity price data in the year 2008 and a gas compression system (GCS) dataset obtained from a Nigerian-based power plant. The performance metrics utilised in this chapter are the RMSE and MAE which are rewritten here for ease of reference in Equations 5.47 and 5.48 respectively:

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y^a - y)^2} \tag{5.47}$$

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|y^a - y| \tag{5.48}$$

where N is the number of test samples, y^a and y are the actual and predicted outputs respectively.

5.5.1 System Identification

The proposed IT2AIFLS-DEKF model is applied to a dynamic system dataset generated using the differential equation expressed in Equation 6.20 [121]:

$$y(t+1) = \frac{y(t)}{1 + y^2(t)} + u^3(t) + f(t)$$

where

$$f(t) = \begin{cases} 0, & 1 \leq t \leq 1000 \\ 1.0, & 1001 \leq t \leq 2000 \\ 0, & 2001 \leq t \end{cases} \tag{5.49}$$

The inputs to the proposed model are $u(t)$ and $y(t)$ while $y(t+1)$ is the desired output. Similar to Juang *et al.* [121], the 2001 training data samples are generated using $u(t) = sin(2\pi t/100)$. There are 4 rules and 40 tunable parameters for the IT2AIFLS. A TSK type-1 AIFLS and an IT2FLS trained with DEKF are also constructed and evaluated on the system identification problem. The number of rules in the three models remain the same with 36 and 24 tunable parameters for the AIFLS and IT2FLS respectively. The RMSE is adopted as the performance metric. The RMSE is computed over 30 simulations for each model. Shown in Figure 5.1 is the actual and predicted outputs of the identification problem using IT2AIFLS. As presented in Table 5.1, IT2AIFLS outperforms both AIFLS and IT2FLS in this problem instance.

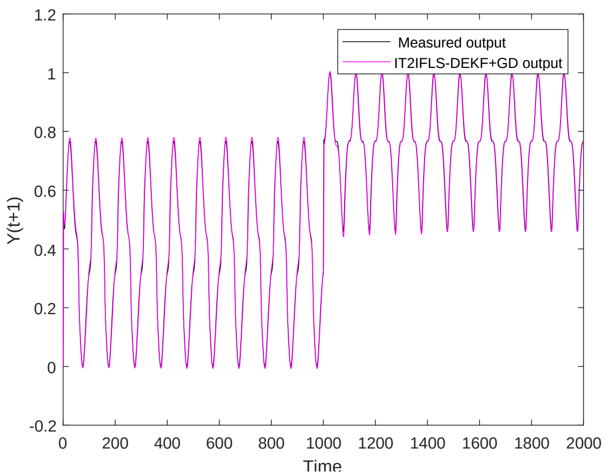

Figure 5.1: Actual and predicted output using IT2AIFLS for identification problem

Table 5.1: Comparison of IT2AIFLS vs AIFLS and IT2FLS on second-order identification problem

Model	Rules	RMSE(trn)	RMSE(tst)
IT2FLS	4	0.0155	0.0082
AIFLS	4	0.0155	0.0079
IT2AIFLS	4	0.0164	0.0068

5.5.2 NSW Electricity Load Forecast

The proposed EKF-based learning IT2AIFLS model is evaluated using a real world datasets from the Australia's National Electricity Market (NEM) namely New South Wales (NSW) electricity market. Similar to [254], the NSW electricity market for the year 2008 is used for the analysis. The dataset is downloaded from [255] and consists of 17568 instances with attributes of regional reference price as the input. The price data are treated as time series data and are partitioned into four separate datasets according to [254] as representatives of the four seasons in Australia. The input data for analysis is generated from four previous values $[x(t-4), x(t-3), x(t-2), x(t-1)]$ with $x(t+1)$ as the output. There are a total of 336 data samples for each season which reduces to 331 after input generation. The first 231 data points are used for training while the remaining 100 data samples are used for testing in each season. There are 16 rules generated with a total of 8(4) + 2*16(4+1) = 192 parameters. The performance metrics employed are the RMSE and MAE. The data for the analysis have been normalised to a small range of [0,1]. The partitioning of the dataset for each season are shown in Table 5.2. The performance of

Table 5.2: NSW 2008 electricity price dataset partitions

Period	Input	Total datapoint	Train datapoint	Test datapoint
Summer 24 - 30/01/08	4	331	231	100
Autumn 24 - 30/05/08	4	331	231	100
Winter 24 - 30/08/08	4	331	231	100
Spring 24 - 30/10/08	4	331	231	100

the new learning algorithm of IT2AIFLS-DEKF using NSW electricity data is evaluated on two fronts namely:

- performance comparison with another learning algorithm such as the GD and
- performance comparison with other fuzzy models trained with DEKF such as AIFLS and classical IT2FLS.

The performance of each of the training algorithms was computed over 30 simulations. Figure 5.2 shows the actual and the predicted outputs of IT2AIFLS-DEKF and IT2AIFLS-GD with the corresponding prediction errors for the different seasons. As shown in Tables 5.3 to 5.6, IT2AIFLS-DEKF exhibits superior performance over IT2AIFLS-GD.

Table 5.3: Performance of different models and algorithms during Summer season

Period		Summer			
Model		AIFLS-DEKF	IT2FLS-DEKF	IT2AIFLS-GD	IT2AIFLS-DEKF
RMSE	Trn	0.0229	0.0243	0.0243	0.0225
	Tst	0.1112	0.2284	0.1599	0.0979
MAE		0.0315	0.0683	0.0502	0.0284

Table 5.4: Performance of different models and algorithms during Autumn season

Period		Autumn			
Model		AIFLS-DEKF	IT2FLS-DEKF	IT2AIFLS-GD	IT2AIFLS-DEKF
RMSE	Trn	0.0889	0.0891	0.0896	0.0871
	Tst	0.0407	0.0410	0.0789	0.0409
MAE		0.0164	0.0161	0.0393	0.0167

Table 5.5: Performance of different models and algorithms during Winter season

Period		Winter			
Model		AIFLS-DEKF	IT2FLS-DEKF	IT2AIFLS-GD	IT2AIFLS-DEKF
RMSE	Trn	0.0836	0.0846	0.0916	0.0791
	Tst	0.0439	0.0429	0.0553	0.0422
MAE		0.0182	0.0184	0.0239	0.0175

Table 5.6: Performance of different models and algorithms during Spring season

Period		Spring			
Model		AIFLS-DEKF	IT2FLS-DEKF	IT2AIFLS-GD	IT2AIFLS-DEKF
RMSE	Trn	0.0715	0.0754	0.0802	0.0723
	Tst	0.0960	0.0954	0.1333	0.0821
MAE		0.0342	0.0367	0.0477	0.0335

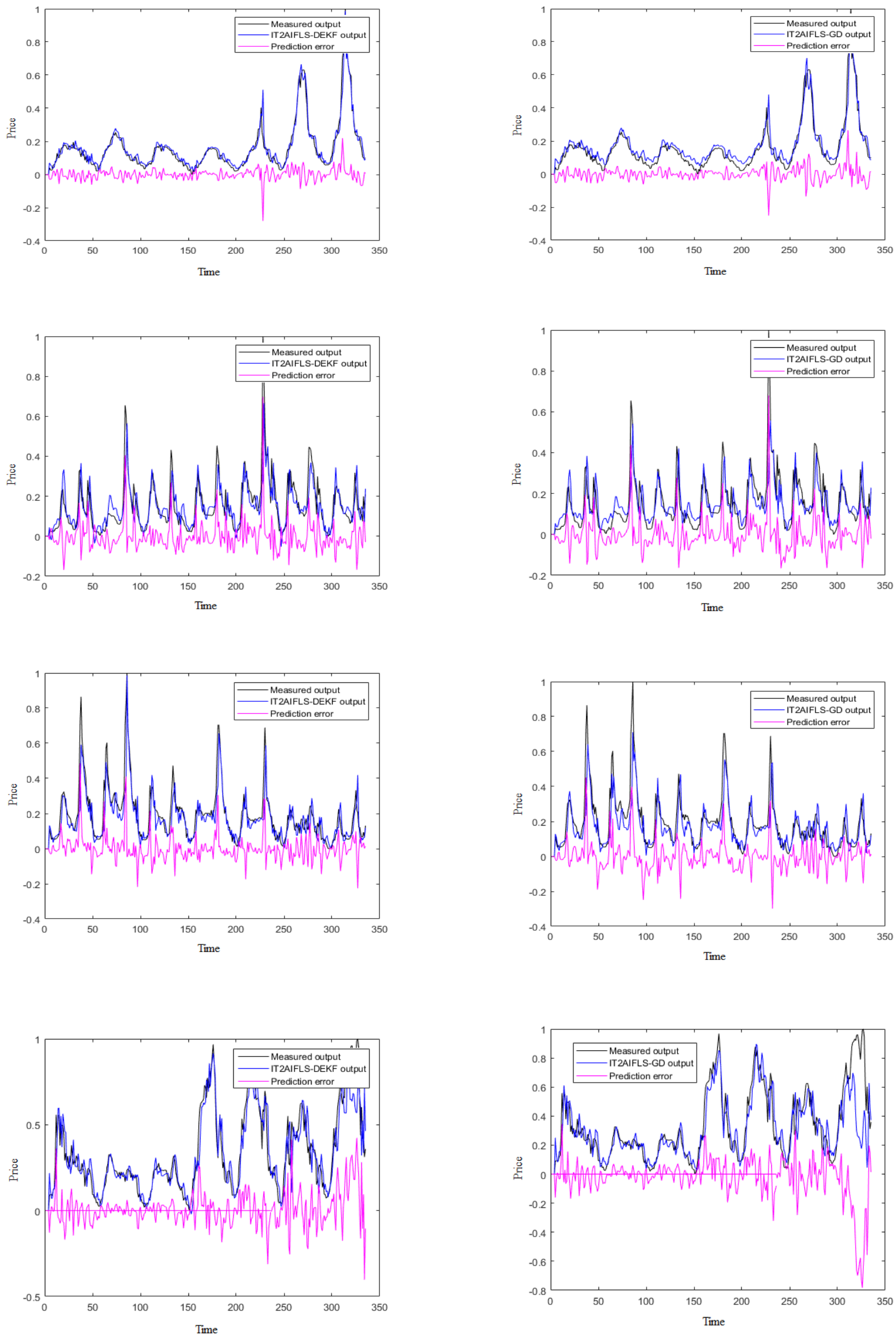

Figure 5.2: Price prediction in summer, autumn, winter and spring using IT2AIFLS-DEKF and IT2AIFLS-GD respectively

It is conjectured that this could be as a result of the EKF-based algorithm's ability to overcome local minima problems and to account for interdependence between outputs at each iterations. In Table 5.3, the DEKF-based AIFLS, IT2FLS and IT2AIFLS for autumn season are very close in their modelling error. The AIFLS performs slightly better than the type-2 models for the autumn season. This shows that a type-1 FLS can model uncertainty and non-linearity to some degree [125, 256]. In the overall, Tables 5.3 to 5.6 show that IT2AIFLS performs better than both AIFLS and IT2FLS trained with the same DEKF algorithm with reduced RMSE and MAE. Hence, using IT2AIFLS can be a preferred option for handling uncertainty in many real world applications.

5.5.3 Gas Compression System Time Series Prediction

In this subsection, the IT2AIFLS is used for the future prediction of another real-world dataset - gas compression system (GCS) dataset of a gas turbine obtained from a Nigerian-based power plant. The GCS data is a complex dataset consisting of different operational conditions of a gas plant. There are a total of 825 data points. The purpose of this simulation is to statistically analyse the performance of IT2AIFLS, IT2FLS and AIFLS. The DEKF learning approach is adopted for this experimental analysis because of its theoretical strength, faster convergence and its ability at finding good solutions [115]. The GCS data is modeled as a time series dataset using input generating format: $[x(t-3), x(t-2, x(t-1)]$ with $x(t)$ as the output. The input are normalised to lie between small range of [0,1], so that larger input values do not overshadow the smaller values, thereby leading to poor prediction and learning with the embedded ANN architecture. For each run of the experiments, the data are randomly sampled and split into 70% training and 30% testing set. In this approach, each data point has equal probability of being sampled for training and testing in the simulation runs. For a clear and objective discussion and evaluation of the three models of IT2IFLS, IT2FLS and AIFLS, the Kalman filter parameters R, Q and P for both membership functions and non-membership functions are initially set as 40, $0.01I_{32}$ and $1.0I_{32}$ respectively for all experiments with 100 epochs for each run. The performance metric adopted for this analysis is the RMSE. The simulation is conducted for 30 runs. This allows for objective evaluation of the performance of the different models under investigation. The test RMSEs averaged over 30 runs for the different fuzzy logic models considered here are presented in Table 5.7

As shown in the box-and-whisker plots in Figure 5.7, IT2AIFLS has the smallest error value on average. This observation points to the merits of non-membership and IF-indices as integral parts of IT2FLS. Shown in Figures 5.3 and 5.4 are the membership and non-membership functions of a single input attribute of GCS data before and after training

Table 5.7: Performance comparison of IT2FLS, AIFLS and IT2AIFLS using GCS dataset

Models	RMSE(trn)	RMSE(tst)
IT2FLS	0.1504	0.1425
AIFLS	0.1496	0.1423
IT2AIFLS	0.1202	0.1199

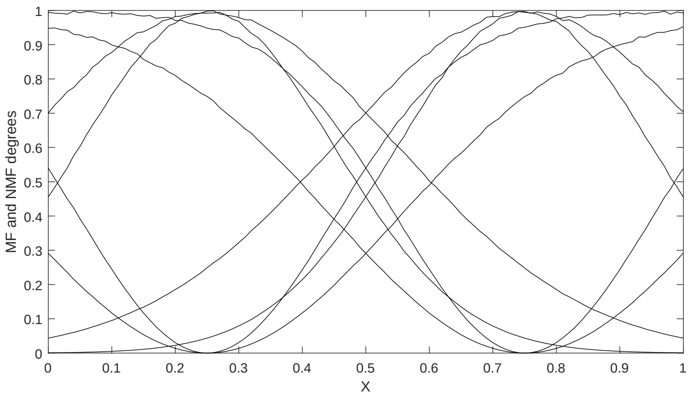

Figure 5.3: GCS membership and non-membership functions before training with IT2AIFLS

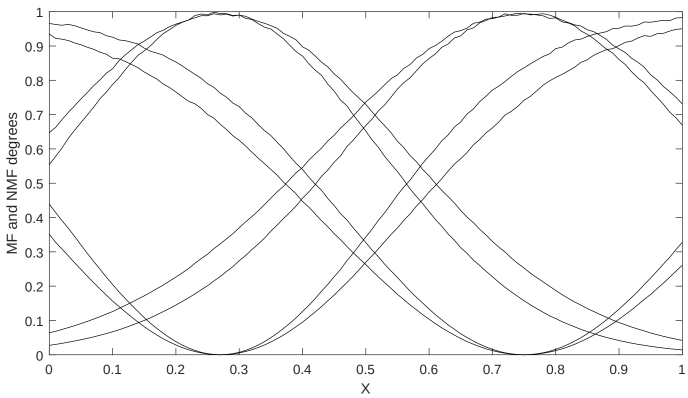

Figure 5.4: GCS membership and non-membership functions after training with IT2AIFLS

using IT2AIFLS. Figures 5.5 and 5.6 show the membership functions before and after training with classical IT2FLS. As depicted in Figure 5.4, IT2AIFLS is able to minimise the effects of membership and non-membership functions uncertainties as shown on their reduced FOU sizes.

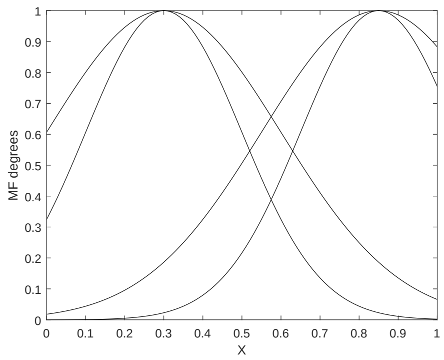

Figure 5.5: GCS membership function before training with IT2FLS

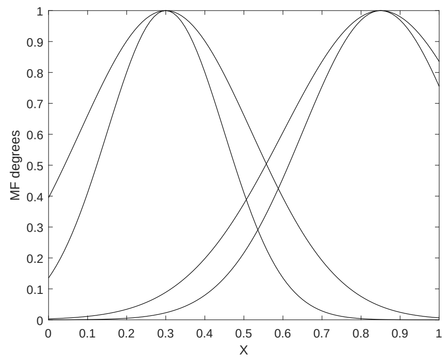

Figure 5.6: GCS membership function after training with IT2FLS

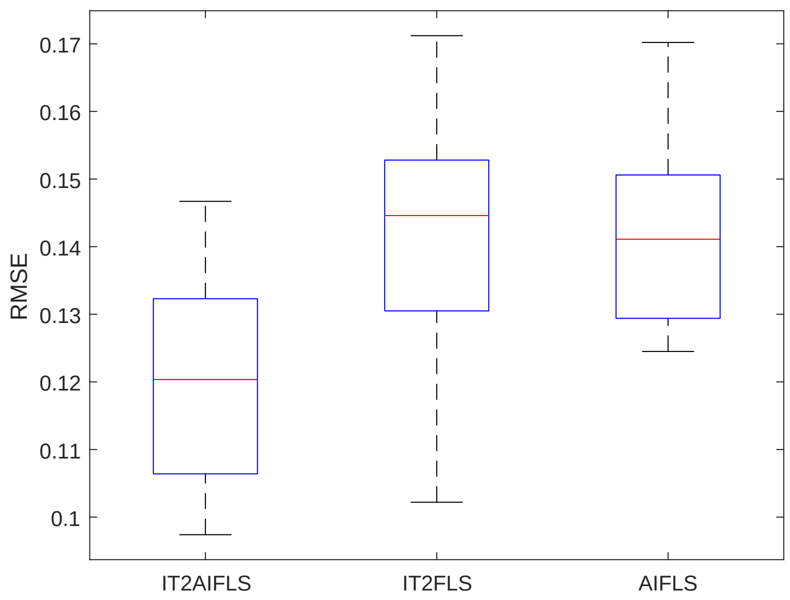

Figure 5.7: Box-and-whisker plot showing the performance of IT2AIFLS, IT2FLS and AIFLS.

5.6 Statistical Evaluation

In this section, statistical evaluation is conducted to test the hypothesis of this research. The main interest is to understand the effectiveness of integrating fuzzy non-membership function and IF-indices into IT2FLSs where more uncertainty is captured in the fuzzy set description. The second is to investigate the performance of the proposed framework of IT2AIFLS with its type-1 counterpart. Statistical comparison is also made between AIFLS and classical IT2FLS. To explore these, three experiments are conducted. The following hypotheses form the basis of evaluation:

- Hypothesis 1: With the integration of non-membership functions and IF-indices into the classical IT2FLS, the new model of IT2AIFLS is able to model uncertainty in many applications than the classical IT2FLS that do not incorporate non-membership and IF-indices.

- Hypothesis 2: With membership and non-membership functions that are intervals, the new model of IT2AIFLS is able to model uncertainty in many applications than its type-1 variant with membership and non-membership functions that are not intervals.

- Hypothesis 3: With membership and non-membership functions of AIFLS, the model is able to model uncertainty in many applications than the classical IT2FLS with lower and upper membership functions.

Statistical significance of differences between pairs of models are carried out using Wilcoxon signed rank test (α level = 0.05). The Wilcoxon signed rank test is one of the most commonly used non-parametric statistical hypothesis test for evaluating the predictive capabilities between pairs of models to determine whether there is existence of statistical differences among results [231].

Hypothesis 1: The first set of experiments is focused on assessing the ability of IT2AIFLS framework to provide good estimates than the classical IT2FLS. The null and alternative hypotheses are:

- H_0: There is no significant difference (one-tailed) in the performance of the IT2FLS that incorporates non-membership and IF-indices and those that do not.

- H_1: There is a significant difference (one-tailed) in the performance of IT2FLS that incorporates non-membership and IF-indices and those that do not.

Table 5.8: Wilcoxon's test: IT2AIFLS and IT2FLS using test RMSE

Model	Hypothesis ($\alpha = 0.05$)	p-value
IT2AIFLS vs IT2FLS	Reject H_0	0.0173

For the first hypothesis, the statistical analysis suggests that there is a highly significant difference in the performances of the two approaches (p-value = 0.0173). This leads to the rejection of the null hypothesis. It is concluded that there is a significant difference in the performances of IT2AIFLS and classical IT2FLS.

Hypothesis 2: The second set of experiments is focused on assessing the ability of IT2AIFLS framework to provide good estimates than its type-1 counterpart. The null and alternative hypotheses are:

- H_0: There is no significant difference (one-tailed) in the performance of IT2AIFLS with membership and non-membership functions that are intervals and AIFLS with membership and non-membership functions that are not intervals.

- H_1: There is a significant difference (one-tailed) in the performance of IT2AIFLS with membership and non-membership functions that are intervals and AIFLS with membership and non-membership functions that are not intervals.

Table 5.9: Wilcoxon's test: IT2AIFLS and AIFLS using test RMSE

Model	Hypothesis ($\alpha = 0.05$)	p-value
IT2AIFLS vs AIFLS	Reject H_0	0.0091

For the second hypothesis, the statistical analysis suggests that there is a highly significant difference in the performances of IT2AIFLS and AIFLS (p-value = 0.0091). This leads to the rejection of the null hypothesis. It is concluded that there is a significant difference in the performances of IT2AIFLS and AIFLS. This shows that membership and non-membership functions that are intervals may be more appropriate for uncertainty modelling than those with membership and non-membership functions that are not intervals.

Hypothesis 3: The third set of experiments is to investigate the statistical significance between IT2FLS and AIFLS. The null and alternative hypotheses are:

- H_0: There is no significant difference (one-tailed) in the performance of IT2FLS utilising upper and lower membership functions of IT2FS and AIFLS utilising membership and non-membership functions of AIFS.

- H_1: There is a significant difference (one-tailed) in the performance of IT2FLS utilising upper and lower membership functions of IT2FS and AIFLS utilising only membership and non-membership functions of AIFS.

Table 5.10: Wilcoxon's test: IT2FLS and AIFLS using test RMSE

Model	Hypothesis ($\alpha = 0.05$)	p-value
IT2FLS vs AIFLS	Fail to reject H_0	0.7336

Table 5.10 shows the results of statistical comparison between classical IT2FLS and type-1 AIFLS. The Wilcoxon's signed rank test at 0.05 significance level shows that there is no significant difference (p-value = 0.7336) existing between IT2FLS and AIFLS, hence a failure to reject the null hypothesis. It can be concluded that there is no significant difference (one-tailed) in the performance of IT2FLS utilising upper and lower membership functions of IT2FS and AIFLS utilising membership and non-membership functions of AIFS.

Table 5.11 summarises the results on the test datasets presented in this chapter. The best results are indicated in bold.

Table 5.11: Summary of results

Dataset	Measure	Proposed Model	Best other
System identification	RMSE	**0.0068**	AIFLS - 0.0079
Poland electricity(Summer)	RMSE	**0.0979**	AIFLS - 0.1112
Poland electricity(Autumn)	RMSE	0.0409	**AIFLS - 0.0407**
Poland electricity(Winter)	RMSE	**0.0422**	IT2FLS - 0.0429
Poland electricity(Spring)	RMSE	**0.0821**	IT2FLS - 0.0954
GCS	RMSE	**0.1199**	AIFLS - 0.1423

5.7 Summary

In this chapter, a new application of DEKF learning algorithm for IT2AIFLS is proposed and evaluated. To aid comparison with existing FLSs, AIFLS and classical IT2FLS are also constructed and parameters updated using the DEKF. The viability of the resulting systems are validated by rigorous study cases and statistical tests. Particularly, the systems are used for system identification problem and evaluation of two real world datasets namely: NSW 2008 electricity dataset obtained from Australia's electricity market and GCS dataset obtained from a Nigerian-based power plant. Statistical analyses reveal that there is a significant performance improvement of IT2AIFLS over AIFLS and classical

IT2FLS trained with the same learning apparatus. It is conjectured that the improved performance of IT2AIFLS is because IT2AIFSs posses extra degrees of freedom, in terms of the non-membership functions, with the capacity to model non-linear input-output relationships better. The results presented in Table 5.3 to 5.6 reveal that IT2AIFLS trained with DEKF exhibits superior performance to that trained with GD algorithm.

Chapter 6

Hybrid Learning of IT2AIFLS as applied to Identification and Prediction Problems

A fuzzy future is a bright future.

Anonymous

6.1 Introduction

This chapter presents a novel application of a hybrid learning approach to the optimisation of membership and non-membership function parameters of the newly developed IT2AIFLS. The hybrid algorithm consisting of DEKF and GD is used to tune the parameters of the IT2AIFLS for the first time [40]. The DEKF is used to tune the consequent parameters in the forward pass while the GD method is used to tune the antecedents parts during the backward pass of the hybrid learning. The hybrid algorithm is described and evaluated, prediction and identification results together with the runtime are compared with similar existing studies in the literature. Performance comparison is made between the proposed hybrid learning model of IT2AIFLS, a type-1 AIFLS and a classical IT2FLS on the different datasets under investigation. The empirical comparison is made on the designed systems using three artificially generated datasets and four real world datasets. Analysis of results reveal that IT2AIFLS outperforms its type-1 variants, IT2FLS and most of the existing models in the literature. Moreover, the minimal run time of the proposed hybrid learning model for IT2AIFLS also puts this model forward as a good candidate for application in real time systems.

The GD (first-order derivative based) methods have been widely used as an optimisation strategy for the parameters of fuzzy systems [12]. As discussed in Section 2.6, the important aspect of GD is that it is guaranteed to reach a minimum (local in this case), but the difficulties with GD method are slow convergence and the possibility of getting stuck in local minima, leading to poor solutions [124]. To address these shortcomings, in this chapter, the first-order GD is combined with a second-order optimisation method such as the DEKF algorithm which have a smaller possibility of getting stuck in local minima [125]. The combination of these two approaches, apart from guaranteeing the goal of reaching a minimum, may also speed up the learning process. Hence, a new learning algorithm of DEKF and GD for tuning the parameters of IT2AIFLS [2] is introduced for the first time in this chapter with the aim of achieving improved system performance in terms of error minimisation and faster convergence.

6.2 IT2AIFLS Rule Structure

In this section, the generic rule structure is recalled for convenience. The generic TSK rule representation for IT2AIFLS is as expressed in Equation 6.1:

$$R_k : \text{IF } x_1 \text{ is } \tilde{A}^*{}_{1k} \text{ and } x_2 \text{ is } \tilde{A}^*{}_{2k} \text{ and } \cdots \text{and } x_n \text{ is } \tilde{A}^*{}_{nk} \text{THEN } y_k = \sum_{i=1}^{n} w_{ik} x_i + b_k \tag{6.1}$$

where $\tilde{A}^*{}_{1k}$,$\tilde{A}^*{}_{2k}$, $\cdots$,$\tilde{A}^*{}_{ik}$,$\cdots$,$\tilde{A}^*{}_{nk}$ are IT2AIFS and y_k is the output of the kth rule.

6.3 Parameter Updates

In this section, the two-pass learning algorithm for the parameters of IT2AIFLS is described. During the forward pass, the antecedent parameters are kept fixed while the consequent parameters are updated using the DEKF. During the backward pass, the consequent parameters are kept fixed while the antecedent parameters are updated using GD method. The hybrid learning procedure of DEKF and GD is as shown in Figure 6.1 and Algorithm 1.

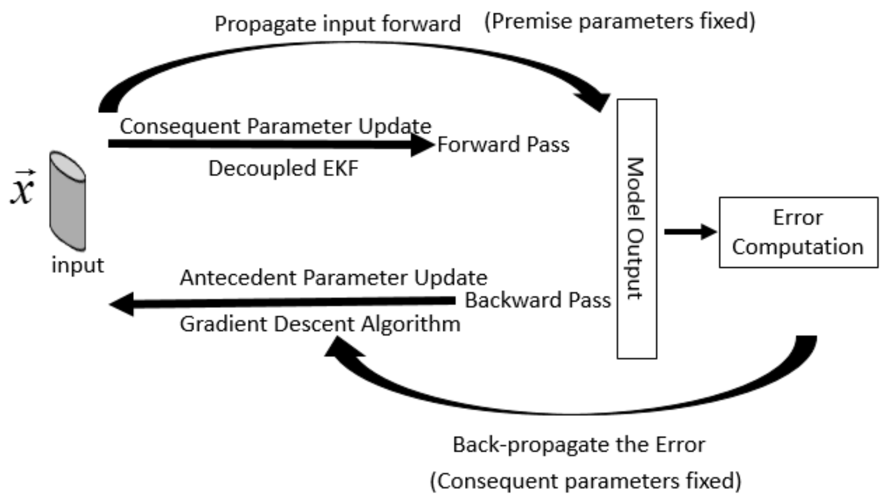

Figure 6.1: Hybrid learning procedure of IT2AIFLS using DEKF and GD

6.3.1 Consequent Parameter Updates

The decoupled EKF, is utilised to train the consequent parts of the IT2AIFLS model because it is less complex. The parameter update rules for the consequent parts of the membership and non-membership functions follow the Kalman filtering recursive procedures as earlier discussed in Chapter 5 and rewritten here in Equation 6.2 to 6.4 for ease of reference:

$$K_t^\mu = P_t^\mu H_t^\mu [(H_t^\mu)^T P_t^\mu H_t^\mu + R^\mu]^{-1} \tag{6.2}$$

$$\hat{\theta}_t^\mu = \hat{\theta}_{t-1}^\mu + K_t^\mu [y_t - h(\hat{\theta}_{t-1})] \tag{6.3}$$

$$P_{t+1}^\mu = P_t^\mu - K_t^\mu P_t^\mu (H_t^\mu)^T + Q^\mu \tag{6.4}$$

and the updates for the non-membership functions follow the same recursive procedure but utilises non-membership function parameters as in Equation 6.5 to 6.7:

$$K_t^\nu = P_t^\nu H_t^\nu [(H_t^\nu)^T P_t^\nu H_t^\nu + R^\nu]^{-1} \tag{6.5}$$

$$\hat{\theta}_t^\nu = \hat{\theta}_{t-1}^\nu + K_t^\nu [y_t - h(\hat{\theta}_{t-1})] \tag{6.6}$$

$$P_{t+1}^\nu = P_t^\nu - K_t^\nu P_t^\nu (H_t^\nu)^T + Q^\nu \tag{6.7}$$

6.3.2 Antecedent Parameter Updates

To adjust the antecedent parameters of the IT2AIFLS, GD algorithm is executed. The cost function for a single output and the inference mechanism for IT2AIFLS are rewritten for ease of reference in Equations 6.8 and 6.9 respectively.

$$E = \frac{1}{2}\left(y^a - y\right)^2 \tag{6.8}$$

where y^a is the actual output and y is the IT2AIFLS output defined as [2,46,47]:

$$y = \frac{(1-\beta)\sum_{k=1}^{M}\left(\underline{f}_k^{\mu} + \overline{f}_k^{\mu}\right) y_k^{\mu}}{\sum_{k=1}^{M}\underline{f}_k^{\mu} + \sum_{k=1}^{M}\overline{f}_k^{\mu}} + \frac{\beta\sum_{k=1}^{M}\left(\underline{f}_k^{\nu} + \overline{f}_k^{\nu}\right) y_k^{\nu}}{\sum_{k=1}^{M}\underline{f}_k^{\nu} + \sum_{k=1}^{M}\overline{f}_k^{\nu}} \tag{6.9}$$

where $\underline{f}_k^{\mu}$, $\overline{f}_k^{\mu}$, $\underline{f}_k^{\nu}$ and $\overline{f}_k^{\nu}$ are the lower membership, upper membership, lower non-membership and upper non-membership functions firing strengths respectively. The generic GD update rules for tuning the antecedent parameters (membership and non-membership) of the proposed framework is recalled for convenience.

$$\theta_{ik}(t+1) = \theta_{ik}(t) - \gamma\frac{\partial E}{\partial \theta_{ik}} \tag{6.10}$$

where γ is the learning rate and θ is the generic parameter.

6.4 Experimental Analysis and Evaluation

In this section, the experimental analyses on publicly available system identification and prediction problems are presented. Similar to previous studies on these datasets, the performance evaluation is on the basis of same datasets and performance metric, RMSE in this case, to evaluate the prediction quality of the proposed model. The RMSE is defined as:

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(y^a - y\right)^2} \tag{6.11}$$

where y^a is the desired output and y is the predicted output, N is the number of testing data points. The β value for all experiments is initialised to 0.5. The initial values of membership and non-membership functions consequent parameters are randomly generated from unit interval [0,1].

6.4.1 Application to Artificially Generated Mackey-Glass Time Series

Mackey-Glass benchmark time series for modelling a physiological system defined by the differential delay equation in (6.12) is examined:

$$\frac{dx\left(t\right)}{dt} = \frac{a * x\left(t-\tau\right)}{1 + x\left(t-\tau\right)^n} - b * x\left(t\right) \tag{6.12}$$

Algorithm 1 Hybrid Learning of IT2AIFLS

INPUT: Train set $(\overrightarrow{x}_t,\ y_t),\ t = 1...N$

OUTPUT: Prediction error

1: *initialise all antecedent* $(c,\ \underline{\sigma}^{\mu},\ \overline{\sigma}^{\mu},\ \underline{\sigma}^{\nu},\ \overline{\sigma}^{\nu})$ *and consequent parameters* $(w^{\mu}, b^{\mu}, w^{\nu}, b^{\nu})$ *of the IT2AIFS.*

2: *set the number of training epochs to unity*

3: *set the training sample point* (t) *to unity*

4: *propagate the input* $(\overrightarrow{x}_t)$ *through the IT2AIFLS hybrid model*

5: *tune the consequent parameters using DEKF according to Equation 6.2 to 6.4 for MF parameters and Equation 6.5 to 6.7 for NMF parameters*

6: *compute the output of the hybrid-IT2AIFLS using Equation 6.9*

7: *compute the difference between the actual output and predicted output of the hybrid-IT2AIFLS model and use RMSE as the cost function*

8: *back-propagate the error*

9: *tune the antecedent parameters using gradient descent back-propagation algorithm*

10: *increment the training sample point by 1* $(\overrightarrow{x}_{t+1})$

11: **If** *trained sample point* $\leq$ *total number of training sample points* **Then**

12:     *go to step 4*

13: **Else**

14:     *increment training epoch by 1.*

15: **End If**

16: **If** *maximum epoch is reached*

17:     *End*

18: **Else**

19:     *go to step 4*

20: **End If**

where a, b and n are constant values, t is the current time and τ is the time delay constant. The proposed model is evaluated with $\tau = 17$. Similar to [2,257–260], a dataset consisting of 1000 data points are generated using Equation 6.12. The first 500 data points are used for training and the remaining 500 are used for testing.

For a fair comparison with existing studies, the data generating vector is $[x(t-18), x(t-12), x(t-6), x(t); x(t+6)]$ with $x(t+6)$ as the target, where $t = 118$ to 1117. There are a total of 16 rules with 192 tunable parameters. The Kalman filter parameters Q and P for both membership function and non-membership function were initially set as $0.001I_{80}$ and $1.0I_{80}$ respectively with $R = 40$. The learning rate is fixed at 0.01 with 500 training epochs and 10 simulation runs. Figure 6.2 shows the actual and the predicted outputs of Mackey-Glass using IT2AIFLS while Figure 6.3 shows the evolution of the adaptive user define parameter, β. Comparison of results is made between IT2AIFLS trained with DEKF and GD, and its type-1 variants on Mackey-Glass benchmark dataset. Table 6.1 shows that IT2AIFLS outperforms its type-1 counterpart. A comparison of the hybrid learning approach of IT2AIFLS with some existing models in the literature is also shown in Table 6.1 with IT2AIFLS exhibiting superior predictive performance to many others but having very close predictive power to local linear wavelet neural network (LLWNN) trained with particle swarm optimisation with diversity learning and GD (LLWNN + hybrid) and LLWNN with GD alone in this problem domain.

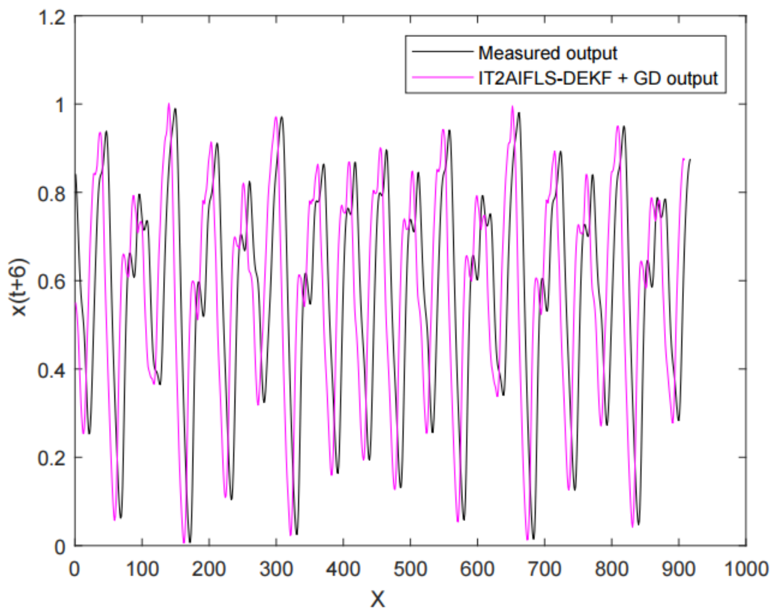

Figure 6.2: Actual and predicted output of Mackey-Glass time series

Table 6.1: Performance comparison of IT2AIFLS on Mackey-Glass time series forecasting with existing models

Model	Rules	RMSE(tst)
SuPFuNIS [261]	15	0.014
Fuzzy-Singular Value Decomposition [212]	- 10	0.012
MDE-RBF NN [211]	-	0.013
Genetic Fuzzy Ensemble [213]	-	0.0264
Radial Basis Function AFS [210]	-	0.0114
RBF-AFS [210]	21	0.013
HyFIS [113]	16	0.012
NEFPROX [262]	129	0.0332
HyFIS-Yager-gDIC [263]		0.0190
T2-HyFIS-Yager [263]		0.0694
D-FNN [264]	10	0.008
WNN + gradient [257]	-	0.0071
WNN + hybrid [257]	-	0.0059
LLWNN + gradient [257]	-	0.0041
LLWNN + hybrid [257]	-	0.0036
MLMVN [258]	-	0.0056
GEFREX [265]	-	0.0061
SA-T2FLS [259]	16	0.0089
TSK-SVR I [260]	-	0.008
TSK-SVR II [260]	-	0.007
AIFLS	16	0.0054
IT2AIFLS	16	**0.0040**

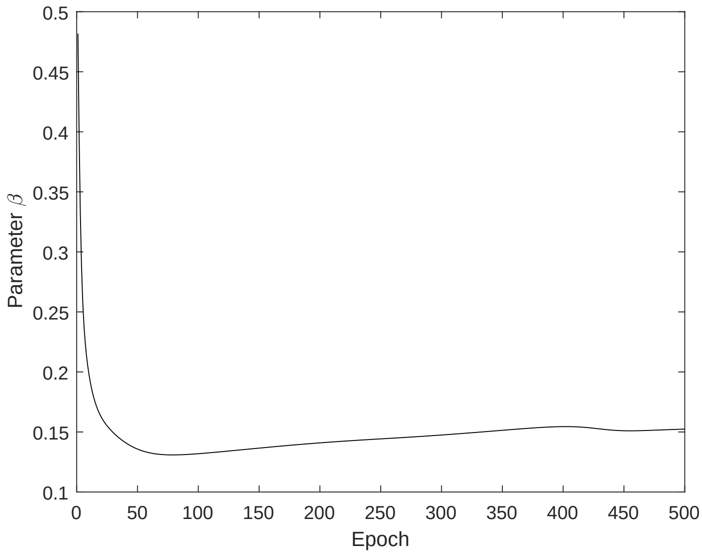

Figure 6.3: The adaptation of the parameter β for Mackey-Glass prediction problem

6.4.2 System Identification Problem #1

A second-order time-varying system is investigated using the hybrid learning model of IT2AIFLS. This first system identification problem involves a dynamic system that is defined by Equation 6.13.

$$y(t+1) = f(y(t), y(t-1), y(t-2), u(t), u(t-1)) \tag{6.13}$$

where

$$f(x_1, x_2, x_3, x_4, x_5) = \frac{x_1 x_2 x_3 x_5 (x_3 - b) + c x_4}{a + x_2^2 + x_3^2} \tag{6.14}$$

and a, b, c are time-varying parameters as shown in Figure 6.4 and defined as in Equation 6.15 to 6.17:

$$a(t) = 1.2 - 0.2cos(2\pi t/T) \tag{6.15}$$

$$b(t) = 1.0 - 0.4sin(2\pi t/T) \tag{6.16}$$

$$c(t) = 1.0 + 0.4sin(2\pi t/T) \tag{6.17}$$

Here, $T = 1000$ represents the total number of sample points. All computational procedures are arranged as closely as possible to those reported in [95, 122, 243]. Two inputs

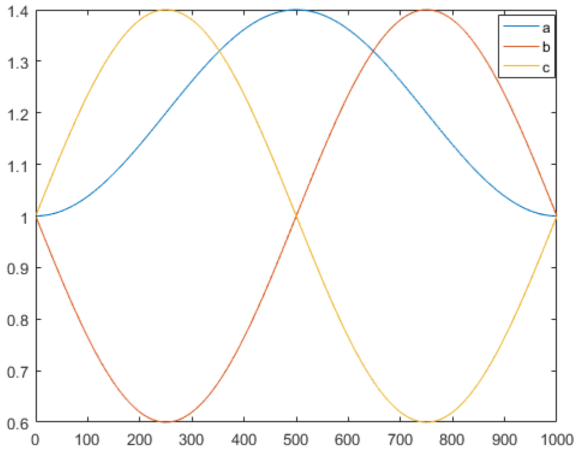

Figure 6.4: Time varying parameters for second-order system identification problem #1

values are utilised which are $u(t)$ and $y(t)$.

$$u(t) = \begin{cases} \sin(\pi t/25) & t < 250 \\ 1.0, & 250 \leq t < 500 \\ -1.0 & 500 \leq t < 750 \\ 0.3\sin(\pi t/25) + 0.1\sin(\pi t/32) & \\ +0.6\sin(\pi t/10) & 750 \leq t < 1000 \end{cases} \tag{6.18}$$

Similar to Lin *et al.* [122], the simulation is conducted for 1000 time steps with 100 training epochs. A total of 4 rules with 40 tunable parameters are generated. The learning rate was set to 0.01 while the Kalman filter parameters for P and Q are initially chosen as $1I_{12}$ and $0.001I_{12}$ respectively for the membership function and non-membership function with R chosen as 40 where I is the identity matrix. The higher value of R is chosen to increase the level of uncertainty in the data. In order to assess the performance of IT2AIFLS-DEKF and GD on the time-varying dynamic system, the test signal in Equation 6.18 is used.

Figure 6.5 shows the actual versus the predicted output for 200 data points of the second-order identification problem #1 using IT2AIFLS-DEKF and GD. As shown in Table 6.2, the hybrid model of IT2AIFLS-DEKF and GD outperforms other existing models except interval type-2 fuzzy neural network (IT2FNN) trained with EKF (IT2FNN-EKF). Although IT2AIFLS-DEKF and GD performs better than IT2FNN-EKF on the training set, IT2FNN-EKF outperforms IT2AIFLS-DEKF and GD on the test set. This could be as a result of utilising the predictive power of EKF on both the antecedent

Table 6.2: Performance comparison of IT2AIFLS with other models on second-order system identification problem #1

Model	Rules	Epoch	RMSE(trn)	RMSE(tst)
Type-1 TSK FNS [243]	9	100	0.0282	0.0598
Type-2 TSK FNS [243]	4	100	0.0284	0.0601
Feedorward Type-2 FNN	3	100	0.0281	0.0593
SIT2FNN [95]	4	100	0.0351	0.0560
SEIT2FNN [121]	3	100	0.0274	0.0574
TSCIT2FNN [122]	3	100	0.0279	0.0576
IT2FNN-GD [42]	-	200	0.0540	0.0613
IT2FNN-EKF [42]	-	200	0.0275	0.0261
IT2FNN-SMC [42]	-	200	0.0360	0.0390
IT2FNN- PSO + SMC [42]	-	200	0.0199	0.0390
IT2AIFLS	4	100	**0.0250**	**0.0310**

and consequent parameters tuning of IT2FNN-EKF. Most notably is the comparison of IT2AIFLS-DEKF and GD with self evolving interval type-2 fuzzy neural network (SEIT2FNN) and TSK-type-based self evolving compensatory interval type-2 fuzzy neural network (TSCIT2FNN). Similar to IT2AIFLS, both SEIT2FNN and TSCIT2FNN utilise Kalman filter-based methodologies to adapt their consequent parameters and GD to optimise the antecedent parameters respectively with A2-C0 TSK-type fuzzy inference. The proposed framework of IT2AIFLS outperforms both existing methods of SEIT2FNN and TSCIT2FNN in this problem instance.

For a fair comparison of the runtime of IT2AIFLS - DEKF and GD with those reported in Kayacan and Khanesar [42], 200 simulations of the experiments are conducted. As shown in Table 6.3, IT2AIFLS-DEKF and GD has the lowest runtime of 82.04 seconds, close to that of IT2FNN trained with sliding mode control (IT2FNN-SMC) algorithm with the runtime of 84.39 seconds. The reason for this short execution time is that the DEKF is only applied to learn the consequent parts of the model which has only two parameters.

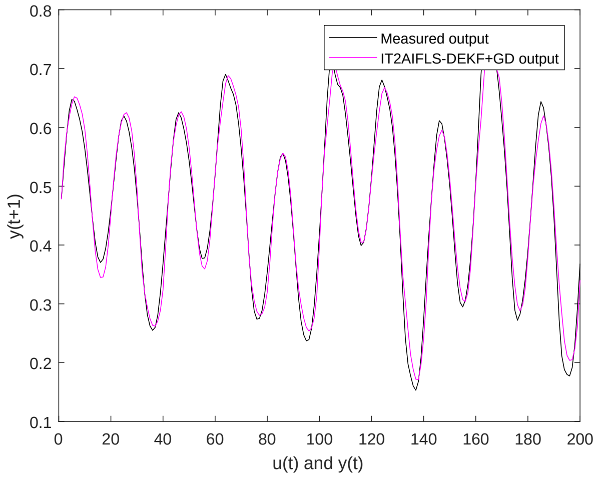

Figure 6.5: Actual and predicted output using hybrid IT2AIFLS for second-order system identification problem #1

Thus, with the superior identification accuracy and computational efficiency in terms of run time in this particular problem, the proposed IT2AIFLS-DEKF and GD model may be a more appropriate choice for real time applications compared to those reported in [42].

Table 6.3: Comparison of runtime of IT2AIFLS with other approaches on second-order identification problem #1

Model	Epoch	Run Time (s)
IT2FNN-GD [42]	200	124.12
IT2FNN-EKF [42]	200	229.71
IT2FNN-SMC [42]	200	84.39
IT2FNN PSO + SMC [42]	200	7086.78
IT2AIFLS	200	**82.04**

6.4.3 System Identification Problem #2

For further evaluation, IT2AIFLS is applied to a non-linear system identification problem where the dataset is generated by the following differential equation:

$$y(t+1) = \frac{y(t)}{1+y^2(t)} + u^3(t) \tag{6.19}$$

The variables $u(t)$ and $y(t)$ are used as inputs while $y(t+1)$ is the desired output. Training samples are generated using $u(t) = sin(2\pi t/100)$. Similar computational set up in $[95, 121, 122, 243, 266]$ are adopted with 200 samples generated and trained for 500 epochs. The proposed approach is compared with three evolving T2FLSs namely, self evolving interval type-2 fuzzy neural network (SEIT2FNN) utilising IT2FS in the antecedents and TSK interval type-1 set in the consequent, TSK-type-based self-evolving compensatory IT2FNN (TSCIT2FNN) which utilises IT2FS in the antecedent and a linear model in the consequent and evolving type-2 neural fuzzy inference system (eT2FIS) with antecedent T2FS and Mamdani-type consequent. Figure 6.6 shows the actual and predicted output for this non-linear system identification problem. As shown in Table 6.4,

Table 6.4: Performance comparison of hybrid IT2AIFLS with other models on non-linear system identification#2

Models	Rules	Parameter	RMSE(tst)
T2FLS (Singleton) [121]	5	49	0.034
T2FLS (TSK) [121]	3	36	0.0388
eT2FIS [266]	14	70	0.053
Type-2 TSK FNS [243]	4	24	0.03239
Feedforward Type-2 FNN [95]	3	36	0.0281
SIT2FNN [95]	3	36	0.0241
TSCIT2FNN [122]	3	34	0.0084
SEIT2FNN [121]	3	36	0.0062
AIFLS-GD [2]	4	36	0.0146
IT2AIFLS-GD [2]	4	40	0.0052
AIFLS	4	36	0.0101
IT2AIFLS	4	40	**0.0030**

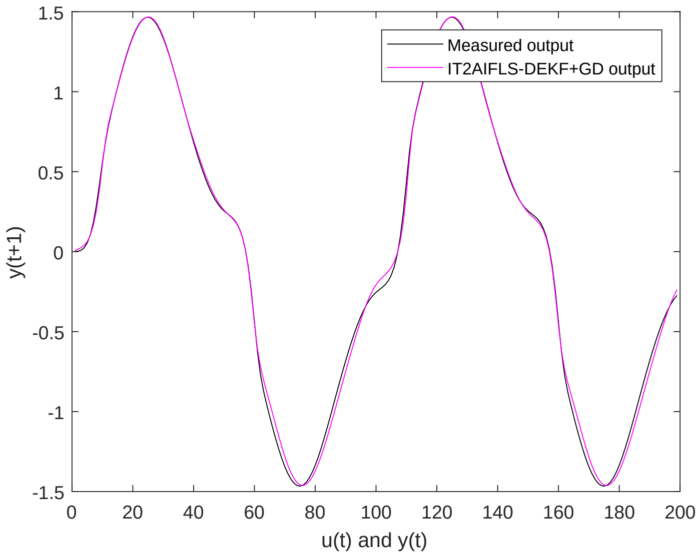

Figure 6.6: Actual and predicted values for system identification#2 using IT2AIFLS

IT2AIFLS exhibits a low level of RMSE over these evolving T2FLSs. In particular, the performance of IT2AIFLS is compared with Type-2 TSK Fuzzy Neural System (Type-2 TSK FNS) [243], TSK-type-based self evolving compensatory interval type-2 fuzzy neural network (TSCIT2FNN) [122] and SIT2FNN [95], which also utilised the T2FLS version of the inference mechanism proposed in this thesis. The results show a clear performance improvement of IT2AIFLS over Type-2 TSK FNS, TSCIT2FNN and SIT2FNN. The performance of hybrid IT2AIFLS trained with DEKF and GD is also compared with the IT2AIFLS trained with GD algorithm alone. As it can be seen in Table 6.4, the hybrid learning method of IT2AIFLS outperforms the GD-based IT2AIFLS. An AIFLS is also constructed in order to compare the performance of the IT2AIFLS with its T1 model on this system identification problem. From Table 6.4, there is a significant performance improvement of IT2AIFLS over AIFLS on system identification problem#2.

6.4.4 System Identification Problem #3

The system in Subsection 6.4.3 is modified by adding a time varying parameter, f. In this case, the parameter of the system varies with the time. The proposed hybrid model, IT2AIFLS - DEKF and GD is applied to this dynamic system with dataset generated by the differential equation [121]:

$$y(t+1) = \frac{y(t)}{1+y^2(t)} + u^3(t) + f(t)$$

where

$$f(t) = \begin{cases} 0, & 1 \le t \le 1000 \\ 1.0, & 1001 \le t \le 2000 \\ 0, & 2001 \le t \end{cases} \tag{6.20}$$

The inputs to the model are $u(t)$ and $y(t)$ while $y(t+1)$ is the desired output. The 2001 training data samples are generated using $u(t) = sin(2\pi t/100)$. There are 4 rules and 40 tunable parameters for the IT2AIFLS-DEKF and GD model. An AIFLS and IT2FLS trained with DEKF and GD are also constructed and evaluated on the system identification problem #3. The number of rules in the three models remain the same with 36 and 24 tunable parameters for the AIFLS and IT2FLS respectively. The RMSE is computed over 10 simulations for each model. As presented in Table 6.5, IT2AIFLS outperforms both AIFLS and IT2FLS in this problem instance.

Table 6.5: A Comparison of IT2AIFLS, AIFLS and IT2FLS on second-order identification problem #3

Model	RMSE(trn)	RMSE(tst)
IT2FLS	0.0173	0.0074
AIFLS	0.0172	0.0073
IT2AIFLS	**0.0151**	**0.0064**

In the following subsections, the performance of the proposed model is evaluated on three real world problems. These are Poland electricity load, Santa Fe A laser and gas furnace datasets.

6.4.5 Application to Real World Electricity Load Forecasting

Similar to system identification problem in 6.4.4, this experiment is conducted to evaluate the performance of hybrid learning of IT2AIFLS with AIFLS and IT2FLS using the same learning procedure on a real world problem. The dataset selected is the Poland electricity load data obtained from (http://research.cs.aalto.fi/) and contains electricity load values of Poland in the 1990's. Table 6.6 shows the first four input samples of Poland electricity load data.

Table 6.6: Excerpt from Poland electricity data

	x1	x2	x3	x4	y
1	0.8743	0.9631	1.0622	1.0731	1.0477
2	1.0477	0.8743	0.9631	1.0622	1.0781
3	1.0781	1.0477	0.8743	0.9631	1.0838
4	1.0838	1.0781	1.0477	0.8743	1.1063
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]

The training dataset consist of 1400 samples while 201 data samples constitute the testing set. The number of epochs is 100 with the RMSE computed over 10 simulations. A one-step-ahead prediction model is constructed with the output defined by Equation 6.21.

The input vector consists of some previous values and the current value of the time series for the prediction. The current value of the electricity load provides an up-to-date measurement to the prediction while the previous values keep track of the trend.

$$y(t+1) = [(x(t), x(t-1), \cdots, x(t-p+1)]$$

where p is the size of input with $t \geq p$. The input size of four is adopted and the input generating equation becomes:

$$y(t+1) = [(x(t), x(t-1), x(t-2), x(t-3)] \tag{6.21}$$

with $y(t+1)$ as the output.

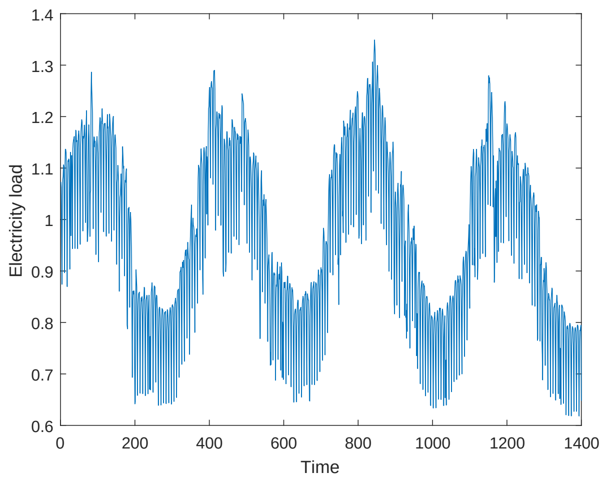

Figure 6.7: Plot of Poland electricity load training data

Figure 6.7 shows the training dataset for Poland electricity load while Figure 6.8 shows the actual and the predicted values of the test dataset. Table 6.7 shows that the perfor-

Table 6.7: Comparison of IT2AIFLS versus AIFLS and IT2FLS on Poland electricity load forecast

Model	Train/Test set	RMSE(trn)	RMSE(tst)
IT2FLS	1395/196	0.0564	0.0595
AIFLS	1395/196	0.0589	0.0599
IT2AIFLS	1395/196	**0.0560**	**0.0572**

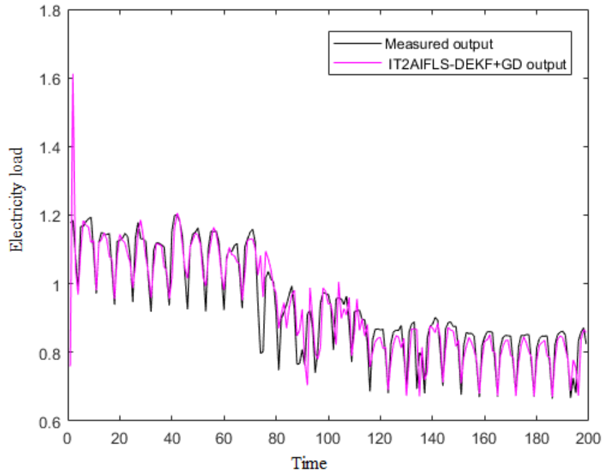

Figure 6.8: Actual and predicted values of Poland electricity load with IT2AIFLS on test dataset

mance of IT2AIFLS is superior to those of AIFLS and IT2FLS trained with the same hybrid algorithm of DEKF and GD.

6.4.6 Gas Furnace Time Series

The gas furnace time series is one of the most researched benchmark datasets for model evaluation which is generated by the combustion process of methane-air mixture. The dataset has the gas flow rate as the process input and the carbon-dioxide (CO_2) concentration as the process output. The gas furnace dataset is downloaded from [267]. The dataset consist of 296 data pairs. Shown in Table 6.8 are the first four input samples of the gas furnace data. From existing studies, the best input-output model structure for this application domain is: $y(t) = f(u(t-4), y(t-1))$. For ease of comparison with earlier studies, the simulation settings are arranged to be as close as possible to those reported in [222, 223, 257]. The task is to forecast the amount of CO_2 concentration in the gas at time (t) using the input data with methane flow rate at time $(t-4)$ and the amount of CO_2 produced at time $(t-1)$, i.e. y(t) $= [u(t-4), y(t-1)]$. After conversion to $[u(t-4), y(t-1); y(t)]$ input-output pairs, the dataset is reduced to 292 sample points of which 200 data points are used for training and 92 samples used for testing. Figure 6.9 shows the actual and predicted outputs of gas furnace time series problem. As shown in

Table 6.8: Excerpt from gas furnace data

	x1	x2	y
1	-0.109	53.5	53.4
2	0	53.4	53.1
3	0.178	53.1	52.7
4	0.339	52.7	52.4
5	[illegible]	[illegible]	[illegible]

Table 6.9, IT2AIFLS trained with DEKF and GD performs better than its type-1 counterpart with the same training procedure. Comparison with existing studies on the other hand shows IT2AIFLS performing better than or comparatively with other works in the literature.

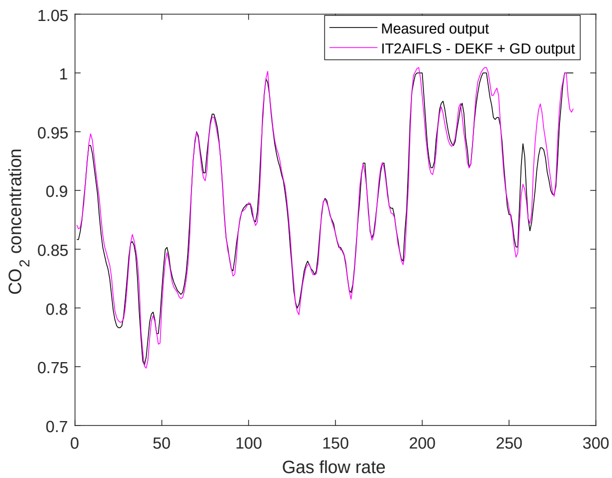

Figure 6.9: Actual and predicted output of gas furnace time series using IT2AIFLS trained with DEKF+GD

Table 6.9: Performance comparison of hybrid-IT2AIFLS on gas furnace time series

Model	Rules	Parameter	RMSE(tst)
ARMA [268]	-	-	0.843
Tongs' model [269]	19	-	0.685
Pedrycz's model [270]	81	-	0.566
Xu's model [271]	25		0.573
Sugeno's model [272]	6	-	0.596
Surmann's model [273]	25	-	0.400
Lee's model [274]	25	-	0.638
Lin's model [275]	4		0.511
Nie's model [276]	45	225	0.412
ANFIS [277]	4	24	0.085
Neural Tree [278]	-	-	0.0257
eTS [279]	5	-	0.04904
Simpl-eTS3 [279]	3	-	0.04849
WNN + gradient [257]		40	0.084
WNN + hybrid [257]		40	0.081
LWNN + gradient [257]		56	0.01643
LWNN + hybrid [257]		56	0.01378
FWNN-S (2MFs) [222]	-	32	0.03085
FWNN-S (3MFs) [222]	-	66	0.02778
FWNN-R (2MFs) [222]	-	28	0.03171
FWNN-R (3MFs) [222]	-	57	0.02794
FWNN-M (2MFs) [222]	-	32	0.02963
FWNN-M (3MFs) [222]	-	66	0.02324
LLNF (2 inputs) [223]	-	-	0.0462
OSSA-LLNF [223] (2 inputs)		-	0.0321
AIFLS	4	36	0.0273
IT2AIFLS 2(MFs)	**4**	**40**	**0.0249**

6.4.7 Santa Fe A Time Series

The proposed hybrid-IT2AIFLS model is also applied to the Santa Fe A time series in order to evaluate the performance of the hybrid model on another real world application.

The Santa Fe A time series had earlier been presented in Chapter 4, subsection 4.4.4. The same input-output generating vector: $[x(t-1), x(t-2), x(t-3), x(t-4), x(t-5); x(t)]$ is utilised. The training dataset consists of 90% of the entire dataset while the remaining 10% are used for testing with 500 training epochs and 10 number of trials. The Kalman filter parameters P and Q are set to $1.0I_{192}$ (I_{192} is the 192 by 192 identity matrix) and $0.001I_{192}$ for membership and non-membership functions respectively and R is chosen as 40. The user design parameter β is initially set to 0.5 with the learning rate, $\gamma = 0.5$. There are 32 rules generated with 8(5) + 2*32(5+1) = 424 tunable parameters. The actual and predicted output of the Santa Fe time series is shown in Figure 6.10.

Table 6.10: Performance comparison of hybrid-IT2AIFLS with other models on Santa Fe A time series dataset

Model	Rule	Parameter	RMSE(trn)	RMSE(tst)
ES [242]	-	-	-	56.20
NN [242]	-	-	-	24.6
PMRS [242]	-	-	-	14.23
SONFIN* [202]	9	144	6.956	5.983
T2FLS-G* [118]	5	135	8.50	7.16
SEIT2FNN* [121]	5	135	7.677	5.766
IT2FNN-SVR(N) [199]	5	106	13.565	4.337
IT2FNN-SVR(F) [199]	5	106	9.094	3.474
SVR-FM* ($\epsilon = 0.1$) [244]	31	188	14.370	9.707
SVR-FM* ($\epsilon = 0.001$) [244]	747	4484	7.069	1.650
IT2AIFLS	32	424	**6.075**	**1.668**

* These results are adapted from [199]

In Table 6.10 the results obtained from IT2AIFLS and AIFLS, both trained with DEKF and GD are shown together with other existing approaches in the literature. As shown in Table 6.10, IT2AIFLS model outperforms AIFLS trained with the same hybrid algorithm. The IT2AIFLS also performs better than other models in the literature with very low RMSE on the test set, thus demonstrating a good generalisation and predictive capability of the proposed IT2AIFLS model. Shown in Table 6.11 is a summary of all the datasets and the corresponding results on the test sets presented in this chapter. The best

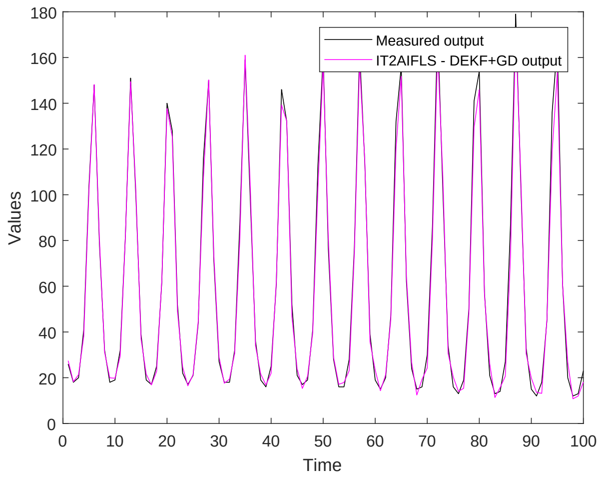

Figure 6.10: Actual and predicted values of Santa Fe A time series with IT2AIFLS

results are shown in bold face.

Table 6.11: Summary of results

Dataset	Measure	Proposed Model	Best other
Mackey-Glass	RMSE	0.0040	**LLWNN + hybrid - 0.0036**
System identification #1	RMSE	0.0310	**IT2FNN(EKF) - 0.0261**
System identification #2	RMSE	**0.0030**	SEIT2FNN - 0.0062
System identification #3	RMSE	**0.0064**	AIFLS - 0.0073
Poland electricity	RMSE	**0.0560**	IT2FLS - 0.0595
Gas furnace	RMSE	0.0249	**LWNN + hybrid - 0.0138**
Santa Fe	RMSE	1.6680	**SVM-FM - 1.6500**

6.5 Summary

This chapter presents a novel application of a hybrid approach of DEKF and GD to optimise the parameters of IT2AIFLS. The DEKF is used to learn the consequent parameters of the model while GD is applied to the tuning of the antecedent parameters. The hybrid learning algorithm consisting of DEKF and GD is also used to tune the parameters of type-1 AIFLS and IT2FLS for performance comparison between its type-1 version and conventional IT2FLS.

From simulation analyses, IT2AIFLS exhibits superior performance quality to those of AIFLS and IT2FLS. In the overall, the developed model of IT2AIFLS exhibits better or comparatively good prediction and identification performances compared to similar studies in the literature. The run time of the proposed IT2AIFLS - DEKF and GD is very short compared to other previous models on the same problem domain. This is an indication that the proposed hybrid IT2AIFLS model may be more appropriate for real time applications compared to those in [42].

Chapter 7

Conclusions and Discussion

> For I know the thoughts that I think toward you, saith the Lord, thoughts of peace, and not of evil, to give you an expected end.
>
> Jeremiah 29:11

To conclude the thesis, this chapter summarises the studies of the preceding chapters, highlights the thesis contributions, discusses the limitations of the model and outlines potential avenues for future work.

7.1 Discussion

In recent times, there have been intensive modelling of uncertainties using the classical IT2FLS, whose non-membership is complementary to the membership (lower or upper). It is observed that this may not necessarily fit within the context of natural/human language representations. This is especially the case in situations where some people are indifferent to certain situations or are hesitant about certain concepts. This research has investigated this short coming by modelling uncertainties using separately defined membership and non-membership functions that are intervals, such that the sum of lower membership and upper non-membership degrees is less than or equal to 1; and the sum of upper membership and lower non-membership degrees is less than or equal to 1, thus relaxing the complementarity notion of the classical IT2FSs and enabling hesitation.

In Chapter 2, a survey of related literature of FLSs is presented with a particular focus on IT2FLS and AIFLS, both of which are key components in the model presented in this thesis. Within these context, reviews of relevant studies are conducted. The strengths and weaknesses of these approaches are highlighted.

In Chapter 3, the IT2AIFLS framework that utilises IT2AIFS is formulated and the different components are discussed. The IT2AIFS design is achieved through a rigorous and systematic integration of AIFS with IT2FS. The ensued interval membership and non-membership functions incorporate IF-indices. This concept of integrating non-membership functions and IF-indices into the classical IT2FLS gives the new framework (IT2AIFLS) an advantage over classical IT2FLS as demonstrated in the simulation examples in Chapters 4, 5 and 6. As argued in Eyoh *et al.* [3], this concept provides a synergistic capability for uncertainty modelling with improved system performance. Atanassov and Gargov [73] presented a similar approach that utilised interval membership and non-membership functions, the so-called IVAIFS. As pointed out in Section 3.4, for IVAIFS, the sum of upper membership and upper non-membership is less than or equal to 1. For the proposed model, the sum of upper membership and lower non-membership lies in a closed interval of 0 and 1, and the sum of lower membership and upper non-membership also lies in a closed interval of 0 and 1. These present a point of departure of the proposed model from existing IVAIFS. Moreover, going by Bustince *et al.* [84], the IT2AIFS is a broader concept and can be used in many more ways than IVAIFS (which is a specific representation of IT2AIFS). For a FLS to be optimally applied, the parameters have to be tuned.

In Chapter 4, a novel application of GD algorithm to tune the parameters of the antecedent and consequent parts of the IT2AIFLS is presented. The GD is one of the popular FLS parameter optimisation tools [116, 117]. The GD-based IT2AIFLS is evaluated using publicly available benchmark time series and regression problems. The experimental analyses reveal that with the integration of non-membership functions and IF-indices into IT2FLS, the new framework (IT2AIFLS) can increase the level of performance in terms of prediction accuracy in many application domains. This new model is shown to outperform its type-1 variant and some existing approaches (see [2, 3, 46]).

In Chapter 5, for the first time, the parameters of the antecedent and consequent parts of the IT2AIFLS are optimised using the DEKF algorithm and evaluated on one synthetic and two real world problems. The idea of adopting the second order derivative-based method is to address the shortcomings of the first order GD methods such as slow convergence and the greater possibility of getting entrapped in local minima. The IT2AIFLSs trained with DEKF is shown to outperform its GD counterparts in terms of prediction accuracies (see [47]). The IT2AIFLS also exhibits statistically significant performance compared to its type-1 counterpart and classical IT2FLS.

In Chapter 6, a novel application of hybrid approach of DEKF and GD to adjust the antecedent and consequent parameters of IT2AIFLS is demonstrated. The hybrid algorithm is evaluated using benchmark identification and prediction problems. The IT2AIFLS is shown to outperform its type-1 and many other existing models in these problem domains (see [40]).

7.2 Summary of Contribution

The focus of this research has been on investigating the effectiveness of integrating non-membership FOUs and IF-indices into the classical IT2FSs. The overarching question this research attempted to answer is: can the integration of non-membership function FOUs and IF-indices into the classical IT2FSs improve the performance of a FLS?

In summary, the contributions of the research presented in this thesis are:

- Introduction of a new and enhanced fuzzy framework namely, IT2AIFLS, that extends on existing approach of IT2FLS through the integration of non-membership function FOUs and IF-indices into IT2FS and thus, enabling hesitation.

- A rigorous analysis of different optimisation methods to tune the parameters of the developed framework. By tuning the antecedent and consequent parameters with different learning algorithms adopted in this research, IT2AIFLS significantly increases the performance of the system in uncertain environments.

- Tuning the contributions of the membership and non-membership functions in the final output using the learning algorithms which allows the non-uniformity of uncertainties in the rule-base of IT2AIFLS to be effectively managed.

- Analysis and comparison of the new method against existing methods on benchmark problem instances. The developed model was evaluated using benchmark problems. The IT2AIFLS is shown to outperform type-1 AIFLSs, many classical T2FLSs and non-fuzzy approaches on these problem instances.

- Statistical analysis and comparison of IT2AIFLS with alternative approaches. The IT2AIFLS is evaluated using a real world problem instance obtained from a Nigerian-based power plant. The IT2AIFLS (having membership functions, non-membership functions that are intervals and IF-indices) is shown to perform significantly better than classical IT2FLS that have only the interval membership functions and type-1 AIFLS with single membership and non-membership functions in making accurate estimate in uncertain environments.

In the overall, the clear conclusion that the new class of IT2FLS (IT2AIFLS) can enhance the capabilities of IT2FLSs in uncertainty modelling is an important advancement in type-2 fuzzy logic system's research and may serve to open up more promising research areas for uncertainty modelling.

7.3 Limitations of the Proposed Framework

The major limitation of the proposed model is the restrictive use of grid partitioning for generating fuzzy rules. This approach leads to an exponential growth of the rules as input space increases (the so-called curse of dimensionality). Based on this premise, the developed model in its current state may only be appropriate for small dimensional datasets. Other resultant limitations of the proposed model is the increased runtime and intensive resource utilisation in terms of memory usage.

Although IT2AIFLS outperforms its type-1 counterparts and most of the classical IT2FLSs in the simulated examples, these are achieved at the expense of greater computational burden for IT2AIFLS than either its type-1 or classical IT2FLS. Nevertheless, if accuracy of prediction is the main essence of the evaluation, then this computational issue may be overlooked or regarded as a small price to pay for achieving better performance in the face of uncertainties [15].

7.4 Future Research Directions

This research has introduced a new framework namely, the IT2AIFLS for uncertainty modelling. It is shown through simulation examples that IT2AIFLS can be a potential tool in obtaining as accurate an estimate as possible under uncertain environments. The promising results obtained shows that IT2AIFLS framework is indeed a viable tool that can be adopted for uncertainty modelling in diverse problem domains and characteristics. Following this, a number of opportunities that should be explored to further benefits from the strengths of this class of FLSs are discussed.

7.4.1 Non-derivative Methods for Tuning the Parameters of IT2AIFLS

One way the proposed model could be enhanced is the application of non-derivative based methods for its parameters update. The entire learning algorithms in the framework of this research are derivative-based methods (first and second order) which include partial derivatives. Computing these derivatives may be quite tedious and complex. In large and complex search spaces, the convergence speed may be relatively slow compared to other learning algorithms. The choice of learning rate for algorithms like the GD may also be challenging. In order to assess the developed model further, non-derivative approaches such as genetic algorithms, simulated annealing, particle swarm optimisation, sliding mode control algorithms and their hybrids may be exploited for tuning the parameters of the IT2AIFLS.

7.4.2 Structure Learning of the Proposed Framework

The FLS model presented in this thesis only considers parameter tuning of the model. As discussed in Section 2.6, FLS design methodology involves both parameter and structure learning. The believe is that learning the structure of the proposed model can add some interesting dimensions to the current model and to uncertainty modelling process in general. For instance, it would be interesting to explore how adapting the model structure, that is, adjusting the number of rules, may affect the overall performance of the proposed model. This also points to ways of tackling the exponential growth of the model parameters as the input dimension increases. These can be achieved through the use of fuzzy clustering approaches and other input partitioning strategies such as scatter partitioning.

7.4.3 Use of Adaptive Learning Rate and IF-indices

Another potential area for enhancing the model proposed here is the adaptive tuning of the learning rate and IF-indices. Learning rate plays important role in the learning process and affects the speed of convergence and stability of the FLS learning process [280]. The choice of learning rate in a GD algorithm can be very challenging. Although a small learning rate may lead to a reliable training, the learning process will take a lot of time because small steps are taken towards the minimum of the loss function which can lead to possible entrapment in local minima [117]. On the other hand, if the learning rate is large, then there may be oscillation and possible instability in the learning process and training may not converge or even diverge. It is possible to start with a small/large learning rate, and gradually increase/decrease the learning rate as the training progresses. Two possible guidelines for increasing or decreasing the learning rate are provided in [42]. In Abiyev and Kaynak [243], an acceptable learning rate can also be obtained using the Lyapunov function. With GD algorithm, the IF-indices can be tuned by following the GD parameter update rule (see Subsection 2.6.1). It will be interesting to investigate further, the performance of the proposed model with learning rate and IF-indices that are tuned as compared to those with fixed values.

7.4.4 Application of other Membership and Non-membership Functions, $t-norms$ and Fuzzification Procedure

The FLS framework presented in this thesis is restricted to only the Gaussian functions. One of the challenges in FLSs is the membership (non-membership) functions specification. This is because the particular membership function (non-membership) and the associated parameters significantly influence the performance of FLSs. A variety of strategies to improve upon the membership functions selection have been researched for both the type-1 and type-2 FLSs based on the use of human experts, evolutionary approaches and neural networks [155]. Recently, Mendel [16] pointed out that appriopriate choice of membership functions in a FLS contributes to increased performance. Mendel laid more emphasis on triangular and trapezoidal membership functions because they tend to have greater sculpting (partitioning) capability of the input space. Hence, another potential research direction is to exploit ways of selecting appropriate membership and non-membership functions for the proposed IT2AIFLS. Albeit, Wu [193] has provided twelve important considerations for selecting between Gaussian and trapezoidal membership functions in a classical IT2FLS. Borrowing from the words of Wagner and Hagras [155], *"there is still much work needed to standardise and simplify the selection of appropriate membership* (non-membership) *functions"* in an IT2AIFLS.

Also, the performance of IT2AIFLS using only the product $t-norm$ is investigated in this research. It will be interesting in the future to see the effects other aggregation functions such as minimum $t-norm$ and Lukasiewicz $t-norm$ may have on the performance of the system.

Moreover, this research employs singleton type-2 fuzzification where the inputs are assumed to be perfect. With singleton type-2 fuzzification, the uncertainty in the input is handled through linguistic labels in the antecedent fuzzy sets. Another interesting research direction is to investigate the modelling of input uncertainty by using non-singleton type-2 fuzzification.

7.4.5 Stability Analysis of IT2AIFLS

Furthermore, with the inference mechanism of IT2AIFLS represented in closed form, an interesting research opportunity therefore presents itself; which is the stability analysis of the IT2AIFLS. Stability is a fundamental concepts in complex dynamic systems to establish the necessary conditions for its safety. The stability analyses have been conducted for classical IT2FLS based on Lyapunov stability theory [281, 282]. This is a non−trivial problem and it will be worth exploring the developed model further in this direction.

7.4.6 Integration of AIFS with GT2FLS

This research focused strictly on the simplified version of the T2FLS. Another interesting research direction is to investigate ways of integrating AIFS with a GT2FLS. A GT2FS, apart from the primary membership functions, also has supports on the third dimension called the secondary membership degrees that defines the possibilities for the primary memberships. The current model representation based on interval type-2 cannot model the variations of the uncertainty within the FOU because the uncertainties are weighted evenly across the FOU and this leads to loss of some important information [155]. It is well known that the third dimension of a GT2FSs offers extra degrees of freedom for a GT2FS, and learning this third dimension provides greater capability, thereby making it possible for GT2FLSs to model uncertainties better than IT2FLSs [207]. In recent years, there has been a steady growth in the applications of GT2FLSs [155, 207, 283–285] to uncertainty modelling, thanks to the simplification of GT2FS by Liu [19] and Mendel *et al.* [17] through the α-plane representation. In this way, it is possible to represent a GT2FS, in a much more straightforward way, as a union of 2-D α-planes and it therefore stands as a realistic and promising alternative for the future.

7.4.7 Formulating the Model based on Mamdani Fuzzy Inference

Two widely used types of fuzzy inferencing are the Mamdani [286] and TSK [91, 92] fuzzy inference. The major difference between the two is that the consequents of Mamdani are represented by fuzzy sets while those of TSK are represented as functions of the inputs. Their applicability depends on the level of interpretability or accuracy expected of the designed system. Thus, the two models differ in their representation and output evaluation which ultimately affect their levels of interpretability and accuracy. In this research, the aim is to optimise the parameters of the system in order to obtain a fuzzy system that closely approximates the input-output relationship of the modelled system, hence the use of a TSK-fuzzy inference for IT2AIFLS. Another future research is to investigate ways of formulating the proposed model based on Mamdani fuzzy inference, where the uncertainties are captured by both the antecedent and consequent parts of the rules. That is, both the antecedent and consequent parts are IT2AIFSs. The current work, presented in this thesis, used a TSK fuzzy inference where only the antecedent part of the rule is uncertain and the consequent is a linear combination of the inputs. This constraint may limit the applicability of IT2AIFLS-TSK fuzzy inference, as some real-world applications may require the consequents to also be IT2AIFSs, in order to support system interpretability. Thus, in such cases, with uncertain antecedent and consequent parts, Mamdani fuzzy inference suffices. It will be worth exploring the Mamdani-based IT2AIFLS in the future and compare their performances.

7.4.8 Analysis of Data Stream

The work presented in this thesis relies on algorithms that use static information. The assumption here is that the entire training set is available. It is, however, rarely the case that real-world problems are static. Another interesting research direction therefore, is the application of the proposed methodology to data stream analysis that requires special treatment due to changing concept, different from the traditional approaches where every new instance contribute to the overall concept. Data stream comes in large volume and speed that it is impossible to store the whole data on disk and to process these information on the fly. Due to changes in the data, the model built on old data may change with the new data as they arrive, requiring regular updating of the model to accommodate new instances - a process often referred to as concept drift. Algortihms must be designed in order to account for these changes in the evolving structure of the system. Although the GD method used in this thesis is incremental in nature, studies show that tracking concept drifting characteristics is beyond ordinary incremental learning procedures [287]. There may be some processing delay and this may not be acceptable in some problem domains. Also, GD may not be able to appropriately cope with changes in data relationships such as a change from linear to non-linear relationship. It may therefore be beneficial to exploit some form of incremental learning algorithm adaptation for IT2AIFLS that scales/evolves accordingly with the incoming data and this may require the use of some forms of forgetting mechanisms [288,289]. The forgetting mechanism allows the model to track changes of the observed phenomenon such that parts of the knowledge which do not reflect current observations are removed. The idea of the application of fuzzy systems to data streams has been investigated in [289–291] for type-1 fuzzy models. Recently, Pratama [292] has proposed the use of classical type-2 recurrent fuzzy neural networks for managing concepts drift in data streams. This is a virgin area, and it will be interesting to investigate an adaptation of the proposed model to data stream analysis with concept drifting characteristics.

7.5 Summary

In this chapter, a reflection of the research presented in this thesis is provided, its drawbacks and possible future directions to assess it further are enumerated. The contributions of this research are also highlighted.

As a final remark, the high level of uncertainties encountered in modern day society has called for more enhanced methods for handling these uncertainties. Whilst the classical IT2FLSs have made significant waves in modelling large amounts of uncertainties, they are not able to manage indeterminate (hesitant) state well, therefore, a new and enhanced class of IT2FLS that enables hesitation, the so-called IT2AIFLS, has been introduced in this thesis for managing these high level of uncertainties more efficiently (better). The model presented takes cognisance of non-membership function FOUs and IF-indices. The non-membership function FOUs allow IT2AIFLS to capture more information than the classical IT2FLSs while the IF-indices allow evaluation of concepts to be more meaningful and consistent with human reasoning and natural language representation than other representative FLSs such as classical IT2FLSs. Whilst the proposed model has been defined in general terms, care has been taken to investigate some concrete examples and application settings involving classical IT2FLS and the proposed model. As demonstrated in this thesis through simulation examples, the IT2AIFLS significantly improves prediction performances compared to some existing approaches. It is believed that the value of this new and enhanced class of IT2FLS has far-reaching impact especially in environments where the description of a problem in terms of only fuzzy membership functions seems too restrictive. By harnessing the potentials of the classical IT2FS and AIFS, this research has contributed to addressing the existing problem of uncertainty modelling.

Bibliography

[1] J. Mendel, H. Hagras, W.-W. Tan, W. W. Melek, and H. Ying, *Introduction to type-2 fuzzy logic control: theory and applications*. John Wiley & Sons, 2014.

[2] I. Eyoh, R. John, and G. De Maere, "Interval type-2 intuitionistic fuzzy logic system for non-linear system prediction," in *2016 IEEE International Conference on Systems, Man and Cybernetics*, pp. 1063–1068, 2016.

[3] I. Eyoh, R. John, and G. D. Maere, "Interval type-2 intuitionistic fuzzy logic for regression problems," *IEEE Transactions on Fuzzy Systems*, vol. PP, no. 99, pp. 1–1, 2017. DOI: 10.1109/TFUZZ.2017.2775599.

[4] M. A. Khanesar, E. Kayacan, M. Teshnehlab, and O. Kaynak, "Analysis of the noise reduction property of type-2 fuzzy logic systems using a novel type-2 membership function," *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, vol. 41, no. 5, pp. 1395–1406, 2011.

[5] `https://www.kaggle.com/arthurtok/feature-ranking`. Accessed: 2017-10-21.

[6] L. A. Zadeh, "Fuzzy sets," *Information and control*, vol. 8, no. 3, pp. 338–353, 1965.

[7] L. A. Zadeh, "The concept of a linguistic variable and its application to approximate reasoning-i," *Information Sciences*, vol. 8, pp. 199–249, 1975.

[8] K. T. Atanassov, "Intuitionistic fuzzy sets," *Fuzzy sets and Systems*, vol. 20, no. 1, pp. 87–96, 1986.

[9] L. A. Zadeh, "Fuzzy logic= computing with words," *Fuzzy Systems, IEEE Transactions on*, vol. 4, no. 2, pp. 103–111, 1996.

[10] J. T. Rodríguez, C. Franco, D. Gómez, and J. Montero, "Paired structures, imprecision types and two-level knowledge representation by means of opposites," in *Novel*

Developments in Uncertainty Representation and Processing, pp. 3–15, Springer, 2016.

[11] H. A. Hagras, "A hierarchical type-2 fuzzy logic control architecture for autonomous mobile robots," *IEEE Transactions on Fuzzy Systems*, vol. 12, no. 4, pp. 524–539, 2004.

[12] J. M. Mendel, "Uncertain rule-based fuzzy logic system: introduction and new directions," 2001.

[13] M. B. Gorzałczany, "A method of inference in approximate reasoning based on interval-valued fuzzy sets," *Fuzzy sets and systems*, vol. 21, no. 1, pp. 1–17, 1987.

[14] M. Gehrke, C. Walker, and E. Walker, "Some comments on interval valued fuzzy sets!," *structure*, vol. 1, p. 2, 1996.

[15] J. M. Mendel and R. B. John, "Type-2 fuzzy sets made simple," *Fuzzy Systems, IEEE Transactions on*, vol. 10, no. 2, pp. 117–127, 2002.

[16] J. M. Mendel, "Explaining the performance potential of rule-based fuzzy systems as a greater sculpting of the state space," *IEEE Transactions on Fuzzy Systems*, vol. PP, no. 99, pp. 1–1, 2017. DOI:10.1109/TFUZZ.2017.2774190.

[17] J. M. Mendel, F. Liu, and D. Zhai, "α-plane representation for type-2 fuzzy sets: Theory and applications," *IEEE Transactions on Fuzzy Systems*, vol. 17, no. 5, pp. 1189–1207, 2009.

[18] S. Coupland and R. John, "Towards more efficient type-2 fuzzy logic systems," in *Fuzzy Systems, 2005. FUZZ'05. The 14th IEEE International Conference on*, pp. 236–241, IEEE, 2005.

[19] F. Liu, "An efficient centroid type-reduction strategy for general type-2 fuzzy logic system," *Information Sciences*, vol. 178, no. 9, pp. 2224–2236, 2008.

[20] S. Greenfield, F. Chiclana, R. John, and S. Coupland, "The sampling method of defuzzification for type-2 fuzzy sets: experimental evaluation," *Information Sciences*, vol. 189, pp. 77–92, 2012.

[21] J. M. Mendel, "On answering the question where do i start in order to solve a new problem involving interval type-2 fuzzy sets?," *Information Sciences*, vol. 179, no. 19, pp. 3418–3431, 2009.

[22] D. Wu and J. M. Mendel, "Uncertainty measures for interval type-2 fuzzy sets," *Information sciences*, vol. 177, no. 23, pp. 5378–5393, 2007.

[23] S. Coupland, "Type-2 fuzzy sets: geometric defuzzification and type-reduction," in *Foundations of Computational Intelligence, 2007. FOCI 2007. IEEE Symposium on*, pp. 622–629, IEEE, 2007.

[24] J. M. Mendel, R. I. John, and F. Liu, "Interval type-2 fuzzy logic systems made simple," *IEEE Transactions on Fuzzy Systems*, vol. 14, no. 6, pp. 808–821, 2006.

[25] J. Y. Ahn, K. S. Han, S. Y. Oh, and C. D. Lee, "An application of interval-valued intuitionistic fuzzy sets for medical diagnosis of headache," *International Journal of Innovative computing, Information and control*, vol. 7, no. 5, pp. 2755–2762, 2011.

[26] J. M. Mendel, "Computing with words: Zadeh, turing, popper and occam," *IEEE computational intelligence magazine*, vol. 2, no. 4, pp. 10–17, 2007.

[27] P. Melin and O. Castillo, "A review on type-2 fuzzy logic applications in clustering, classification and pattern recognition," *Applied soft computing*, vol. 21, pp. 568–577, 2014.

[28] O. Castillo and P. Melin, "A review on interval type-2 fuzzy logic applications in intelligent control," *Information Sciences*, vol. 279, pp. 615–631, 2014.

[29] E. Celik, M. Gul, N. Aydin, A. T. Gumus, and A. F. Guneri, "A comprehensive review of multi criteria decision making approaches based on interval type-2 fuzzy sets," *Knowledge-Based Systems*, vol. 85, pp. 329–341, 2015.

[30] T. Dereli, A. Baykasoglu, K. Altun, A. Durmusoglu, and I. B. Türksen, "Industrial applications of type-2 fuzzy sets and systems: A concise review," *Computers in Industry*, vol. 62, no. 2, pp. 125–137, 2011.

[31] H. Hagras and C. Wagner, "Towards the wide spread use of type-2 fuzzy logic systems in real world applications," *IEEE Computational Intelligence Magazine*, vol. 7, no. 3, pp. 14–24, 2012.

[32] K. Tai, A.-R. El-Sayed, M. Biglarbegian, C. I. Gonzalez, O. Castillo, and S. Mahmud, "Review of recent type-2 fuzzy controller applications," *Algorithms*, vol. 9, no. 2, p. 39, 2016.

[33] D. Wu, "On the fundamental differences between interval type-2 and type-1 fuzzy logic controllers," *IEEE Transactions on Fuzzy Systems*, vol. 20, no. 5, pp. 832–848, 2012.

[34] H. Hagras, "Type-2 FLCs: a new generation of fuzzy controllers," *Computational Intelligence Magazine, IEEE*, vol. 2, no. 1, pp. 30–43, 2007.

[35] S. Naim and H. Hagras, "A type 2-hesitation fuzzy logic based multi-criteria group decision making system for intelligent shared environments," *Soft Computing*, vol. 18, no. 7, pp. 1305–1319, 2014.

[36] L. A. Zadeh, "Fuzzy logic − a personal perspective," *Fuzzy Sets and Systems*, vol. 281, pp. 4–20, 2015.

[37] H. Singh, M. M. Gupta, T. Meitzler, Z.-G. Hou, K. K. Garg, A. M. Solo, and L. A. Zadeh, "Real-life applications of fuzzy logic," *Advances in Fuzzy Systems*, vol. 2013, 2013.

[38] A.-K. T. Cornelis, Chris and E. Kerre, "Intuitionistic fuzzy sets and interval valued fuzzy sets: a critical comparison," in *Proc. Eusflat03*, Citeseer, 2003.

[39] D.-F. Li, *Decision and game theory in management with intuitionistic fuzzy sets*, vol. 308. Springer, 2014.

[40] I. Eyoh, R. John, G. D. Maere, and E. Kayacan, "Hybrid learning for interval type-2 intuitionistic fuzzy logic systems as applied to identification and prediction problems," *IEEE Transactions on Fuzzy Systems*, vol. PP, no. 99, pp. 1–1, 2018. DOI: 10.1109/TFUZZ.2018.2803751.

[41] O. Castillo, A. Alanis, M. Garcia, and H. Arias, "An intuitionistic fuzzy system for time series analysis in plant monitoring and diagnosis," *Applied Soft Computing*, vol. 7, no. 4, pp. 1227–1233, 2007.

[42] E. Kayacan and M. Khanesar, *Fuzzy Neural Networks for Real Time Control Applications: Concepts, Modeling and Algorithms for Fast Learning.* Elsevier Science, 2015.

[43] S. Chopra, R. Mitra, and V. Kumar, "Fuzzy controller: choosing an appropriate and smallest rule set," *international journal of computational cognition*, vol. 3, no. 4, pp. 73–78, 2005.

[44] K. Hornik, M. Stinchcombe, and H. White, "Multilayer feedforward networks are universal approximators," *Neural networks*, vol. 2, no. 5, pp. 359–366, 1989.

[45] B. Kosko, "Fuzzy systems as universal approximators," *IEEE transactions on computers*, vol. 43, no. 11, pp. 1329–1333, 1994.

[46] I. Eyoh, R. John, and G. De Maere, "Time series forecasting with interval type-2 intuitionistic fuzzy logic systems," in *Fuzzy Systems (FUZZ-IEEE), 2017 IEEE International Conference on*, pp. 1–6, IEEE, 2017.

[47] I. Eyoh, R. John, and G. De Maere, "Extended kalman filter-based learning of interval type-2 intuitionistic fuzzy logic system," in *2017 IEEE International Conference on Systems, Man and Cybernetics*, pp. 728–733, 2017.

[48] R. Belohlavek, G. J. Klir, H. W. Lewis, and E. C. Way, "Concepts and fuzzy sets: Misunderstandings, misconceptions, and oversights," *International journal of approximate reasoning*, vol. 51, no. 1, pp. 23–34, 2009.

[49] L. A. Zadeh, "Toward extended fuzzy logic − a first step," *Fuzzy sets and systems*, vol. 160, no. 21, pp. 3175–3181, 2009.

[50] H. J. Zimmermann, "Fuzzy set theory," *Wiley Interdisciplinary Reviews: Computational Statistics*, vol. 2, no. 3, pp. 317–332, 2010.

[51] S. Logvinov, "Applications of the theory of fuzzy sets in complex control systems," in *Soft Computing and Measurements (SCM), 2017 XX IEEE International Conference on*, pp. 887–889, IEEE, 2017.

[52] S. Hatagar and S. Halase, "Three input–one output fuzzy logic control of washing machine," *vol*, vol. 4, pp. 57–62, 2015.

[53] O. Castillo, H. Neyoy, J. Soria, P. Melin, and F. Valdez, "A new approach for dynamic fuzzy logic parameter tuning in ant colony optimization and its application in fuzzy control of a mobile robot," *Applied soft computing*, vol. 28, pp. 150–159, 2015.

[54] O. Yazdanbakhsh and S. Dick, "Forecasting of multivariate time series via complex fuzzy logic," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2017.

[55] S. Askari and N. Montazerin, "A high-order multi-variable fuzzy time series forecasting algorithm based on fuzzy clustering," *Expert Systems with Applications*, vol. 42, no. 4, pp. 2121–2135, 2015.

[56] O. Yazdanbakhsh and S. Dick, "Time-series forecasting via complex fuzzy logic," in *Frontiers of Higher Order Fuzzy Sets*, pp. 147–165, Springer, 2015.

[57] I. . R. T. Geman, Oana ; Chiuchisan, "Application of adaptive neuro-fuzzy inference system for diabetes classification and prediction," in *2017 E-Health and Bioengineering Conference (EHB)*, pp. 639 – 642, IEEE, 2017.

[58] S. Taneja, B. Suri, H. Narwal, A. Jain, A. Kathuria, and S. Gupta, "A new approach for data classification using fuzzy logic," in *Cloud System and Big Data Engineering (Confluence), 2016 6th International Conference*, pp. 22–27, IEEE, 2016.

[59] R. M. Rodríguez, A. Labella, and L. Martínez, "An overview on fuzzy modelling of complex linguistic preferences in decision making," *International Journal of Computational Intelligence Systems*, vol. 9, no. sup1, pp. 81–94, 2016.

[60] S. Naim, H. Hagras, and J. M. Garibaldi, "A fuzzy logic based multi-criteria group decision making system for the assesement of umbilical cord acid-base balance," in *Fuzzy Systems (FUZZ-IEEE), 2012 IEEE International Conference on*, pp. 1–8, IEEE, 2012.

[61] N. Korenevskiy, "Application of fuzzy logic for decision-making in medical expert systems," *Biomedical Engineering*, vol. 49, no. 1, pp. 46–49, 2015.

[62] D. Chaturvedi, A. Sinha, and O. Malik, "Short term load forecast using fuzzy logic and wavelet transform integrated generalized neural network," *International Journal of Electrical Power & Energy Systems*, vol. 67, pp. 230–237, 2015.

[63] H. H. Çevik and M. Çunkaş, "Short-term load forecasting using fuzzy logic and anfis," *Neural Computing and Applications*, vol. 26, no. 6, pp. 1355–1367, 2015.

[64] A. Tehlan and V. Kumar, "Fuzzy logic based load forecasting," *International Journal of Applied Engineering Research*, vol. 11, no. 9, pp. 6625–6626, 2016.

[65] D. Sonika, S. Darshan, and K. Daljeet, "Long term load forecasting using fuzzy logic methodology," *International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering*, vol. 4, no. 6, pp. 5578–5585, 2015.

[66] S. Kar, S. Das, and P. K. Ghosh, "Applications of neuro fuzzy systems: A brief review and future outline," *Applied Soft Computing*, vol. 15, pp. 243–259, 2014.

[67] A. Mardani, A. Jusoh, and E. K. Zavadskas, "Fuzzy multiple criteria decision-making techniques and applications–two decades review from 1994 to 2014," *Expert Systems with Applications*, vol. 42, no. 8, pp. 4126–4148, 2015.

[68] C. Kahraman, S. C. Onar, and B. Oztaysi, "Fuzzy multicriteria decision-making: a literature review," *International Journal of Computational Intelligence Systems*, vol. 8, no. 4, pp. 637–666, 2015.

[69] L. Suganthi, S. Iniyan, and A. A. Samuel, "Applications of fuzzy logic in renewable energy systems–a review," *Renewable and Sustainable Energy Reviews*, vol. 48, pp. 585–607, 2015.

[70] G. Deschrijver and E. E. Kerre, "On the relationship between some extensions of fuzzy set theory," *Fuzzy sets and systems*, vol. 133, no. 2, pp. 227–235, 2003.

[71] H. Bustince, E. Barrenechea, M. Pagola, J. Fernandez, Z. Xu, B. Bedregal, J. Montero, H. Hagras, F. Herrera, and B. De Baets, "A historical account of types of fuzzy sets and their relationships," *IEEE Transactions on Fuzzy Systems*, vol. 24, no. 1, pp. 179–194, 2016.

[72] J. A. Goguen, "L-fuzzy sets," *Journal of mathematical analysis and applications*, vol. 18, no. 1, pp. 145–174, 1967.

[73] K. Atanassov and G. Gargov, "Interval valued intuitionistic fuzzy sets," *Fuzzy sets and systems*, vol. 31, no. 3, pp. 343–349, 1989.

[74] D. Julong, "Introduction to grey system theory," *The Journal of grey system*, vol. 1, no. 1, pp. 1–24, 1989.

[75] W.-L. Gau and D. J. Buehrer, "Vague sets," *IEEE transactions on systems, man, and cybernetics*, vol. 23, no. 2, pp. 610–614, 1993.

[76] V. Torra, "Hesitant fuzzy sets," *International Journal of Intelligent Systems*, vol. 25, no. 6, pp. 529–539, 2010.

[77] F. Smarandache, "Neutrosophic set–a generalization of the intuitionistic fuzzy set," *Multispace & Multistructure. Neutrosophic Transdisciplinarity*, p. 403, 2010.

[78] G. Deschrijver and E. E. Kerre, "On the position of intuitionistic fuzzy set theory in the framework of theories modelling imprecision," *Information Sciences*, vol. 177, no. 8, pp. 1860–1866, 2007.

[79] C. Cornelis, G. Deschrijver, and E. E. Kerre, "Implication in intuitionistic fuzzy and interval-valued fuzzy set theory: construction, classification, application," *International journal of approximate reasoning*, vol. 35, no. 1, pp. 55–95, 2004.

[80] H. Bustince and P. Burillo, "Vague sets are intuitionistic fuzzy sets," *Fuzzy sets and systems*, vol. 79, no. 3, pp. 403–405, 1996.

[81] C.-M. Own, "Switching between type-2 fuzzy sets and intuitionistic fuzzy sets: an application in medical diagnosis," *Applied Intelligence*, vol. 31, pp. 283–291, 2009.

[82] Y. Yang and R. John, "Grey sets and greyness," *Information Sciences*, vol. 185, no. 1, pp. 249–264, 2012.

[83] D. Dubois and H. Prade, "Interval-valued fuzzy sets, possibility theory and imprecise probability.," in *EUSFLAT Conf.*, pp. 314–319, 2005.

[84] H. Bustince, J. Fernandez, H. Hagras, F. Herrera, M. Pagola, and E. Barrenechea, "Interval type-2 fuzzy sets are generalization of interval-valued fuzzy sets: toward a wider view on their relationship," *IEEE Transactions on Fuzzy Systems*, vol. 23, no. 5, pp. 1876–1882, 2015.

[85] S. Coupland and R. John, "A fast geometric method for defuzzification of type-2 fuzzy sets," *IEEE Transactions on Fuzzy Systems*, vol. 16, no. 4, pp. 929–941, 2008.

[86] J. M. Mendel, "Type-2 fuzzy sets and systems: An overview [corrected reprint]," *IEEE computational intelligence magazine*, vol. 2, no. 2, pp. 20–29, 2007.

[87] A. Niewiadomski, "Interval-valued and interval type-2 fuzzy sets: A subjective comparison," in *Fuzzy systems conference, 2007. fuzz-IEEE 2007. IEEE international*, pp. 1–6, IEEE, 2007.

[88] J. Casillas, O. Cordón, F. H. Triguero, and L. Magdalena, *Accuracy improvements in linguistic fuzzy modeling*, vol. 129. Springer, 2013.

[89] J. Castro, O. Castillo, and L. G. Martinez, "Interval type-2 fuzzy logic toolbox," *Engineering Letters*, vol. 15, no. 1, 2007.

[90] Q. Liang and J. M. Mendel, "An introduction to type-2 tsk fuzzy logic systems," in *Proceedings IEEE International Fuzzy Systems Conference (FUZZ-IEEE)*, vol. 3, pp. 1534–1539, 1999.

[91] T. Takagi and M. Sugeno, "Fuzzy identification of systems and its applications to modeling and control," *Systems, Man and Cybernetics, IEEE Transactions on*, no. 1, pp. 116–132, 1985.

[92] M. Sugeno and G. Kang, "Structure identification of fuzzy model," *Fuzzy sets and systems*, vol. 28, no. 1, pp. 15–33, 1988.

[93] S. Greenfield and F. Chiclana, "Accuracy and complexity evaluation of defuzzification strategies for the discretised interval type-2 fuzzy set," *International Journal of Approximate Reasoning*, vol. 54, no. 8, pp. 1013–1033, 2013.

[94] N. N. Karnik and J. M. Mendel, "Centroid of a type-2 fuzzy set," *Information Sciences*, vol. 132, no. 1, pp. 195–220, 2001.

[95] Y.-Y. Lin, S.-H. Liao, J.-Y. Chang, and C.-T. Lin, "Simplified interval type-2 fuzzy neural networks," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 25, no. 5, pp. 959–969, 2014.

[96] A. D. Torshizi, M. H. F. Zarandi, and H. Zakeri, "On type-reduction of type-2 fuzzy sets: A review," *Applied Soft Computing*, vol. 27, pp. 614–627, 2015.

[97] S. Greenfield and F. Chiclana, "Type-reduction of the discretised interval type-2 fuzzy set: Approaching the continuous case through progressively finer discretisation.," 2011.

[98] M. B. Begian, W. W. Melek, and J. M. Mendel, "Parametric design of stable type-2 tsk fuzzy systems," in *IEEE Annual Meeting of the North American Fuzzy Information Processing Society, (NAFIPS)*, pp. 1–6, 2008.

[99] Q. Liang and J. M. Mendel, "Equalization of nonlinear time-varying channels using type-2 fuzzy adaptive filters," *IEEE Transactions on Fuzzy Systems*, vol. 8, no. 5, pp. 551–563, 2000.

[100] H. Wu and J. M. Mendel, "Uncertainty bounds and their use in the design of interval type-2 fuzzy logic systems," *IEEE Transactions on fuzzy systems*, vol. 10, no. 5, pp. 622–639, 2002.

[101] D. Wu and W. W. Tan, "Computationally efficient type-reduction strategies for a type-2 fuzzy logic controller," in *Fuzzy Systems, 2005. FUZZ'05. The 14th IEEE International Conference on*, pp. 353–358, IEEE, 2005.

[102] C. Li, J. Yi, and D. Zhao, "A novel type-reduction method for interval type-2 fuzzy logic systems," in *Fuzzy Systems and Knowledge Discovery, 2008. FSKD'08. Fifth International Conference on*, vol. 1, pp. 157–161, IEEE, 2008.

[103] M. Nie and W. W. Tan, "Towards an efficient type-reduction method for interval type-2 fuzzy logic systems," in *Fuzzy Systems, 2008. FUZZ-IEEE 2008.(IEEE World Congress on Computational Intelligence). IEEE International Conference on*, pp. 1425–1432, IEEE, 2008.

[104] C. Ulu, M. Güzelkaya, and I. Eksin, "A new type reduction method for piecewise linear interval type 2 fuzzy sets," in *2nd World Conference on Soft Computing, W Con SC*, vol. 12, pp. 494–498, 2012.

[105] C. Ulu, M. Guzelkaya, and I. Eksin, "A closed form type reduction method for piecewise linear interval type-2 fuzzy sets," *International Journal of Approximate Reasoning*, vol. 54, no. 9, pp. 1421–1433, 2013.

[106] T.-C. Lu, "Genetic-algorithm-based type reduction algorithm for interval type-2 fuzzy logic controllers," *Engineering Applications of Artificial Intelligence*, vol. 42, pp. 36–44, 2015.

[107] D. Wu, "An overview of alternative type-reduction approaches for reducing the computational cost of interval type-2 fuzzy logic controllers," in *Fuzzy Systems (FUZZ-IEEE), 2012 IEEE International Conference on*, pp. 1–8, IEEE, 2012.

[108] R. John and S. Coupland, "Type-2 fuzzy logic: A historical view," *IEEE computational intelligence magazine*, vol. 2, no. 1, pp. 57–62, 2007.

[109] G. J. Klir, *Uncertainty and information: foundations of generalized information theory.* John Wiley & Sons, 2005.

[110] G. J. Klir and T. A. Folger, "Fuzzy sets, uncertainty, and information," 1988.

[111] S. Hassan, M. A. Khanesar, E. Kayacan, J. Jaafar, and A. Khosravi, "Optimal design of adaptive type-2 neuro-fuzzy systems: A review," *Applied Soft Computing*, vol. 44, pp. 134–143, 2016.

[112] F. Guély and P. Siarry, "Gradient descent method for optimizing various fuzzy rule bases," in *Fuzzy Systems, 1993., Second IEEE International Conference on*, pp. 1241–1246, IEEE, 1993.

[113] J. Kim and N. Kasabov, "Hyfis: adaptive neuro-fuzzy inference systems and their application to nonlinear dynamical systems," *Neural Networks*, vol. 12, no. 9, pp. 1301–1319, 1999.

[114] A. Habbi and M. Zelmat, "An improved self-tuning mechanism of fuzzy control by gradient descent method," in *Proceedings of the 17th European Simulation Multiconference, Nottingham, UK*, 2003.

[115] D. Simon, "Training fuzzy systems with the extended kalman filter," *Fuzzy sets and systems*, vol. 132, no. 2, pp. 189–199, 2002.

[116] J. R. Castro, O. Castillo, P. Melin, and A. Rodríguez-Díaz, "A hybrid learning algorithm for a class of interval type-2 fuzzy neural networks," *Information Sciences*, vol. 179, no. 13, pp. 2175–2193, 2009.

[117] E. Kayacan, E. Kayacan, and M. A. Khanesar, "Identification of nonlinear dynamic systems using type-2 fuzzy neural networksa novel learning algorithm and a comparative study," *IEEE Transactions on Industrial Electronics*, vol. 62, no. 3, pp. 1716–1724, 2015.

[118] J. M. Mendel, "Computing derivatives in interval type-2 fuzzy logic systems," *IEEE Transactions on Fuzzy Systems*, vol. 12, no. 1, pp. 84–98, 2004.

[119] X. Wang, "Method of steepest descent and its applications," *IEEE Microwave and Wireless Components Letters*, vol. 12, pp. 24–26, 2008.

[120] C.-H. Wang, C.-S. Cheng, and T.-T. Lee, "Dynamical optimal training for interval type-2 fuzzy neural network (t2fnn)," *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 34, no. 3, pp. 1462–1477, 2004.

[121] C.-F. Juang and Y.-W. Tsao, "A self-evolving interval type-2 fuzzy neural network with online structure and parameter learning," *IEEE Transactions on Fuzzy Systems*, vol. 16, no. 6, pp. 1411–1424, 2008.

[122] Y.-Y. Lin, J.-Y. Chang, and C.-T. Lin, "A tsk-type-based self-evolving compensatory interval type-2 fuzzy neural network (TSCIT2FNN) and its applications," *IEEE Transactions on Industrial Electronics*, vol. 61, no. 1, pp. 447–459, 2014.

[123] O. Castillo and P. Melin, "A review on the design and optimization of interval type-2 fuzzy controllers," *Applied Soft Computing*, vol. 12, no. 4, pp. 1267–1278, 2012.

[124] G. V. Puskorius and L. A. Feldkamp, "Neurocontrol of nonlinear dynamical systems with kalman filter trained recurrent networks," *IEEE Transactions on neural networks*, vol. 5, no. 2, pp. 279–297, 1994.

[125] M. A. Khanesar, E. Kayacan, M. Teshnehlab, and O. Kaynak, "Extended kalman filter based learning algorithm for type-2 fuzzy logic systems and its experimental evaluation," *IEEE Transactions on Industrial Electronics*, vol. 59, no. 11, pp. 4443–4455, 2012.

[126] C.-F. Juang, R.-B. Huang, and Y.-Y. Lin, "A recurrent self-evolving interval type-2 fuzzy neural network for dynamic system processing," *IEEE Transactions on Fuzzy Systems*, vol. 17, no. 5, pp. 1092–1105, 2009.

[127] C. Slim, "Neuro-fuzzy network based on extended kalman filtering for financial time series," *World Academy of Science, Engineering and Technology*, vol. 22, pp. 134–139, 2006.

[128] F. Yihong, L. Weimin, Z. Xiaoguang, and X. Xin, "Threat assessment based on adaptive intuitionistic fuzzy neural network," in *Fourth IEEE International Symposium on Computational Intelligence and Design (ISCID),*, vol. 1, pp. 262–265, 2011.

[129] C. M. Takenga, K. R. Anne, K. Kyamakya, and J. C. Chedjou, "Comparison of gradient descent method, kalman filtering and decoupled kalman in training neural

networks used for fingerprint-based positioning," in *Vehicular Technology Conference, 2004. VTC2004-Fall. 2004 IEEE 60th*, vol. 6, pp. 4146–4150, IEEE, 2004.

[130] A. Najafi, A. Amirkhani, K. Mohammadi, and A. Naimi, "A novel soft computing method based on interval type-2 fuzzy logic for classification of celiac disease," in *Biomedical Engineering and 2016 1st International Iranian Conference on Biomedical Engineering (ICBME), 2016 23rd Iranian Conference on*, pp. 257–262, IEEE, 2016.

[131] P. Herman, G. Prasad, and T. M. McGinnity, "Design and on-line evaluation of type-2 fuzzy logic system-based framework for handling uncertainties in bci classification," in *Engineering in Medicine and Biology Society, 2008. EMBS 2008. 30th Annual International Conference of the IEEE*, pp. 4242–4245, IEEE, 2008.

[132] J. Zarei, M. M. Arefi, and H. Hassani, "Bearing fault detection based on interval type-2 fuzzy logic systems for support vector machines," in *Modeling, Simulation, and Applied Optimization (ICMSAO), 2015 6th International Conference on*, pp. 1–6, IEEE, 2015.

[133] B. Yao, H. Hagras, J. J. Lepley, R. Peall, and M. Butler, "An evolutionary optimization based interval type-2 fuzzy classification system for human behaviour recognition and summarisation," in *Systems, Man, and Cybernetics (SMC), 2016 IEEE International Conference on*, pp. 004706–004711, IEEE, 2016.

[134] D. Enke, M. Grauer, and N. Mehdiyev, "Stock market prediction with multiple regression, fuzzy type-2 clustering and neural networks," *Procedia Computer Science*, vol. 6, pp. 201–206, 2011.

[135] V. Sumati, C. Patvardhan, and V. M. Swarup, "Application of interval type-2 subsethood neural fuzzy inference system in classification," in *Humanitarian Technology Conference (R10-HTC), 2016 IEEE Region 10*, pp. 1–6, IEEE, 2016.

[136] D. Bernardo, H. Hagras, and E. Tsang, "A genetic type-2 fuzzy logic based system for financial applications modelling and prediction," in *Fuzzy Systems (FUZZ), 2013 IEEE International Conference on*, pp. 1–8, IEEE, 2013.

[137] W. Song and H. Hagras, "A big-bang big-crunch type-2 fuzzy logic based system for soccer video scene classification," in *Fuzzy Systems (FUZZ-IEEE), 2016 IEEE International Conference on*, pp. 2059–2066, IEEE, 2016.

[138] A. Khosravi and S. Nahavandi, "Load forecasting using interval type-2 fuzzy logic systems: Optimal type reduction," *IEEE Transactions on Industrial Informatics*, vol. 10, no. 2, pp. 1055–1063, 2014.

[139] O. Mendoza, P. Melín, and O. Castillo, "Interval type-2 fuzzy logic and modular neural networks for face recognition applications," *Applied Soft Computing*, vol. 9, no. 4, pp. 1377–1387, 2009.

[140] P. Melin, "Interval type-2 fuzzy logic applications in image processing and pattern recognition," in *Granular Computing (GrC), 2010 IEEE International Conference on*, pp. 728–731, IEEE, 2010.

[141] S. Yadav and V. P. Vishwakarma, "Interval type-2 fuzzy based pixel wise information extraction: An improved approach to face recognition," in *Computational Techniques in Information and Communication Technologies (ICCTICT), 2016 International Conference on*, pp. 409–414, IEEE, 2016.

[142] K. Qin, L. Kong, Y. Liu, and Q. Xiao, "Sea surface temperature clustering based on type-2 fuzzy theory," in *Geoinformatics, 2010 18th International Conference on*, pp. 1–5, IEEE, 2010.

[143] L. Yu, J. Xiao, and G. Zheng, "Robust interval type-2 possibilistic c-means clustering and its application for fuzzy modeling," in *Fuzzy Systems and Knowledge Discovery, 2009. FSKD'09. Sixth International Conference on*, vol. 4, pp. 360–365, IEEE, 2009.

[144] J. Cuevas-Martinez, A. Yuste-Delgado, and A. Trivino-Cabrera, "Cluster head enhanced election type-2 fuzzy algorithm for wireless sensor networks," *IEEE Communications Letters*, 2017.

[145] W.-D. Kim, S.-K. Oh, and K. Seo, "Design of interval type-2 fcm-based neural networks," in *Soft Computing and Intelligent Systems (SCIS) and 17th International Symposium on Advanced Intelligent Systems, 2016 Joint 8th International Conference on*, pp. 759–763, IEEE, 2016.

[146] D. A. R. Wati, "Interval type-2 fuzzy logic controller for multi input multi output system: A shower system case study," in *Systems, Process and Control (ICSPC), 2016 IEEE Conference on*, pp. 154–159, IEEE, 2016.

[147] O. Linda and M. Manic, "Uncertainty modeling with interval type-2 fuzzy logic systems in mobile robotics," in *IECON 2011-37th Annual Conference on IEEE Industrial Electronics Society*, pp. 2441–2446, IEEE, 2011.

[148] M. Ri, J. Huang, S. Ri, H. Yun, and C. Kim, "Design of interval type-2 fuzzy logic controller for mobile wheeled inverted pendulum," in *Intelligent Control and Automation (WCICA), 2016 12th World Congress on*, pp. 535–540, IEEE, 2016.

[149] Y. Bai and D. Wang, "On the comparison of an interval type-2 fuzzy interpolation system and other interpolation methods used in industrial modeless robotic calibrations," in *Computational Intelligence and Virtual Environments for Measurement Systems and Applications (CIVEMSA), 2016 IEEE International Conference on*, pp. 1–6, IEEE, 2016.

[150] S.-H. Cheng, S.-M. Chen, and Z.-C. Huang, "A novel autocratic decision making method using group recommendations based on ranking interval type-2 fuzzy sets," in *Advanced Computational Intelligence (ICACI), 2016 Eighth International Conference on*, pp. 84–88, IEEE, 2016.

[151] S.-M. Chen, M.-W. Yang, L.-W. Lee, and S.-W. Yang, "Fuzzy multiple attributes group decision-making based on ranking interval type-2 fuzzy sets," *Expert Systems with Applications*, vol. 39, no. 5, pp. 5295–5308, 2012.

[152] S.-M. Chen and L.-W. Lee, "Fuzzy multiple attributes group decision-making based on the interval type-2 topsis method," *Expert systems with applications*, vol. 37, no. 4, pp. 2790–2798, 2010.

[153] C.-H. Chen, T.-P. Hong, and Y. Li, "Fuzzy association rule mining with type-2 membership functions," in *Intelligent Information and Database Systems*, pp. 128–134, Springer, 2015.

[154] O. Castillo, L. Amador-Angulo, J. R. Castro, and M. Garcia-Valdez, "A comparative study of type-1 fuzzy logic systems, interval type-2 fuzzy logic systems and

generalized type-2 fuzzy logic systems in control problems," *Information Sciences*, vol. 354, pp. 257–274, 2016.

[155] C. W. H. Hagras, "Novel methods for the design of general type-2 fuzzy sets based on device characteristics and linguistic labels surveys," *Proc. Int. Fuzzy Syst. Assoc. World Congr*, pp. 537–543, 2009.

[156] T. Nguyen, A. Khosravi, D. Creighton, and S. Nahavandi, "Medical data classification using interval type-2 fuzzy logic system and wavelets," *Applied Soft Computing*, vol. 30, pp. 812–822, 2015.

[157] P. Ejegwa, A. Akubo, and O. Joshua, "Intuitionistic fuzzy set and its application in career determination via normalized euclidean distance method," *European Scientific Journal*, vol. 10, no. 15, 2014.

[158] E. Barrenechea, M. Pagola, J. Fernandez, and J. Sanz, "Generalized atanassovs intuitionistic fuzzy index. construction method," in *IFSA-EUSFLAT conf.*, pp. 478–482, 2009.

[159] J. Ye, "Fuzzy decision-making method based on the weighted correlation coefficient under intuitionistic fuzzy environment," *European Journal of Operational Research*, vol. 205, no. 1, pp. 202–204, 2010.

[160] D. Yu and S. Shi, "Researching the development of atanassov intuitionistic fuzzy set: Using a citation network analysis," *Applied Soft Computing*, vol. 32, pp. 189–198, 2015.

[161] L. Dengfeng and C. Chuntian, "New similarity measures of intuitionistic fuzzy sets and application to pattern recognitions," *Pattern recognition letters*, vol. 23, no. 1, pp. 221–225, 2002.

[162] D.-F. Li, "Multiattribute decision making models and methods using intuitionistic fuzzy sets," *Journal of computer and System Sciences*, vol. 70, no. 1, pp. 73–85, 2005.

[163] Z. Xu, "Intuitionistic preference relations and their application in group decision making," *Information sciences*, vol. 177, no. 11, pp. 2363–2379, 2007.

[164] E. Szmidt and J. Kacprzyk, "Intuitionistic fuzzy sets in some medical applications," in *Computational Intelligence. Theory and Applications*, pp. 148–151, Springer, 2001.

[165] V. Olej and P. Hájek, "If-inference systems design for prediction of ozone time series: the case of pardubice micro-region," in *Artificial Neural Networks–ICANN 2010*, pp. 1–11, Springer, 2010.

[166] C. Cornelis, M. De Cock, and E. Kerre, "Assessing degrees of possibility and certainty within an unreliable environment," in *Proceedings of Fourth International Conference on Recent Advances in Soft Computing*, pp. 194–199, 2002.

[167] T. Shinoj and J. Sunil, "Intuitionistic fuzzy multisets and its application in medical diagnosis," *World Academy of Science, Engineering and Technology*, vol. 6, no. 1, pp. 1418–1421, 2012.

[168] E. Szmidt and J. Kacprzyk, "Medical diagnostic reasoning using a similarity measure for intuitionistic fuzzy sets," *Note on IFS*, vol. 10, no. 4, pp. 61–69, 2004.

[169] T. Chaira, "Medical image thresholding scheme using atanassov intuitionistic fuzzy set," in *Methods and Models in Computer Science, 2009. ICM2CS 2009. Proceeding of International Conference on*, pp. 1–5, IEEE, 2009.

[170] V. Khatibi and G. A. Montazer, "Intuitionistic fuzzy set vs. fuzzy set application in medical pattern recognition," *Artificial Intelligence in Medicine*, vol. 47, no. 1, pp. 43–52, 2009.

[171] S. K. De, R. Biswas, and A. R. Roy, "An application of intuitionistic fuzzy sets in medical diagnosis," *Fuzzy sets and Systems*, vol. 117, no. 2, pp. 209–213, 2001.

[172] P. Hájek and V. Olej, "Intuitionistic fuzzy neural network: The case of credit scoring using text information," in *Engineering Applications of Neural Networks*, pp. 337–346, Springer, 2015.

[173] P. Hájek and V. Olej, "Adaptive intuitionistic fuzzy inference systems of takagi-sugeno type for regression problems," in *Artificial Intelligence Applications and Innovations*, pp. 206–216, Springer, 2012.

[174] V. Olej and P. Hájek, "Comparison of fuzzy operators for if-inference systems of takagi-sugeno type in ozone prediction," in *Artificial Intelligence Applications and Innovations*, pp. 92–97, Springer, 2011.

[175] P. Intarapaiboon, "An application of intuitionistic fuzzy sets in text classification," in *Information Science, Electronics and Electrical Engineering (ISEEE), 2014 International Conference on*, vol. 1, pp. 604–608, IEEE, 2014.

[176] E. Szmidt and J. Kacprzyk, "Using intuitionistic fuzzy sets in text categorization," in *International Conference on Artificial Intelligence and Soft Computing*, pp. 351–362, Springer, 2008.

[177] M. Akram, S. Shahzad, A. Butt, and A. Khaliq, "Intuitionistic fuzzy logic control for heater fans," *Mathematics in Computer Science*, vol. 7, no. 3, pp. 367–378, 2013.

[178] M. Akram, S. Habib, and I. Javed, "Intuitionistic fuzzy logic control for washing machines," *Indian Journal of Science and Technology*, vol. 7, no. 5, pp. 654–661, 2014.

[179] P. Hajek and V. Olej, "Defuzzification methods in intuitionistic fuzzy inference systems of takagi-sugeno type: The case of corporate bankruptcy prediction," in *11th IEEE International Conference on Fuzzy Systems and Knowledge Discovery (FSKD), 2014*, pp. 232–236.

[180] E. Szmidt and J. Kacprzyk, "Using intuitionistic fuzzy sets in group decision making," *Control and cybernetics*, vol. 31, pp. 1055–1057, 2002.

[181] L. Lin, X.-H. Yuan, and Z.-Q. Xia, "Multicriteria fuzzy decision-making methods based on intuitionistic fuzzy sets," *Journal of computer and System Sciences*, vol. 73, no. 1, pp. 84–88, 2007.

[182] H.-W. Liu and G.-J. Wang, "Multi-criteria decision-making methods based on intuitionistic fuzzy sets," *European Journal of Operational Research*, vol. 179, no. 1, pp. 220–233, 2007.

[183] E. Sotirova, K. Atanassov, A. Shannon, T. Kim, M. Krawczak, P. Melo-Pinto, and B. Riečan, "Intuitionistic fuzzy evaluations for analysis of a student's knowledge of mathematics in university e-learning courses," in *Intelligent Systems (IS), 2016 IEEE 8th International Conference on*, pp. 535–537, IEEE, 2016.

[184] S. Sotirov, E. Sotirova, P. Melin, O. Castillo, and K. T. Atanassov, "Modular neural network preprocessing procedure with intuitionistic fuzzy intercriteria analysis method.," in *FQAS*, pp. 175–186, 2015.

[185] S. Naim and H. Hagras, "A hybrid approach for multi-criteria group decision making based on interval type-2 fuzzy logic and intuitionistic fuzzy evaluation," in *Fuzzy Systems (FUZZ-IEEE), 2012 IEEE International Conference on*, pp. 1–8, IEEE, 2012.

[186] B. C. Cng, T. H. Anh, and B. D. Hi, "Some operations on type-2 intuitionistic fuzzy sets.," *Journal of Computer Science and Cybernetics*, vol. 28, no. 3, pp. 274–283, 2012.

[187] D. K. Jana, "Novel arithmetic operations on type-2 intuitionistic fuzzy and its applications to transportation problem," *Pacific Science Review A: Natural Science and Engineering*, vol. 18, no. 3, pp. 178–189, 2016.

[188] S. Singh and H. Garg, "Distance measures between type-2 intuitionistic fuzzy sets and their application to multicriteria decision-making process," *Applied Intelligence*, vol. 46, no. 4, pp. 788–799, 2017.

[189] D. D. Nguyen, L. T. Ngo, and L. T. Pham, "Interval type-2 fuzzy c-means clustering using intuitionistic fuzzy sets," in *IEEE Third World Congress on Information and Communication Technologies (WICT)*, pp. 299–304, 2013.

[190] T. Van Nghiem, D. D. Nguyen, and L. T. Ngo, "Intuitionistic type-2 fuzzy set approach to image thresholding," in *IEEE International Conference of Soft Computing and Pattern Recognition (SoCPaR)*, pp. 207–212, 2013.

[191] S. Naim, H. Hagras, and A. Bilgin, "Employing an interval type-2 fuzzy logic and hesitation index in a multi criteria group decision making system for lighting level selection in an intelligent environment," in *Advances in Type-2 Fuzzy Logic Systems (T2FUZZ), 2013 IEEE Symposium on*, pp. 1–8, IEEE, 2013.

[192] E. Hisdal, "The if then else statement and interval-valued fuzzy sets of higher type," *International Journal of Man-Machine Studies*, vol. 15, no. 4, pp. 385–455, 1981.

[193] D. Wu, "Twelve considerations in choosing between gaussian and trapezoidal membership functions in interval type-2 fuzzy logic controllers," in *Fuzzy Systems (FUZZ-IEEE), 2012 IEEE International Conference on*, pp. 1–8, IEEE, 2012.

[194] J.-S. R. Jang, "ANFIS: Adaptive-network-based fuzzy inference system," *IEEE Transactions on Systems, Man and Cybernetics*, vol. 23, no. 3, pp. 665–685, 1993.

[195] R. I. John and C. Czarnecki, "A type 2 adaptive fuzzy inferencing system," in *IEEE International Conference on Systems, Man, and Cybernetics, 1998.*, vol. 2, pp. 2068–2073.

[196] J. H. Friedman, "Multivariate adaptive regression splines," *The annals of statistics*, pp. 1–67, 1991.

[197] P. Picton, "What is a neural network?," in *Introduction to Neural Networks*, pp. 1–12, Springer, 1994.

[198] `http://www.keel.es/`. Accessed: 2016-06-23.

[199] C.-F. Juang, R.-B. Huang, and W.-Y. Cheng, "An interval type-2 fuzzy-neural network with support-vector regression for noisy regression problems," *IEEE Transactions on Fuzzy Systems*, vol. 18, no. 4, pp. 686–699, 2010.

[200] J. G. Carney and P. Cunningham, "Tuning diversity in bagged ensembles," *International Journal of Neural Systems*, vol. 10, no. 04, pp. 267–279, 2000.

[201] E. W. Lee, C. P. Lim, R. K. Yuen, and S. Lo, "A hybrid neural network model for noisy data regression," *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 34, no. 2, pp. 951–960, 2004.

[202] C.-F. Juang and C.-T. Lin, "An online self-constructing neural fuzzy inference network and its applications," *IEEE Transactions on Fuzzy Systems*, vol. 6, no. 1, pp. 12–32, 1998.

[203] O. Cordón, F. Herrera, and L. Sánchez, "Solving electrical distribution problems using hybrid evolutionary data analysis techniques," *Applied Intelligence*, vol. 10, no. 1, pp. 5–24, 1999.

[204] O. Cordón, F. Herrera, and P. Villar, "Generating the knowledge base of a fuzzy rule-based system by the genetic learning of the data base," *IEEE Transactions on fuzzy systems*, vol. 9, no. 4, pp. 667–674, 2001.

[205] O. Cordón, F. Herrera, L. Magdalena, and P. Villar, "A genetic learning process for the scaling factors, granularity and contexts of the fuzzy rule-based system data base," *Information Sciences*, vol. 136, no. 1, pp. 85–107, 2001.

[206] O. Cordón, F. Herrera, and I. Zwir, "Linguistic modeling by hierarchical systems of linguistic rules," *Fuzzy Systems, IEEE Transactions on*, vol. 10, no. 1, pp. 2–20, 2002.

[207] M. Almaraashi, R. John, A. Hopgood, and S. Ahmadi, "Learning of interval and general type-2 fuzzy logic systems using simulated annealing: Theory and practice," *Information Sciences*, vol. 360, pp. 21–42, 2016.

[208] M. Almaraashi, "Learning of type-2 fuzzy logic systems using simulated annealing.," 2012.

[209] C. Chen, R. John, J. Twycross, and J. M. Garibaldi, "Type-1 and interval type-2 anfis: A comparison," in *Fuzzy Systems (FUZZ-IEEE), 2017 IEEE International Conference on*, pp. 1–6, IEEE, 2017.

[210] K. B. Cho and B. H. Wang, "Radial basis function based adaptive fuzzy systems and their applications to system identification and prediction," *Fuzzy sets and systems*, vol. 83, no. 3, pp. 325–339, 1996.

[211] H. Dhahri and A. M. Alimi, "The modified differential evolution and the RBF (MDE-RBF) neural network for time series prediction," in *IEEE International Joint Conference on Neural Networks (IJCNN), 2006.*, pp. 2938–2943.

[212] H. Gu and H. Wang, "Fuzzy prediction of chaotic time series based on singular value decomposition," *Applied Mathematics and Computation*, vol. 185, no. 2, pp. 1171–1185, 2007.

[213] D. Kim and C. Kim, "Forecasting time series with genetic fuzzy predictor ensemble," *IEEE Transactions on Fuzzy Systems*, vol. 5, no. 4, pp. 523–535, 1997.

[214] J. Soto, P. Melin, and O. Castillo, "A new approach for time series prediction using ensembles of ANFIS models with interval type-2 and type-1 fuzzy integrators," in *IEEE Conference on Computational Intelligence for Financial Engineering & Economics (CIFEr)*, pp. 68–73, 2013.

[215] `http://www.ngdc.noaa.gov/stp/solar/ssndata.html`. Accessed: 2016-07-30.

[216] H. Tong and K. S. Lim, "Threshold autoregression, limit cycles and cyclical data," *Journal of the Royal Statistical Society. Series B (Methodological)*, pp. 245–292, 1980.

[217] A. S. Weigend, B. A. Huberman, and D. E. Rumelhart, "Predicting the future: A connectionist approach," *International journal of neural systems*, vol. 1, no. 03, pp. 193–209, 1990.

[218] C. Svarer, L. K. Hansen, and J. Larsen, "On design and evaluation of tapped-delay neural network architectures," in *IEEE International Conference on Neural Networks*, pp. 46–51, 1993.

[219] J. R. McDonnell and D. Waagen, "Evolving recurrent perceptrons for time-series modeling," *IEEE Transactions on Neural Networks*, vol. 5, no. 1, pp. 24–38, 1994.

[220] R. A. Aliev, B. Guirimov, B. Fazlollahi, and R. R. Aliev, "Evolutionary algorithm-based learning of fuzzy neural networks. part 2: Recurrent fuzzy neural networks," *Fuzzy Sets and Systems*, vol. 160, no. 17, pp. 2553–2566, 2009.

[221] A. Hussain, "A new neural network structure for temporal signal processing," in *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, vol. 4, pp. 3341–3344, 1997.

[222] S. Yilmaz and Y. Oysal, "Fuzzy wavelet neural network models for prediction and identification of dynamical systems," *IEEE transactions on neural networks*, vol. 21, no. 10, pp. 1599–1609, 2010.

[223] M. Abdollahzade, A. Miranian, H. Hassani, and H. Iranmanesh, "A new hybrid enhanced local linear neuro-fuzzy model based on the optimized singular spectrum analysis and its application for nonlinear and chaotic time series forecasting," *Information Sciences*, vol. 295, pp. 107–125, 2015.

[224] Hyndman, R. J. Time Series Data Library. `http://data.is/TSDLdemo`. Accessed: 2016-07-30.

[225] F. M. Pouzols and A. Lendasse, "Evolving fuzzy optimally pruned extreme learning machine for regression problems," *Evolving Systems*, vol. 1, no. 1, pp. 43–58, 2010.

[226] N. K. Kasabov and Q. Song, "Denfis: dynamic evolving neural-fuzzy inference system and its application for time-series prediction," *IEEE transactions on Fuzzy Systems*, vol. 10, no. 2, pp. 144–154, 2002.

[227] P. P. Angelov and D. P. Filev, "An approach to online identification of takagi-sugeno fuzzy models," *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 34, no. 1, pp. 484–498, 2004.

[228] H.-J. Rong, G.-B. Huang, N. Sundararajan, and P. Saratchandran, "Online sequential fuzzy extreme learning machine for function approximation and classification problems," *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 39, no. 4, pp. 1067–1072, 2009.

[229] G. P. Zhang, "Time series forecasting using a hybrid arima and neural network model," *Neurocomputing*, vol. 50, pp. 159–175, 2003.

[230] M. Khashei and M. Bijari, "A new class of hybrid models for time series forecasting," *Expert Systems with Applications*, vol. 39, no. 4, pp. 4344–4357, 2012.

[231] L. Wang, Y. Zeng, and T. Chen, "Back propagation neural network with adaptive differential evolution algorithm for time series forecasting," *Expert Systems with Applications*, vol. 42, no. 2, pp. 855–863, 2015.

[232] M. Khashei and M. Bijari, "An artificial neural network (p, d, q) model for timeseries forecasting," *Expert Systems with applications*, vol. 37, no. 1, pp. 479–489, 2010.

[233] C. H. Aladag, E. Egrioglu, and C. Kadilar, "Forecasting nonlinear time series with a hybrid methodology," *Applied Mathematics Letters*, vol. 22, no. 9, pp. 1467–1470, 2009.

[234] Y. Kajitani, K. W. Hipel, and A. I. McLeod, "Forecasting nonlinear time series with feed-forward neural networks: a case study of canadian lynx data," *Journal of Forecasting*, vol. 24, no. 2, pp. 105–117, 2005.

[235] M. Khashei and M. Bijari, "A novel hybridization of artificial neural networks and arima models for time series forecasting," *Applied Soft Computing*, vol. 11, no. 2, pp. 2664–2675, 2011.

[236] E. Bas, "The training of multiplicative neuron model based artificial neural networks with differential evolution algorithm for forecasting," *Journal of Artificial Intelligence and Soft Computing Research*, vol. 6, no. 1, pp. 5–11, 2016.

[237] R. Adhikari and R. Agrawal, "A homogeneous ensemble of artificial neural networks for time series forecasting," *arXiv preprint arXiv:1302.6210*, 2013.

[238] M. Gan and H. Peng, "Stability analysis of rbf network-based state-dependent autoregressive model for nonlinear time series," *Applied Soft Computing*, vol. 12, no. 1, pp. 174–181, 2012.

[239] R. Samsudin, P. Saad, and A. Shabri, "A hybrid gmdh and least squares support vector machines in time series forecasting," *Neural Network World*, vol. 21, no. 3, p. 251, 2011.

[240] U. Yolcu, E. Egrioglu, and C. H. Aladag, "A new linear & nonlinear artificial neural network model for time series forecasting," *Decision support systems*, vol. 54, no. 3, pp. 1340–1347, 2013.

[241] `http://www-psych.stanford.edu/~andreas/Time-Series/SantaFe.html`. Accessed: 2016-07-21.

[242] S. Singh, "Noise impact on time-series forecasting using an intelligent pattern matching technique," *Pattern Recognition*, vol. 32, no. 8, pp. 1389–1398, 1999.

[243] R. H. Abiyev and O. Kaynak, "Type 2 fuzzy neural structure for identification and control of time-varying plants," *IEEE Transactions on Industrial Electronics*, vol. 57, no. 12, pp. 4147–4159, 2010.

[244] J.-H. Chiang and P.-Y. Hao, "Support vector learning mechanism for fuzzy rule-based modeling: a new approach," *IEEE Transactions on Fuzzy Systems*, vol. 12, no. 1, pp. 1–12, 2004.

[245] M. J. Gacto, M. Galende, R. Alcalá, and F. Herrera, "Metsk-hd e: A multiobjective evolutionary algorithm to learn accurate tsk-fuzzy systems in high-dimensional and large-scale regression problems," *Information Sciences*, vol. 276, pp. 63–79, 2014.

[246] R. Alcalá, M. J. Gacto, and F. Herrera, "A fast and scalable multiobjective genetic fuzzy system for linguistic fuzzy modeling in high-dimensional regression problems," *IEEE Transactions on Fuzzy Systems*, vol. 19, no. 4, pp. 666–681, 2011.

[247] P. Pulkkinen and H. Koivisto, "A dynamically constrained multiobjective genetic fuzzy system for regression problems," *IEEE Transactions on Fuzzy Systems*, vol. 18, no. 1, pp. 161–177, 2010.

[248] M. J. Gacto, R. Alcalá, and F. Herrera, "Adaptation and application of multi-objective evolutionary algorithms for rule reduction and parameter tuning of fuzzy rule-based systems," *Soft Computing*, vol. 13, no. 5, pp. 419–436, 2009.

[249] `https://www.kaggle.com/datasets`. Accessed: 2017-10-21.

[250] M. A. Shoorehdeli, M. Teshnehlab, and A. K. Sedigh, "Training anfis as an identifier with intelligent hybrid stable learning algorithm based on particle swarm optimization and extended kalman filter," *Fuzzy Sets and Systems*, vol. 160, no. 7, pp. 922–948, 2009.

[251] D. Simon, "Training radial basis neural networks with the extended kalman filter," *Neurocomputing*, vol. 48, no. 1, pp. 455–475, 2002.

[252] G. V. Puskorius and L. A. Feldkamp, "Decoupled extended kalman filter training of feedforward layered networks," in *Neural Networks, 1991., IJCNN-91-Seattle International Joint Conference on*, vol. 1, pp. 771–777, IEEE, 1991.

[253] Y. Ma, P. B. Luh, K. Kasiviswanathan, and E. Ni, "A neural network-based method for forecasting zonal locational marginal prices," in *Power Engineering Society General Meeting, 2004. IEEE*, pp. 296–302, IEEE, 2004.

[254] P. Areekul, T. Senjyu, N. Urasaki, and A. Yona, "Neural-wavelet approach for short term price forecasting in deregulated power market," *Journal of International Council on Electrical Engineering*, vol. 1, no. 3, pp. 331–338, 2011.

[255] `https://www.aemo.com.au/Electricity/National-Electricity-Market-NEM/`. Accessed: 2017-02-03.

[256] R. Sepulveda, O. Castillo, P. Melin, O. Montiel, and L. T. Aguilar, "Evolutionary optimization of interval type-2 membership functions using the human evolutionary model," in *Fuzzy Systems Conference, 2007. FUZZ-IEEE 2007. IEEE International*, pp. 1–6, IEEE, 2007.

[257] Y. Chen, B. Yang, and J. Dong, "Time-series prediction using a local linear wavelet neural network," *Neurocomputing*, vol. 69, no. 4, pp. 449–465, 2006.

[258] I. Aizenberg and C. Moraga, "Multilayer feedforward neural network based on multi-valued neurons (MLMVN) and a backpropagation learning algorithm," *Soft Computing*, vol. 11, no. 2, pp. 169–183, 2007.

[259] M. Almaraashi and R. John, "Tuning of type-2 fuzzy systems by simulated annealing to predict time series," in *Proceedings of the World Congress on Engineering*, vol. 2, pp. 976–980, 2011.

[260] V. Uslan, H. Seker, and R. John, "A support vector-based interval type-2 fuzzy system," in *IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*, pp. 2396–2401, 2014.

[261] S. Paul and S. Kumar, "Subsethood-product fuzzy neural inference system (SuPFuNIS)," *IEEE Transactions on Neural Networks*, vol. 13, no. 3, pp. 578–599, 2002.

[262] D. Nauck and R. Kruse, "A neuro-fuzzy approach to obtain interpretable fuzzy systems for function approximation," in *Fuzzy Systems Proceedings, 1998. IEEE World Congress on Computational Intelligence., The 1998 IEEE International Conference on*, vol. 2, pp. 1106–1111, IEEE, 1998.

[263] S. W. Tung, C. Quek, and C. Guan, "T2-hyfis-yager: Type 2 hybrid neural fuzzy inference system realizing yager inference," in *Fuzzy Systems, 2009. FUZZ-IEEE 2009. IEEE International Conference on*, pp. 80–85, IEEE, 2009.

[264] S. Wu and M. J. Er, "Dynamic fuzzy neural networks-a novel approach to function approximation," *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on*, vol. 30, no. 2, pp. 358–364, 2000.

[265] M. Russo, "Genetic fuzzy learning," *IEEE transactions on evolutionary computation*, vol. 4, no. 3, pp. 259–273, 2000.

[266] S. W. Tung, C. Quek, and C. Guan, "eT2FIS: an evolving type-2 neural fuzzy inference system," *Information Sciences*, vol. 220, pp. 124–148, 2013.

[267] `http://openmv.net/info/gas-furnace`. Accessed: 2016-07-21.

[268] G. E. Box and G. M. Jenkins, *Time series analysis: forecasting and control, revised ed.* Holden-Day, 1976.

[269] R. Tong, "The evaluation of fuzzy models derived from experimental data," *Fuzzy sets and systems*, vol. 4, no. 1, pp. 1–12, 1980.

[270] W. Pedrycz, "An identification algorithm in fuzzy relational systems," *Fuzzy sets and systems*, vol. 13, no. 2, pp. 153–167, 1984.

[271] C.-W. Xu and Y.-Z. Lu, "Fuzzy model identification and self-learning for dynamic systems," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 17, no. 4, pp. 683–689, 1987.

[272] M. Sugeno and T. Yasukawa, "Linguistic modeling based on numerical data," in *IFSA*, vol. 91, pp. 264–267, 1991.

[273] H. Surmann, A. Kanstein, and K. Goser, "Self-organizing and genetic algorithms for an automatic design of fuzzy control and decision systems," in *In Proc. EUFIT'93*, Citeseer, 1993.

[274] Y.-C. Lee, E. Hwang, and Y.-P. Shih, "A combined approach to fuzzy model identification," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 24, no. 5, pp. 736–744, 1994.

[275] Y. Lin and G. A. Cunningham, "A new approach to fuzzy-neural system modeling," *IEEE Transactions on Fuzzy systems*, vol. 3, no. 2, pp. 190–198, 1995.

[276] J. Nie, "Constructing fuzzy model by self-organizing counterpropagation network," *IEEE transactions on systems, man, and cybernetics*, vol. 25, no. 6, pp. 963–970, 1995.

[277] J.-S. R. Jang, C.-T. Sun, and E. Mizutani, "Neuro-fuzzy and soft computing, a computational approach to learning and machine intelligence," 1997.

[278] Y. Chen, B. Yang, and J. Dong, "Nonlinear system modelling via optimal design of neural trees," *International Journal of Neural Systems*, vol. 14, no. 02, pp. 125–137, 2004.

[279] P. Angelov and D. Filev, "Simpl_ets: A simplified method for learning evolving takagi-sugeno fuzzy models," in *Fuzzy Systems, 2005. FUZZ'05. The 14th IEEE International Conference on*, pp. 1068–1073, IEEE, 2005.

[280] G. M. MéNdez and M. De Los Angeles HernáNdez, "Hybrid learning mechanism for interval a2-c1 type-2 non-singleton type-2 takagi–sugeno–kang fuzzy logic systems," *Information Sciences*, vol. 220, pp. 149–169, 2013.

[281] M. Biglarbegian, W. W. Melek, and J. M. Mendel, "On the stability of interval type-2 tsk fuzzy logic control systems," *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 40, no. 3, pp. 798–818, 2010.

[282] H.-K. Lam and L. D. Seneviratne, "Stability analysis of interval type-2 fuzzy-model-based control systems," *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 38, no. 3, pp. 617–628, 2008.

[283] S. Coupland and R. John, "Geometric type-1 and type-2 fuzzy logic systems," *IEEE Transactions on Fuzzy Systems*, vol. 15, no. 1, pp. 3–15, 2007.

[284] S. Greenfield and R. John, "Optimised generalised type-2 join and meet operations," in *Fuzzy Systems Conference, 2007. FUZZ-IEEE 2007. IEEE International*, pp. 1–6, IEEE, 2007.

[285] L. A. Lucas, T. M. Centeno, and M. R. Delgado, "General type-2 fuzzy inference systems: analysis, design and computational aspects," in *Fuzzy Systems Conference, 2007. FUZZ-IEEE 2007. IEEE International*, pp. 1–6, IEEE, 2007.

[286] E. H. Mamdani, "Application of fuzzy algorithms for control of simple dynamic plant," in *Proceedings of the Institution of Electrical Engineers*, vol. 121, pp. 1585–1588, IET, 1974.

[287] J. C. Schlimmer and R. H. Granger, "Beyond incremental processing: Tracking concept drift.," in *AAAI*, pp. 502–507, 1986.

[288] J. M. Tomczak and A. Gonczarek, "Decision rules extraction from data stream in the presence of changing context for diabetes treatment," *Knowledge and Information Systems*, vol. 34, no. 3, pp. 521–546, 2013.

[289] E. Lughofer and P. Angelov, "Handling drifts and shifts in on-line data streams with evolving fuzzy systems," *Applied Soft Computing*, vol. 11, no. 2, pp. 2057–2068, 2011.

[290] P. P. Angelov and X. Zhou, "Evolving fuzzy-rule-based classifiers from data streams," *Fuzzy Systems, IEEE Transactions on*, vol. 16, no. 6, pp. 1462–1475, 2008.

[291] M. Pratama, S. G. Anavatti, P. P. Angelov, and E. Lughofer, "Panfis: A novel incremental learning machine," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 25, no. 1, pp. 55–68, 2014.

[292] M. Pratama, J. Lu, E. Lughofer, G. Zhang, and M. J. Er, "An incremental learning of concept drifts using evolving type-2 recurrent fuzzy neural networks," *IEEE Transactions on Fuzzy Systems*, vol. 25, no. 5, pp. 1175–1192, 2017.