Hybrid Learning for Interval Type-2 Intuitionistic Fuzzy Logic Systems as applied to Identification and Prediction Problems

Imo Eyoh, *Student Member, IEEE,* Robert John, *Senior Member, IEEE,* Geert De Maere and Erdal Kayacan, *Senior Member, IEEE*

***Abstract*—This paper presents a novel application of a hybrid learning approach to the optimisation of membership and non-membership functions of a newly developed interval type-2 intuitionistic fuzzy logic system (IT2 IFLS) of a Takagi-Sugeno-Kang (TSK) fuzzy inference system with neural network learning capability. The hybrid algorithms consisting of decoupled extended Kalman filter (DEKF) and gradient descent (GD) are used to tune the parameters of the IT2 IFLS for the first time. 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 IT2 IFLS, a TSK-type-1 intuitionistic fuzzy logic system (IFLS-TSK) and a TSK-type interval type-2 fuzzy logic system (IT2 FLS-TSK) on two instances of the datasets under investigation. The empirical comparison is made on the designed systems using three artificially generated datasets and three real world datasets. Analysis of results reveal that IT2 IFLS outperforms its type-1 variants, IT2 FLS and most of the existing models in the literature. Moreover, the minimal run time of the proposed hybrid learning model for IT2 IFLS also puts this model forward as a good candidate for application in real time systems.**

***Index Terms*—Interval type-2 intuitionistic fuzzy logic system; Decoupled extended Kalman filter; Gradient descent algorithm.**

I. Introduction

FUZZY sets which generalise the conventional sets were introduced by Zadeh in 1965 where elements of a set have membership degrees and defined in the unit interval [0,1] using membership functions (MF). Fuzzy sets can be used to model nonlinear relationships with ill-defined information and as such are able to capture human experiences or knowledge. However, because of the precise nature of the MF of FSs, they may not handle the level of uncertainty inherent in many real world applications. Zadeh [1] in 1975 introduced the concept of type-2 fuzzy set (T2 FS). The T2 FSs are higher order conventional fuzzy sets with fuzzy MFs, in contrast to the type-1 counterpart with precise MFs. With this kind of MF for T2 FSs, researchers are able to *" $\cdots$ model and minimise the effects of uncertainties in rule-based fuzzy logic systems"* [2]. Due to the complex nature of the general T2 FS with non-uniformly weighted secondary MFs, many researchers have focussed on the less complex T2 FS model known as the interval T2 FS (IT2 FS) where all the secondary MFs (membership degrees on the third dimension) take the value 1. The uncertainty about an IT2 FS is therefore captured in their footprints of uncertainties (FOUs).

Because the conventional fuzzy set non-membership function (NMF) (ν) is complementary to the MF (μ), that is, $\nu = 1 - \mu$ with no form of uncertainty whatsoever, Atanassov [3] in 1986 introduced a new kind of fuzzy set known as the intuitionistic fuzzy set (IFS), which is characterised by independently defined MF and NMF together with some degree of indeterminacy or hesitancy (otherwise known as intuitionistic fuzzy index (IF-index)). The IFS therefore defies the claim of the FS that $\mu + \nu = 1$ *" $\cdots$ thus relaxing the enforced duality that $\nu = 1 - \mu$ "* [4] and maintains a set whose sum of MF and NMF is less than or equal to 1 with the capacity to capture more uncertainty in terms of the hesitation degree. With the degree of MF, NMF and hesitation, the IFS becomes more meaningful in the context of human reasoning and natural language representation. However, similar to the notion of a conventional FS of type-1, IFS with precise MF and NMF may not handle uncertainty well. Atanassov and Gargov [5] proposed the extension of IFS to IV IFS based on the idea of interval-valued fuzzy set (IV FS) [6] defined in [0,1] with MFs that are intervals. It has been pointed out in [7]–[10] that IV FSs are special representations of IT2 FSs. In particular, Bustince *et al.* [10] has shown four different representations for IT2 FSs namely as type-1 FS, as IV FS, as multi-FS and as multi-interval FS. Thus, according to the authors, IT2 FSs can conveniently be used to construct other concepts that are not possible with IV FSs, and as such IV FS and IT2 FS should be treated as two distinct sets. Arguably, this also applies to IT2 IFS. We argue that similar to IT2 FSs and its representations, IT2 IFS can be used to model other concepts, a capability that is not possible with IV IFSs. It is necessary to make this distinction in the context of this study as it serves to distinguish the much broader concept of IT2 IFS from the more specific notion of IV IFS. Moreover, the constraint imposed on IV IFS is quite different from those of IT2 IFS. For instance, for IV IFS, the summation of the

Imo Eyoh and Robert John are with the Automated Scheduling, Optimisation and Planning (ASAP) and Laboratory for Uncertainty in Data and Decision Making (LUCID) research groups, School of Computer Science, University of Nottingham, UK (e-mail: {ije, rij}@cs.nott.ac.uk).

Geert De Maere is with the ASAP research group, School of Computer Science, University of Nottingham, UK (e-mail: gdm@cs.nott.ac.uk).

Erdal Kayacan is with the School of Mechanical and Aerospace Engineering, Nanyang Technological University, Singapore (e-mail: erdal@ntu.edu.sg).

upper bound MF and upper bound NMF is less than or equal to 1 [5]. On the contrary, for IT2 IFS, the summation of the upper bound MF and lower bound NMF is less than or equal to 1 and the summation of the lower bound MF and upper bound NMF is less than or equal to 1 for every element of the fuzzy set. These notions present IT2 IFS as a novel class of fuzzy set entirely different from IV IFS. Moreover, apart from the novelty in the constraint definitions of our proposed framework, the model also takes into account the contributions of the NMF and IF-indices in the partitioning of the input space. The capturing of additional uncertainties in the form of IF-indices in the FOUs of the MF and NMF through a process of scaling and shifting presents IT2 IFLS as a novel approach completely different from existing approaches. In this way, we relax the assumption of IT2 FS that NMF is complementary to MF (upper or lower).

The use of T2 and IFSs is not a new approach in the FL community and has been successfully applied in diverse domains. In fact, Castillo *et al.* [11], used IFS to represent IT2 FS and this is achieved with a suitable choice of interval function (g). The same authors in [12] also investigate the use of IFS and its multidimensional (MDIFS) variant to interpret FSs, IT2 FSs and generalised T2 FSs (GT2 FSs). Some research has shown interest in the arithmetic operations of T2 IFS. In Cuong *et al.* [13] some operations for T2 IFS and their properties are discussed with the remark that many applications will benefit from the use of such sets. Similarly, Jana [14] has proposed some novel arithmetic operations on generalised T2 IFS on the basis of (α, β)-cut methods with application to transportation problems. Recently, Singh and Garg [15] have proposed some distance measures for T2 IFS and applied the proposed measures to multi-criteria decision making.

In the work of Nguyen *et al.* [16], IT2 fuzzy C-mean (IT2FCM) and IFS are applied for image clustering with improved performance in the clustering quality compared to FCM and IT2FCM. Nghiem *et al.* [17] applied intuitionistic T2 FS to image thresholding using Sugeno intuitionistic fuzzy generator. Results show improved thresholding quality compared to standard algorithms such as type-1 FS and IFS. In Naim and Hagras [18], a hybrid approach using IT2 and IFS for multi-criteria group decision making (MCGDM) is proposed. In their study, IT2 FS is used to handle the linguistic uncertainty with intuitionistic evaluation used in the design of the NMF degrees. The proposed hybrid approach was evaluated on 10 candidates in a postgraduate study. Results of evaluation show better agreement with the human experts decision than IFS, FS and IT2 fuzzy systems. In the same vein, Naim *et al.* [19] presented a fuzzy logic-MCGDM (FL-MCGDM) for choosing appropriate and convenient lighting level to meet particular individual reading needs. The hybrid system adopted the concepts of IT2 FS and the hesitation indices of IFS with intuitionistic values used to represent the MFs of the IT2 FS for the left and right end-points. Results show that with T2 FS and IFS, the capability of FL-MDGDM is enhanced with improved capacity for decision making. In Naim and Hagras [20], FL-MCGDM is proposed for intelligent shared environment. Analysis of results indicate consistency of the proposed method with the human decision as compared to conventional fuzzy MCGDM. In a study by Own [7], a switching between T2 FSs and IFSs is proposed. In Own's paper, the switching relation between T2 FSs and IFSs is defined axiomatically. The advantage of T2 FSs is 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 of using both IFS and IT2 FS in the design of the proposed framework of IT2 IFLS. However, the motive and approach for the framework proposed in this study are quite distinct from those advanced in the above models. Among other things, no learning whatsoever is involved in these previous models and no framework listed above obviously shows the benefit of explicitly using MFs and NMFs that are intervals together with IF-indices for uncertainty modeling and although, these different approaches use combination of IFS and IT2 FS, most of these approaches are less relevant to the problem domain investigated in this study. In addition, some of these approaches utilise only a single IT2 FS and evaluate the hesitation on the primary MF of the IT2 FS. For instance, the works of [18]–[20] while effective in handling MCGDM, do not consider the specification of NMF as a separate region but rather IT2 FS is employed with intuitionistic evaluation (hesitation) on the MF FOUs. They do not explicitly apply MFs and NMFs that are intervals. On the contrary, the new framework incorporates the hesitation indices of each element into the FOUs of the MFs and NMFs through scaling and shifting as shown in Equations (10) to (13) thereby satisfying the constraints imposed on Definition 4. Hence, our proposed model of IT2 IFLS provides a point of departure from existing approaches in the literature by using a framework that fuses both concepts (IFS and IT2 FS) systematically and models uncertainty using separately defined MFs and NMFs that are intervals with a neural network learning capability.

Type-1 FSs have been the main stay in uncertainty modeling in the last few decades and have been applied in many applications with great success. Despite the widespread use of FS and its representation of uncertainty, FS handles uncertainty about the meaning of words by using MFs that are precise [21]; which is not necessarily realistic [22]. Real world applications are fraught with higher order uncertainties that make it difficult to determine the exact MFs for the antecedent and consequent parts of a fuzzy set [23]. With these level of uncertainties, it becomes inappropriate in certain applications to employ type-1 FS. According to Hagras [24], 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 modeling cannot be properly accomplished with type-1 FS, the use of T2 FS which is characterised by MFs that are themselves fuzzy can be more appropriate. Mendel [21] states that T2 FLSs control the effects of the uncertainties associated with the meaning of words by *modeling the uncertainties* and concluded in [25] that IT2 FLS is a scientifically correct model for modeling uncertainties associated with words. What

is more, some studies have shown in a more analytical way that the performance of T2 FLSs is superior to their type-1 counterparts [26], [27] and T2 FLSs can be more appropriate for real time environments [28].

The IT2 FSs have been extensively used in the literature to model uncertainty (see [23], [29]–[32]). Despite literature being replete with several works revolving round IT2 FSs, they only make use of the MFs alone in uncertainty modeling with an implicit assertion that NMF is complementary to the MF (lower or upper). In a real life context, it is not necessarily the case that NMF is complementary to MF as there may exist some degree of hesitation or indeterminacy, otherwise known as intuitionistic fuzzy index (IF-index) or neutral degree. The conventional IT2 FLS cannot singularly model these IF-indices in a fuzzy set. Barrenechea *et al.* [33] pointed out that valuable information of an element can be obtained using the IF-index of IFS. The authors in [33] also noted that the IF-index plays a very important role in algorithm's performance. Our study is an attempt in this direction to enhance the capabilities of IT2 FLS by incorporating IFL into IT2 FLS. As earlier discussed, with the capability of the IT2 FSs to adequately model uncertainty in their FOUs and the ability of the IFS to separately cater for MF and NMF of an element with some level of hesitation, we are motivated to integrate these two concepts (IT2 FS and IFS) to design a new TSK-type interval T2 intuitionistic fuzzy logic system (IT2 IFLS-TSK) [34]. The new framework apart from incorporating fuzzy NMF into the conventional IT2 FS is able to deal with indeterminate (hesitant) states which are not well managed by alternative fuzzy approaches such as IT2 FLSs. The introduction of additional NMF and IF-indices into IT2 FS increases the fuzziness of the model. According to Hisdal [35], "increased fuzziness in a description means increased ability to handle inexact information in a logically correct manner." We believe that the fusion of these two kinds of fuzzy sets is able to provide a synergistic capability in managing the effects of uncertainties in data. The proposed framework of IT2 IFLS is enhanced with a neural network learning capability similar to adaptive neuro-fuzzy inference system (ANFIS) and T2-ANFIS [36] for modeling uncertainty in data. The combination of these two approaches, fuzzy logic and artificial neural network (ANN), have been very popular with increasing interest in recent years. With the integration of ANN into FLS, the FLS is enhanced with the learning and generalisation capabilities of ANN.

Different approaches have been used over the years to learn or adjust the parameters of fuzzy systems and are classified into derivative-based and non-derivative-based optimization methods. Among the derivative-based approaches are gradient descent (GD) methods, least squares, extended Kalman filter (EKF) while the non-derivative-based methods include genetic algorithms (GA), particle swarm optimization (PSO), simulated annealing (SA) and sliding mode control (SMC) - based theory algorithms. Kayacan *et al.* [37] have investigated some of these learning algorithms for adjustment of the parameters of IT2 fuzzy neural network (IT2 FNN) including hybrid method of PSO and GD. Other hybrid methods for tuning the parameters of the conventional TSK-type IT2 FLSs are reported in [38]–[42]. The GD (first-order derivative based) methods have been widely used as an optimisation strategy for the parameters of fuzzy systems [21], [37]. The difficulties associated with GD methods however, are slow convergence and the possibility of getting stuck in local minima, leading to poor solutions [43]. This can be compensated for by combining the first-order GD with a higher-order derivative-based method such as the Kalman filter (KF)-based algorithms which have a smaller possibility of getting stuck in local minima [44]. In a different application domain, the hybrid learning utilising KF-based and GD techniques has shown good performance. For instance, Mendez *et al.* [38] proposed a hybrid learning approach for IT2 FLS of TSK-type otherwise known as interval type-1 non-singleton type-2 TSK FLS ANFIS (IT2 NSFLS1 ANFIS) utilising recursive Kalman-type filter (REFIL) to tune the consequent parameters and the steepest descent back propagation method to tune the antecedent parameters. The developed model was applied to the prediction of transfer bar surface temperature. Experimental evaluation revealed that the IT2 NSFLS1 ANFIS trained with hybrid REFIL-BP had the lowest prediction error on test data compared to other learning approaches investigated in their study. However, the basic KF works well for linear dynamic systems with white process and measurement noise but real world problems are non-linear. Hence, for nonlinear systems, we have extended the linear KF used in [38] 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 fuzzy logic systems [44], [45] and intuitionistic fuzzy systems of type-1 [46], [47]. However, because of the high dimensionality of the fuzzy system parameters, using the standard EKF can be more complicated [44], [48] 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 [48] with the most useful properties of the EKF still preserved [49]. The DEKF algorithm has been used previously in [44] to train a T2 FLS where the parameters of both the antecedent and consequent parts of the T2 FLS were gathered into two separate vectors (antecedent and consequent parameter vectors). Similar to [38], we adopt a hybrid learning methodology (KF-based and GD) to adjust the antecedent and consequent parameters of the proposed model. While GD is also adopted for the update of the antecedent parameters, our model utilises the DEKF approach, different from [38], to adjust the consequent parameters of the new and extended framework of IT2 FLS, otherwise known as IT2 IFLS [34] for the first time in this study with the aim of achieving improved system performance in terms of error minimisation and faster convergence. To the best knowledge of the authors, there is previously no work in the literature where DEKF and GD is used for the optimisation of IT2 IFLS-TSK parameters.

In Section II, we provide definitions for IFS, GT2 IFS and IT2 IFS. Section III discusses the design of IT2 IFLS-TSK model and the hybrid learning algorithms are presented in Section IV. We present our experimental results in Section V, and conclude in Section VI.

II. Definitions

In this section the basic definitions surrounding the formulation of the proposed framework are given.

Intuitionistic Fuzzy Set (IFS)

Definition 1. *An intuitionistic fuzzy set A^* in X is of the form: $A^* = \{(x, \mu_{A^*}(x), \nu_{A^*}(x)) : x \in X)\}$, where $\mu_{A^*}(x) : X \to [0,1]$ is the MF degree and $\nu_{A^*}(x) : X \to [0,1]$ is the NMF degree of element $x \in X$ restricted by $0 \leq \mu_{A^*}(x) + \nu_{A^*}(x) \leq 1$* [3].

A conventional fuzzy set A is obtained if $\nu_{A^*}(x) = 1 - \mu_{A^*}(x)$ for every $x \in X$.

Intuitionistic Fuzzy Index - IF-index

Definition 2. *Given an IFS, A^*, an IF-index, denoted by π and defined in the unit interval [0,1] is the complement of the sum of the degree of MF and NMF of an element x to 1 i.e. $\pi_{A^*}(x) = 1 - (\mu_{A^*}(x) + \nu_{A^*}(x))$. Obviously $0 \leq \pi_{A^*}(x) \leq 1$* [3].

Generalised Type-2 Intuitionistic Fuzzy Set (GT2 IFS)

Definition 3. *A GT2 IFS denoted by $\tilde{A}^*$ is composed of a T2 MF $\mu_{\tilde{A}^*}(x,u)$, and a T2 NMF $\nu_{\tilde{A}^*}(x,u)$* [34]*, i.e.,*

$$\tilde{A}^* = \{(x,u), \mu_{\tilde{A}^*}(x,u), \nu_{\tilde{A}^*}(x,u) \mid \forall x \in X, \forall u \in J_x^\mu, \forall u \in J_x^\nu\} \quad (1)$$

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

$$J_x^\mu = \left\{(x,u) : u \in \left[\underline{\mu}_{\tilde{A}^*}(x), \overline{\mu}_{\tilde{A}^*}(x)\right]\right\} \quad (2)$$

$$J_x^\nu = \{(x,u) : u \in [\underline{\nu}_{\tilde{A}^*}(x), \overline{\nu}_{\tilde{A}^*}(x)]\} \quad (3)$$

Equations (2) and (3) for GT2 IFS represent the supports of secondary MF and secondary NMF values of element $x \in X$ on the third dimension respectively. When uncertainties disappear, a T2 IFS must collapse to an IFS. A T2 IFS can also be expressed as,

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

where $\int\int\int$ represents union over all admissible values of x and u over a continuous universe of discourse (UoD) and $\sum$ instead for discrete UoD. When $\mu_{\tilde{A}^*}(x,u) = 1$ and $\nu_{\tilde{A}^*}(x,u) = 1$, a T2 IFS simplifies to an IT2 IFS (see Figure 1).

Interval Type-2 Intuitionistic Fuzzy Set (IT2 IFS)

Definition 4. *An IT2 IFS, $\tilde{A}^*$, consist of a T2 MF and a T2 NMF defined as $[\underline{\mu}_{\tilde{A}^*}(x), \bar{\mu}_{\tilde{A}^*}(x)]$ and $[\underline{\nu}_{\tilde{A}^*}(x), \bar{\nu}_{\tilde{A}^*}(x)]$ respectively for all $x \in X$ with constraints: $0 \leq \overline{\mu}_{\tilde{A}^*}(x) + \underline{\nu}_{\tilde{A}^*}(x) \leq 1 \, and \ 0 \leq \underline{\mu}_{\tilde{A}^*}(x) + \overline{\nu}_{\tilde{A}^*}(x) \leq 1$* [16].

An IT2 IFS can be expressed as:

$$\begin{aligned}\tilde{A}^* &= \int_{x \epsilon X}\int_{u \epsilon J_x^\mu}\int_{u \epsilon J_x^\nu} 1/(x,u) \\ &= \int_{x \epsilon X}\left[\int_{u \epsilon J_x^\mu}\int_{u \epsilon J_x^\nu} 1/(u)\right]/x\end{aligned} \quad (4)$$

where x is the primary variable, and u is the secondary variable. For IT2 IFS, J_x^μ and J_x^ν represent the primary MF and primary NMF values of element $x \in X$. The footprints of uncertainty (FOU) of an IT2 IFS consist of MF-FOU - Equation (5) and NMF-FOU - Equation (6).

$$FOU_\mu\left(\tilde{A}^*\right) = \bigcup_{\forall x \in X}\left[\underline{\mu}_{\tilde{A}^*}(x), \bar{\mu}_{\tilde{A}^*}(x)\right] \quad (5)$$

$$FOU_\nu\left(\tilde{A}^*\right) = \bigcup_{\forall x \in X}\left[\underline{\nu}_{\tilde{A}^*}(x), \bar{\nu}_{\tilde{A}^*}(x)\right] \quad (6)$$

The FOUs together with the IF-indices completely describe the uncertainty about an IT2 IFS.
The IF-indices utilised for the definition of IT2 IFS are as follows [34]:

$$\pi_c(x) = max\left(0, (1 - (\mu_{\tilde{A}^*}(x) + \nu_{\tilde{A}^*}(x)))\right) \quad (7)$$

$$\overline{\pi}_{var}(x) = max\left(0, (1 - (\overline{\mu}_{\tilde{A}^*}(x) + \underline{\nu}_{\tilde{A}^*}(x)))\right) \quad (8)$$

$$\underline{\pi}_{var}(x) = max\left(0, \left(1 - \left(\underline{\mu}_{\tilde{A}^*}(x) + \overline{\nu}_{\tilde{A}^*}(x)\right)\right)\right) \quad (9)$$

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

III. Interval Type-2 Intuitionistic Fuzzy Logic System

The architecture of a T2 IFLS is the same as that of a conventional T2 FLS and composed of the fuzzifier, rule base, inference engine and output processing unit.

A. Fuzzification

During this process, the input vector is projected onto the IT2 IFS and the degrees of MF, $\mu_{\tilde{A}^*}(x)$ and NMF, $\nu_{\tilde{A}^*}(x)$ are obtained for each input and in this study, interval singleton fuzzification is utilised where the MF and NMF are defined at a single point where $x = x'$ with value 1 respectively and 0 everywhere that $x \neq x'$. We use a Gaussian function with uncertain standard deviation (with some scaling and shifting) to construct the IT2 IFS. The purpose of scaling and shifting in the MF and NMF is to capture the hesitation of the human experts in the specification of the center and the spread of the IT2 IFS. The functions are as defined in Equations (10) to (13) [34].

$$\overline{\mu_{ik}}(x_i) = exp\left(-\frac{(x_i - c_{ik})^2}{2\overline{\sigma}_{2,ik}^2}\right) * (1 - \pi_{c,ik}(x_i)) \quad (10)$$

$$\underline{\mu_{ik}}(x_i) = exp\left(-\frac{(x_i - c_{ik})^2}{2\underline{\sigma}_{1,ik}^2}\right) * (1 - \pi_{c,ik}(x_i)) \quad (11)$$

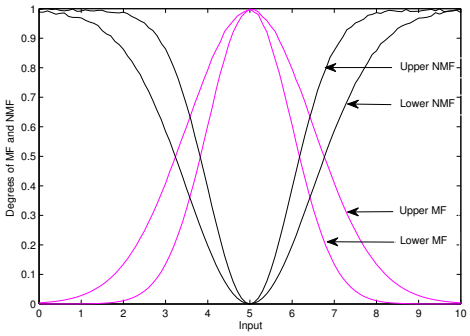

Fig. 1: An IT2 IFS using Gaussian function [34]

$$\overline{\nu_{ik}}\left(x_i\right) = \left(1 - \overline{\pi}_{var,ik}(x_i)\right) - \left[exp\left(-\frac{\left(x_i - c_{ik}\right)^2}{2\bar{\sigma}^2_{1,ik}}\right) * \left(1 - \pi_{c,ik}(x_i)\right)\right] \quad (12)$$

$$\underline{\nu_{ik}}\left(x_i\right) = \left(1 - \underline{\pi}_{var,ik}(x_i)\right) - \left[exp\left(-\frac{\left(x_i - c_{ik}\right)^2}{2\underline{\sigma}^2_{2,ik}}\right) * \left(1 - \pi_{c,ik}(x_i)\right)\right] \quad (13)$$

where $\pi_{c,ik}$ is the IF-index of center and $\pi_{var,ik}$ is the IF-index of variance [47]. The IF-THEN rule of an IT2 IFLS does not differ from the general TSK rule syntax of fuzzy logic. The generic TSK rule representation is as expressed in Equation (14):

$$R_k : \text{IF } x_1 \text{ is } \tilde{A^*}_{1k} \text{ and } x_2 \text{ is } \tilde{A^*}_{2k} \text{ and } \cdots \text{and } x_n \text{ is } \tilde{A^*}_{nk} \quad \text{THEN } y_k = \sum_{i=1}^{n} w_{ik}x_i + b_k \quad (14)$$

where $\tilde{A^*}_{1k}$, $\tilde{A^*}_{2k}$, $\cdots$, $\tilde{A^*}_{ik}$, $\cdots$, $\tilde{A^*}_{nk}$ are IT2 IFS and y_k is the output of the kth rule.
For MF, the rule in Eqn (14) is decomposed to:

$$R^{\mu}_k : IF\ x_1\ is\ \tilde{A^*}^{\mu}_{1k}\ and\ x_2\ is\ \tilde{A^*}^{\mu}_{2k}\ and\ \cdots and\ x_n\ is\ \tilde{A^*}^{\mu}_{nk} \quad THEN\ y^{\mu}_k = \sum_{i=1}^{n} w^{\mu}_{ik}x_i + b^{\mu}_k \quad (15)$$

For NMF, the rule becomes:

$$R^{\nu}_k : IF\ x_1\ is\ \tilde{A^*}^{\nu}_{1k}\ and\ x_2\ is\ \tilde{A^*}^{\nu}_{2k}\ and\ \cdots and\ x_n\ is\ \tilde{A^*}^{\nu}_{nk} \quad THEN\ y^{\nu}_k = \sum_{i=1}^{n} w^{\nu}_{ik}x_i + b^{\nu}_k \quad (16)$$

where y^{μ}_k is the MF output and y^{ν}_k is the NMF output of the kth rule, w and b are the consequent parameters representing the weights and bias respectively.

B. Fuzzy Inference

In this paper, we aim to obtain a fuzzy system that closely approximates the input-output relationship of the modeled system, hence the use of a TSK-fuzzy inference for IT2 IFLS. The proposed hybrid learning IT2 IFLS is developed using a TSK-based fuzzy inference with an IT2 IFS in the antecedent and a crisp value in the consequent (A2-C0) because of its simplicity in representation and evaluation. From our previous study, the final output of IT2 IFLS-TSK is a weighted average defined as follows [34]:

$$y = \frac{(1-\beta)\sum_{k=1}^{M}\left(\underline{f^{\mu}_k} + \overline{f^{\mu}_k}\right)y^{\mu}_k}{\sum_{k=1}^{M}\underline{f^{\mu}_k} + \sum_{k=1}^{M}\overline{f^{\mu}_k}} + \frac{\beta\sum_{k=1}^{M}\left(\underline{f^{\nu}_k} + \overline{f^{\nu}_k}\right)y^{\nu}_k}{\sum_{k=1}^{M}\underline{f^{\nu}_k} + \sum_{k=1}^{M}\overline{f^{\nu}_k}} \quad (17)$$

where $\underline{f^{\mu}_k}$, $\overline{f^{\mu}_k}$, $\underline{f^{\nu}_k}$ and $\overline{f^{\nu}_k}$ are the lower MF, upper MF, lower NMF and upper NMF firing strength respectively. The "prod" t-norm for the implication is expressed as:

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

where $*$ is the "prod" operator, y^{μ}_k is the MF output and y^{ν}_k is the NMF output of the kth rule. The contribution of MF and NMF in the final output depending on the value of β can be expressed as in Equation (18) as follows:

$$y = \begin{cases} MF\ only & \text{if } \beta = 0 \\ MF\ and\ NMF & \text{if } 0 < \beta < 1 \\ NMF\ only, & \text{if } \beta = 1 \end{cases} \quad (18)$$

Hence, the parameter β in the unit interval [0,1] determines the magnitude of MF and NMF in the overall output.

IV. Parameter Updates

In this section, we describe the two-pass learning algorithm of the parameters of IT2 IFLS. 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 as shown in Table I.

TABLE I: Two-pass hybrid learning of IT2 IFLS

	Forward pass	Backward pass
Antecedent parameter	Fixed	GD
Consequent parameter	DEKF	Fixed

A. Consequent Parameter Updates

For linear systems with white process and measurement noise, KF has been shown to be an optimal linear estimator, but real world applications are non-linear, hence the extended version of KF (EKF) is used instead and a simpler version of the EKF, DEKF, is utilised to train the consequent parts of the IT2 IFLS model because it is less complex. To update the consequent parameters, two steps are involved which are the time update and the measurement update. During the time

update, the current state is projected forward in time and a prior estimate for the next step is obtained. During the measurement update, a new measurement is propagated and the posteriori estimate is obtained. In using the DEKF, all the consequent parameters for both MFs and NMFs of IT2 IFLS are grouped into a single vectors.

1) The Extended Kalman Filter - EKF: Let $y = f(X, \theta)$ be the output of a fuzzy logic system, where $X's$ are the inputs into the system and θ are unknown state variables. The non-linear system model is represented as:

$$\theta_{t+1} = f(\theta_t) + \omega_t \tag{19}$$

$$y_t = h(\theta_t) + \upsilon_t \tag{20}$$

where θ_t is the state variable at time t, $f(.)$ is the non-linear vector function of state, $h(.)$ is the non-linear vector function of the output, y_t is the observation vector, ω_t and υ_t are the process and measurement noise respectively which are assumed to be Gaussian and uncorrelated with:

$$\begin{aligned} E(\theta_0) = \overline{\theta}_0, \qquad & E[(\theta_0 - \overline{\theta}_0)(\theta_0 - \overline{\theta}_0)^T] = P_0 \\ E(\omega_t) = 0, \qquad & E(\omega_t \omega_l^T) = Q\delta_{tl} \\ E(\upsilon_t) = 0, \qquad & E(\upsilon_t \upsilon_l^T) = R\delta_{tl} \end{aligned}$$

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

$$\begin{aligned} f(\theta_t) &= f(\hat{\theta}_t) + F_t(\theta_t - \hat{\theta}_t) + H.O.T \\ h(\theta_t) &= h(\hat{\theta}_t) + H_t(\theta_t - \hat{\theta}_t) + H.O.T \end{aligned} \tag{21}$$

where:

$$F_t = \left.\frac{\partial f(\theta)}{\partial \theta}\right|_{\theta=\hat{\theta_t}}$$

$$H_t^T = \left.\frac{\partial h(\theta)}{\partial \theta}\right|_{\theta=\hat{\theta_t}}$$

and H.O.T stands for the higher order term. Without the H.O.T, Equation (21) can be estimated as in Equation (22) below:

$$\begin{aligned} \theta_{t+1} &= F_t\theta_t + \omega_t + \phi_t \\ y_{t+1} &= H_t^T\theta_t + \upsilon_t + \varphi_t \end{aligned} \tag{22}$$

where ϕ_t and φ_t are random error term for state and observation equation respectively and expressed as:

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

The desired estimation of the parameters in Equation (22) can therefore be obtained using the recursive Kalman procedures in Equations (23) to (25) [45], [50], [51].

$$K_t = P_t H_t[(H_t)^T P_t H_t + R]^{-1} \tag{23}$$

$$\hat{\theta}_t = f(\hat{\theta}_{t-1}) + K_t[y_t - h(\hat{\theta}_{t-1})] \tag{24}$$

$$P_{t+1} = F_k(P_t - K_t P_t (H_t)^T)F_K^T + Q \tag{25}$$

The vector F_k is taken as an identity matrix and Equations (26) to (28) are obtained [44].

$$K_t = P_t H_t[(H_t)^T P_t H_t + R]^{-1} \tag{26}$$

$$\hat{\theta}_t = \hat{\theta}_{t-1} + K_t[y_t - h(\hat{\theta}_{t-1})] \tag{27}$$

$$P_{t+1} = P_t - K_t P_t (H_t)^T + Q \tag{28}$$

where K_t is the Kalman gain at time step t, P_t is the covariance matrix of the state estimation error at step t, R is the measurement noise covariance and Q is the covariance of process noise. The computational cost of EKF is in the order of DB^2 where D is the dimension of the output and B is the total number of parameters. Thus for an IT2 IFLS with n inputs, M number of rules and an output, the total number of parameters to be tuned is $6n + 2M(n + 1)$. The computational expense of EKF for IT2 IFLS is therefore $36n^2 + 4M^2(n^2 + 2n + 1) + 24nM(n + 1)$ which is a very large number in most applications.

2) Decoupled Extended Kalman Filter - DEKF: In order to reduce the computational cost of EKF, the DEKF is employed to adapt the consequent parameters of IT2 IFLS during the forward pass. The consequent parameters for both MF and NMF are grouped in a vector. Hence, the state of the non-linear system can be denoted as:

$$\begin{aligned} \theta = [&w_{11}^{\mu}, w_{21}^{\mu}, \cdots, w_{Mn}^{\mu}, b_1^{\mu}, b_2^{\mu}, \cdots, b_M^{\mu}, \\ &w_{11}^{\nu}, w_{21}^{\nu}, \cdots, w_{Mn}^{\nu}, b_1^{\nu}, b_2^{\nu}, \cdots, b_M^{\nu}]^T \end{aligned} \tag{29}$$

Equation (29) is further decomposed into:

$$\theta^{\mu} = [w_{11}^{\mu}, \cdots w_{n1}^{\mu}, \cdots, w_{M1}^{\mu}, \cdots, w_{Mn}^{\mu}, b_1^{\mu}, b_2^{\mu}, \cdots, b_M^{\mu}]^T \tag{30}$$

for MF parameters and

$$\theta^{\nu} = [w_{11}^{\nu}, \cdots, w_{n1}^{\nu}, \cdots, w_{M1}^{\nu}, \cdots, w_{Mn}^{\nu}, b_1^{\nu}, b_2^{\nu}, \cdots, b_M^{\nu}]^T \tag{31}$$

for NMF parameters, with the MF and NMF having separate Kalman parameters. The derivative matrix, H, is defined as:

$$H^{\mu} = \frac{\partial y}{\partial \theta^{\mu}}, \;\; H^{\nu} = \frac{\partial y}{\partial \theta^{\nu}} \tag{32}$$

which encompasses

$$\frac{\partial y}{\partial w_{ik}} = \frac{\partial y}{\partial y_k}\frac{\partial y_k}{\partial w_{ik}} = \left[\frac{\partial y}{\partial y_k^{\mu}}\frac{\partial y_k^{\mu}}{\partial w_{ik}^{\mu}} + \frac{\partial y}{\partial y_k^{\nu}}\frac{\partial y_k^{\nu}}{\partial w_{ik}^{\nu}}\right] \tag{33}$$

and

$$\frac{\partial y}{\partial b_k} = \frac{\partial y}{\partial y_k}\frac{\partial y_k}{\partial b_k} = \left[\frac{\partial y}{\partial y_k^{\mu}}\frac{\partial y_k^{\mu}}{\partial b_k^{\mu}} + \frac{\partial y}{\partial y_k^{\nu}}\frac{\partial y_k^{\nu}}{\partial b_k^{\nu}}\right] \tag{34}$$

The parameter update rules for the consequent parts of the MF then follow the Kalman filtering recursive procedures as shown in Equation (35) to (37):

$$K_t^{\mu} = P_t^{\mu} H_t^{\mu}[(H_t^{\mu})^T P_t^{\mu} H_t^{\mu} + R^{\mu}]^{-1} \tag{35}$$

$$\hat{\theta}_t^{\mu} = \hat{\theta}_{t-1}^{\mu} + K_t^{\mu}[y_t - h(\hat{\theta}_{t-1})] \tag{36}$$

$$P_{t+1}^{\mu} = P_t^{\mu} - K_t^{\mu} P_t^{\mu} (H_t^{\mu})^T + Q^{\mu} \tag{37}$$

and the updates for the NMFs follow the same recursive procedure but utilises NMF parameters as in Equation (38) to (40):

$$K_t^\nu = P_t^\nu H_t^\nu [(H_t^\nu)^T P_t^\nu H_t^\nu + R^\nu]^{-1} \tag{38}$$

$$\hat{\theta}_t^\nu = \hat{\theta}_{t-1}^\nu + K_t^\nu [y_t - h(\hat{\theta}_{t-1})] \tag{39}$$

$$P_{t+1}^\nu = P_t^\nu - K_t^\nu P_t^\nu (H_t^\nu)^T + Q^\nu \tag{40}$$

With the DEKF, the reduction in the computational cost is in the order $36n^2 + 4M^2(n^2 + 2n + 1)$ and the computational complexity of DEKF to EKF is in the ratio:

$$\frac{36n^2 + 4M^2(n^2 + 2n + 1)}{36n^2 + 4M^2(n^2 + 2n + 1) + 24nM(n + 1)}$$

The DEKF therefore has an advantage over the conventional EKF in terms of resource utilisation and coupled with the complexity of IT2 IFLS, DEKF becomes the preferred learning approach in this study.

B. Antecedent Parameter Updates

To adjust the antecedent parameters of the IT2 IFLS, GD algorithm is executed. For a single output, the cost function is expressed as:

$$E = \frac{1}{2}(y^a - y)^2$$

where y^a is the actual output and y is the IT2 IFLS output. The following GD update rules are used for the tuning of the antecedent parameters:

$$c_{ik}(t+1) = c_{ik}(t) - \gamma \frac{\partial E}{\partial c_{ik}} \tag{41}$$

$$\sigma_{1,ik}(t+1) = \sigma_{1,ik}(t) - \gamma \frac{\partial E}{\partial \sigma_{1,ik}} \tag{42}$$

$$\sigma_{2,ik}(t+1) = \sigma_{2,ik}(t) - \gamma \frac{\partial E}{\partial \sigma_{2,ik}} \tag{43}$$

where γ is the learning rate. The derivatives in Equations (41) - (43) are computed as follows:

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

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

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

The design parameter β is adjusted using Equation (44):

$$\beta(t+1) = \beta(t) - \gamma \frac{\partial E}{\partial \beta} \tag{44}$$

which allows for adaptive contributions of the MF and the NMF during the hybrid learning process.

Algorithm 1 Hybrid Learning of IT2 IFLS

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 IT2 IFS.*
2: *set the number of training epochs to unity*
3: *set the training sample point (p) to unity*
4: *propagate the input (x_p) through the IT2 IFLS hybrid model*
5: *tune the consequent parameters using DEKF according to Equations (36) to (37) for MF parameters and Equations (39) to (40) for NMF parameters*
6: *compute the output of the hybrid IT2 IFLS using Equation (17)*
7: *compute the difference between the actual output and predicted output of the hybrid IT2 IFLS model and use RMSE as the cost function*
8: *back-propagate the error*
9: *tune the antecedent parameters using gradient descent back-propagation algorithm according to Equations (41) to (43)*
10: *increment the training sample point by 1 (x_{p+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** the stopping criterion is reached **then**
17: *End*
18: **else**
19: *go to step 4*
20: **end if**

V. Experimental Analysis and Evaluation

In this section, we present our experimental analysis on publicly available benchmark time series and system identification problems. The datasets and the criteria used in the evaluation (RMSE and runtime) were carefully selected to facilitate comparison of the approach introduced here with existing methods. The RMSE is defined as:

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

where y^a is the desired output and y is the predicted output, N is the number of testing data points. For each input in this study, two IT2 IFS are utilised. The β value for all experiments is initialised to 0.5. The initial values of MF and NMF parameters are randomly generated from unit interval [0,1]. The IF-indices are m-by-n matrices randomly generated from the unit interval [0,1] for all experiments, where m is the number of linguistic terms and n is the number of rules. The entire experiments were conducted using $MATLAB^{©}$ 2016 running on a 64-bit Intel core i3-4130 CPU@3.40GHz /8GB RAM configuration computer.

A. Application to Artificially Generated Datasets

1) Mackey-Glass Time Series: Mackey-Glass benchmark time series for modeling a physiological system defined by the

differential delay equation in (45) is examined:

$$\frac{dx\left(t\right)}{dt} = \frac{a * x\left(t-\tau\right)}{1 + x\left(t-\tau\right)^{n}} - b * x\left(t\right) \tag{45}$$

where a, b and n are constant values, t is the current time and τ is the time delay constant. We evaluate the proposed model with $\tau = 17$. Similar to [34], [52]–[55], a dataset consisting of 1000 data points are generated using Equation (45). 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 184 tunable parameters. The KF parameters Q and P for both MF and NMF were initially set as $0.001I_{80}$ and $1.0I_{80}$ respectively with R = 1.0. The learning rate is fixed at 0.01 with 500 training epochs and 10 simulation runs. Figure 2 shows the actual and the predicted output while Figure 3 shows the evolution of the adaptive user define parameter, β, for Mackey-Glass application problem. Comparison of results is made between IT2 IFLS trained with DEKF and GD, and its type-1 variants on Mackey-Glass benchmark dataset. Table II shows that IT2 IFLS outperforms its type-1 counterpart.

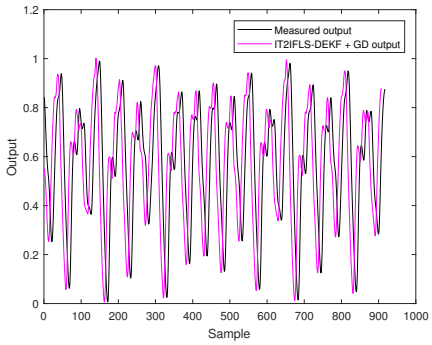

Fig. 2: Actual and predicted output of Mackey-Glass time series using hybrid IT2 IFLS

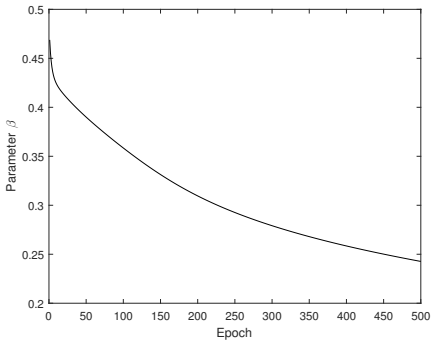

Fig. 3: The adaptation of the parameter β for Mackey-Glass prediction problem

A comparison of the hybrid learning approach of IT2 IFLS with some existing models in the literature is also shown in Table II with IT2 IFLS 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 (PSO) with diversity learning and GD (LLWNN + hybrid) and LLWNN trained with GD (LLWNN + GD) in this problem domain. The NMFs allow IT2 IFLS to capture more information than the conventional IT2 FLSs 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 the conventional IT2 FLSs.

TABLE II: Performance Comparison of Mackey-Glass Time Series Forecasting

Model	Rules	RMSE
SuPFuNIS [56]	15	0.014
Fuzzy-Singular Value Decomposition [57]	- 10	0.012
MDE-RBF NN [58]	-	0.013
Genetic Fuzzy Ensemble [59]	-	0.0264
Radial Basis Function AFS [60]	-	0.0114
RBF-AFS [60]	21	0.013
HyFIS [61]	16	0.012
NEFPROX [62]	129	0.0332
HyFIS-Yager-gDIC [63]		0.0190
T2-HyFIS-Yager [63]		0.0694
D-FNN [64]	10	0.008
WNN + gradient [52]	-	0.0071
WNN + hybrid [52]	-	0.0059
LLWNN + gradient [52]	-	0.0041
LLWNN + hybrid [52]	-	0.0036
MLMVN [53]	-	0.0056
GEFREX [65]	-	0.0061
SA-T2 FLS [54]	16	0.0089
TSK-SVR I [55]	-	0.008
TSK-SVR II [55]	-	0.007
IT2 FNN-1 [66]	16	0.0050
IT2 FNN-2 [66]	16	0.0035
IT2 FNN-3 [66]	16	0.0020
NNT2FW [67]		0.056
NNT2FWGA [67]		0.0431
NNT2FWPSO [67]		0.0456
IFLS - DEKF+GD	16	0.0054
IT2 IFLS -DEKF+GD	16	0.0040

2) System Identification Problem #1: A second-order time-varying system is investigated using the hybrid learning model of IT2 IFLS. This first identification problem involves a dynamic system that is defined by Equation (46).

$$y(t+1) = f(y(t), y(t-1), y(t-2), u(t), u(t-1)) \tag{46}$$

where

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

where a, b, c are time-varying parameters defined as in Equations (47) to (49).

$$a(t) = 1.2 - 0.2cos(2\pi t/T) \tag{47}$$

$$b(t) = 1.0 - 0.4sin(2\pi t/T) \tag{48}$$

$$c(t) = 1.0 + 0.4sin(2\pi t/T) \tag{49}$$

Here, $T = 1000$ represents the total number of sample points. All computational procedures are arranged as closely as possible to those reported in [68]–[70]. Two inputs values are utilised which are $u(t)$ and $y(t)$.

$$u(t) = \begin{cases} \sin(\pi t/25) & t < 250 \\ 1.0, & 250 \leq t < 500 \\ -1.0 & 500 \leq t < 750 \\ 0.3\sin(\pi t/25) + 0.1\sin(\pi t/32) & \\ +0.6\sin(\pi t/10) & 750 \leq t < 1000 \end{cases} \tag{50}$$

Similar to [69], the simulation is conducted for 1000 time steps with 100 training epochs. A total of 4 rules with 36 tunable parameters are generated. The learning rate was set to 0.01 while the KF parameters for P and Q are initially chosen as $1I_{12}$ and $0.001I_{12}$ respectively for the MF and NMF 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 IT2 IFLS-DEKF and GD on the time-varying dynamic system, the test signal in Equation (50) is used.

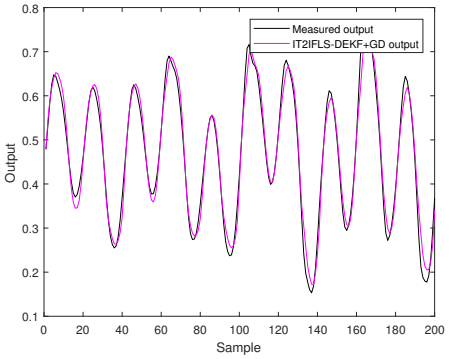

Fig. 4: Actual and predicted output using hybrid IT2 IFLS for second-order system identification problem #1

Figure 4 shows the actual versus the predicted output for 200 data points of the second-order identification problem #1 using IT2 IFLS-DEKF and GD. As shown in Table III, the hybrid model of IT2 IFLS-DEKF and GD outperforms other existing models except interval T2 fuzzy neural network (IT2 FNN) trained with EKF (IT2 FNN-EKF). Although IT2 IFLS-DEKF and GD performs better than IT2 FNN-EKF on the training set, IT2 FNN-EKF outperforms IT2 IFLS-DEKF and GD on the test set. This could be as a result of utilising the predictive power of EKF on both the antecedent and consequent parameters tuning of IT2 FNN-EKF. Most notably is the comparison of IT2 IFLS-DEKF and GD with self evolving IT2 FNN (SEIT2 FNN) and TSK-type-based self

TABLE III: Performance Comparison of hybrid-IT2 IFLS for second-order system identification #1

Model	Rules	Epoch	Training RMSE	Testing RMSE
Type-1 TSK FNS [68]	9	100	0.0282	0.0598
Type-2 TSK FNS [68]	4	100	0.0284	0.0601
Feedorward Type-2 FNN	3	100	0.0281	0.0593
SIT2FNN [70]	4	100	0.0351	0.0560
SEIT2 FNN [71]	3	100	0.0274	0.0574
TSCIT2FNN [69]	3	100	0.0279	0.0576
IT2 FNN-GD [48]	-	200	0.0540	0.0613
IT2 FNN-EKF [48]	-	200	0.0275	0.0261
IT2 FNN-SMC [48]	-	200	0.0360	0.0390
IT2 FNN- PSO + SMC [48]	-	200	0.0199	0.0390
IT2 IFLS - DEKF+GD	4	100	0.0250	0.0310

evolving compensatory IT2 FNN (TSCIT2 FNN). Similar to IT2 IFLS, both SEIT2 FNN and TSCIT2 FNN utilise KF-based methodology to adapt their consequent parameters and GD to optimize the antecedent parameters respectively with A2-C0 TSK-type fuzzy inference. The proposed framework of IT2 IFLS outperforms both existing methods of SEIT2 FNN and TSCIT2 FNN in this problem instance. For a fair comparison of the runtime of IT2 IFLS - DEKF and GD with those reported in [48], 200 simulations of the experiments are conducted. As shown in Table IV, IT2 IFLS-DEKF and GD has the lowest runtime of 82.04 seconds, close to that of IT2 FNN trained with sliding mode control (IT2 FNN-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. Thus, with the superior identification accuracy and computational efficiency in terms of run time, the proposed IT2 IFLS-DEKF and GD model is therefore more appropriate for real time applications.

TABLE IV: Comparison of runtime of IT2 IFLS with other approaches on second-order identification problem #1

Model	Epoch	Run Time (s)
IT2 FNN-GD [48]	200	124.12
IT2 FNN-EKF [48]	200	229.71
IT2 FNN-SMC [48]	200	84.39
IT2 FNN PSO + SMC [48]	200	7086.78
IT2 IFLS DEKF+GD	200	82.04

3) System Identification Problem #2: The proposed hybrid model, IT2 IFLS - DEKF and GD is applied to a dynamic

system with dataset generated by the differential equation [71]:

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

where

$$f(t) = \begin{cases} 0, & 1 \leq t \leq 1000 \\ 1.0, & 1001 \leq t \leq 2000 \\ 0, & 2001 \leq t \end{cases} \quad (51)$$

The inputs to the proposed 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 36 tunable parameters for the IT2 IFLS-DEKF and GD model. A TSK type-1 intuitionistic fuzzy logic model (IFLS-TSK) and a TSK-type IT2 FLS (IT2 FLS-TSK) trained with DEKF and GD are also constructed and evaluated on the system identification problem #2. The number of rules in the three models remain the same with 32 and 24 tunable parameters for the IFLS and IT2 FLS respectively. The RMSE is computed over 10 simulations for each model. Shown in Figure 5 is the actual and predicted outputs of the identification problem #2 for 100 test samples using IT2 IFLS - DEKF and GD. As presented in Table V, IT2 IFLS outperforms both IFLS and IT2 FLS.

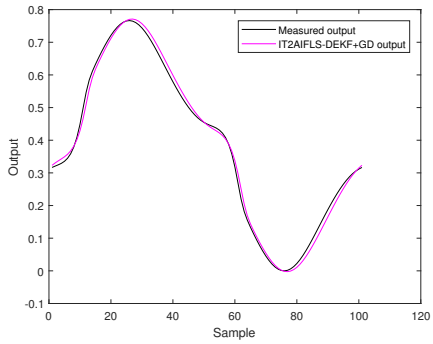

Fig. 5: Actual and predicted output using hybrid IT2 IFLS for identification problem #2

TABLE V: Comparison of IT2 IFLS vs IFLS and IT2 FLS on second-order identification problem #2

Model	Trn RMSE	Tst RMSE
IT2 FLS-DEKF+GD	0.0173	0.0074
IFLS-DEKF+GD	0.0172	0.0073
IT2 IFLS-DEKF+GD	0.0151	0.0064

B. Application to Real World Problems

In this section, the performance of the proposed model on three real world problems is analysed. These are Poland electricity load, Santa-Fe laser and Box-Jenkins gas furnace datasets.

1) Electricity Load Forecasting: Similar to system identification problem #2, this experiment is conducted to evaluate the performance of hybrid learning of IT2 IFLS with IFLS and IT2 FLS using the same learning procedure on a real world problem. The dataset selected is the Poland electricity load dataset obtained from (http://research.cs.aalto.fi/) and contains electricity load values of Poland in the 1990's. 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 (52).

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.

$$\hat{x}_{t+1} = f(x(t), x(t-1), \cdots, x(t-p+1))$$

where p is the size of input with $t \geq p$. We adopt the input size of four and the input generating vector becomes:

$$\hat{x}_{t+1} = f(x(t), x(t-1), x(t-2), x(t-3)) \quad (52)$$

with $\hat{x}_{t+1}$ as the output. Figure 6 shows the training dataset

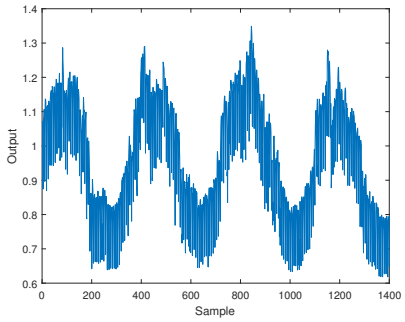

Fig. 6: Training dataset of Poland electricity load

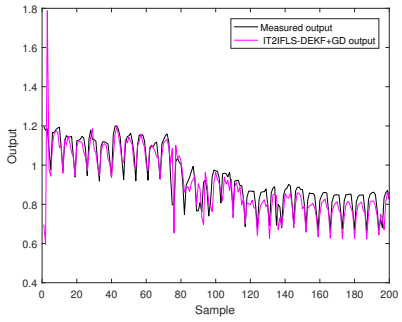

Fig. 7: Actual and predicted values of Poland electricity load with IT2 IFLS-DEKF+GD using test dataset

for Poland electricity load while Figure 7 shows the actual and the predicted values of the test dataset. Table VI shows that the performance of IT2 IFLS is superior to those of IFLS and IT2FLs trained with the same hybrid algorithm of DEKF and GD.

TABLE VI: Comparison of IT2 IFLS vs IFLS and IT2 FLS on Poland electricity load forecast using hybrid algorithm of DEKF and GD

Model	train/chk	Trn RMSE	Tst RMSE
IT2 FLS - DEKF+GD	1395/196	0.0564	0.0595
IFLS-DEKF+GD	1395/196	0.0589	0.0599
IT2 IFLS-DEKF+GD	1395/196	0.0560	0.0572

2) Box-Jenkins Time Series: The Box-Jenkins time series is one of the most used benchmark datasets for model evaluation. It is a gas furnace dataset 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 Box-Jenkins dataset was downloaded from (http://openmv.net/info/gas-furnace). The dataset consist of 296 data pairs. For ease of comparison with earlier studies, the simulation settings are arranged to be as close as possible to those reported in [52], [72], [73]. The task is to forecast the amount of CO_2 concentration in the gas at time (t) using 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. As shown in Table VII, IT2 IFLS 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 IT2 IFLS performing better than or comparatively with other works in the literature.

3) Santa-Fe Time Series: The proposed hybrid IT2 IFLS model is also applied to the Santa-Fe time series in order to evaluate the performance of the hybrid model on another real world scenario. The time series is downloaded from (http://www-psych.stanford.edu/ andreas/Time-Series/SantaFe.html). The univariate Santa-Fe dataset was obtained from a physical system in the laboratory in a chaotic state. To aid comparison with previous studies, the experimental set-up are arranged as closely as possible to those reported in [85], [86]. From the Santa-fe time series, 1000 input-output data pairs are generated using the format: $[y(t-1), y(t-2), y(t-3), y(t-4), y(t-5); y(t)]$ giving five inputs and one output $y(t)$. Similar to [86], all samples are normalised to unit interval $[0, 1]$ so that larger values do not overwhelm the smaller values.

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 consequent parameters w and b are randomly generated in the unit interval [0,1] while the KF parameters P and Q are chosen as $1.0I_{192}$ (I_{192} is the 192 by 192 identity matrix) and $0.001I_{192}$ for MF and NMF respectively and R is chosen as 1.0. The user design parameter β is initially set as 0.5 with the learning rate is set to 0.1. There are 32 rules generated with 6(5) + 2*32(5+1) = 414 tunable parameters.

TABLE VII: Performance comparison of hybrid-IT2 IFLS on Box-Jenkins time series

Model	Rules	No. of Parameters	Testing RMSE
ARMA [74]	-	-	0.843
Tongs' model [75]	19	-	0.685
Pedrycz's model [76]	81	-	0.566
Xu's model [77]	25		0.573
Sugeno's model [78]	6	-	0.596
Surmann's model [79]	25	-	0.400
Lee's model [80]	25	-	0.638
Lin's model [81]	4		0.511
Nie's model [82]	45	225	0.412
ANFIS [83]	4	24	0.085
Neural Tree [84]	-	-	0.0257
WNN + gradient [52]		40	0.084
WNN + hybrid [52]		40	0.081
LWNN + gradient [52]		56	0.01643
LWNN + hybrid [52]		56	0.01378
FWNN-S (2MFs) [73]	-	32	0.03085
FWNN-S (3MFs) [73]	-	66	0.02778
FWNN-R (2MFs) [73]	-	28	0.03171
FWNN-R (3MFs) [73]	-	57	0.02794
FWNN-M (2MFs) [73]	-	32	0.02963
FWNN-M (3MFs) [73]	-	66	0.02324
LLNF (2 inputs) [72]	-	-	0.0462
OSSA-LLNF [72] (2 inputs)		-	0.0321
IFLS-DEKF+GD	4	32	0.0273
IT2 IFLS DEKF+GD 2(MFs)	4	36	0.0249

In Table VIII the results obtained from IT2 IFLS and IFLS, both trained with DEKF and GD are shown together with other existing approaches in the literature. As shown in Table VIII, IT2 IFLS model trained with DEKF and GD outperforms IFLS trained with the same hybrid algorithm. The IT2 IFLS-DEKF and GD also performs better than other models in the literature with very low RMSE on the test set, thus demonstrating a good generalisation of the proposed hybrid learning model.

VI. Conclusion

The EKF and GD-based approaches have been used to learn the parameters of fuzzy systems. This study presents a novel application of a hybrid approach of DEKF and GD to optimise the parameters of a newly developed IT2 IFLS-TSK. 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 IFLS and IT2 FLS for performance comparison between its type-1 version and conventional IT2 FLS.

TABLE VIII: Performance comparison of hybrid IT2 IFLS with other models on Santa-Fe time series A dataset

Model	Rule Number	Parameter number	Training RMSE	Test RMSE
ES [85]	-	-	-	56.20
NN [85]	-	-	-	24.6
PMRS [85]	-	-	-	14.23
SONFIN* [87]	9	144	6.956	5.983
T2 FLS-G* [88]	5	135	8.50	7.16
SEIT2 FNN* [71]	5	135	7.677	5.766
IT2 FNN-SVR(N) [86]	5	106	13.565	4.337
IT2 FNN-SVR(F) [86]	5	106	9.094	3.474
SVR-FM* ($\epsilon = 0.1$) [89]	31	188	14.370	9.707
SVR-FM* ($\epsilon = 0.001$) [89]	747	4484	7.069	1.650
IT2 IFLS-DEKF+GD	32	414	6.075	1.668

* These results are adapted from [86]

From simulation analyses, IT2 IFLS exhibits superior performance quality to those of IFLS and IT2 FLS trained with DEKF and GD. Overall, the developed hybrid learning model of IT2 IFLS-TSK exhibits better or comparatively good prediction and identification performances compared to similar studies in the literature. The run time of the proposed IT2 IFLS - DEKF and GD is very short compared to other previous models on the same problem domain. This is an indication that the proposed hybrid learning model may be more appropriate for real time systems.

In the future, we intend to learn the parameters of the IT2 IFLS using other hybrid approaches such as PSO and GD, PSO and EKF, and compare their performances. In this initial study, we have investigated the performance of IT2 IFLS using the product $t-norm$. It will be interesting in the future to see the effects other aggregation functions such as minimum $t-norm$ and Lukasiewicz $t-norm$ will have on the performance of the system. Furthermore, with the inference mechanism of IT2 IFLS represented in closed form, an interesting research opportunity therefore presents itself; which is the stability analysis of the IT2 IFLS. This is a non-trivial problem and we intend to explore the proposed model further in this direction.

ACKNOWLEDGEMENT

This research work is funded by the Government of Nigeria under the Tertiary Education Trust Fund (TETFund).

REFERENCES

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

[2] 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.

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

[4] A. K. T. Cornelis, Chris and E. Kerre, "Intuitionistic fuzzy sets and intervalvalued fuzzy sets: a critical comparison," in *Proc. Eusflat03*. Citeseer, 2003.

[5] K. Atanassov and G. Gargov, "Interval valued intuitionistic fuzzy sets," *Fuzzy sets and systems*, vol. 31, no. 3, pp. 343–349, 1989.

[6] 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.

[7] 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.

[8] D. Dubois and H. Prade, "Interval-valued fuzzy sets, possibility theory and imprecise probability." in *EUSFLAT Conf.*, 2005, pp. 314–319.

[9] 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.

[10] 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.

[11] O. Castillo, P. Melin, R. Tsvetkov, and K. Atanassov, "Short remark on interval type-2 fuzzy sets and intuitionistic fuzzy sets," in *Proceedings of the 18th International Conference on IFSs*, 2014, pp. 1–5.

[12] O. Castillo, P. Melin, R. Tsvetkov, and K. T. Atanassov, "Short remark on fuzzy sets, interval type-2 fuzzy sets, general type-2 fuzzy sets and intuitionistic fuzzy sets," in *Intelligent Systems' 2014*. Springer, 2015, pp. 183–190.

[13] 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.

[14] 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.

[15] 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.

[16] 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)*, 2013, pp. 299–304.

[17] 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)*, 2013, pp. 207–212.

[18] 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*. IEEE, 2012, pp. 1–8.

[19] 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*. IEEE, 2013, pp. 1–8.

[20] 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.

[21] J. M. Mendel, "Uncertain rule-based fuzzy logic system: introduction and new directions," 2001.

[22] 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.

[23] 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.

[24] H. Hagras, "Type-2 FLCs: a new generation of fuzzy controllers," *Computational Intelligence Magazine, IEEE*, vol. 2, no. 1, pp. 30–43, 2007.

[25] J. M. Mendel, "Computing with words: Zadeh, turing, popper and occam," *IEEE computational intelligence magazine*, vol. 2, no. 4, pp. 10–17, 2007.

[26] 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.

[27] 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.

[28] E. Kayacan and R. Maslim, "Type-2 fuzzy logic trajectory tracking control of quadrotor vtol aircraft with elliptic membership functions," *IEEE/ASME Transactions on Mechatronics*, vol. 22, no. 1, pp. 339–348, 2017.

[29] 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.

[30] 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.

[31] 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.

[32] 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.

[33] E. Barrenechea, M. Pagola, J. Fernandez, and J. Sanz, "Generalized atanassovs intuitionistic fuzzy index. construction method," in *IFSA-EUSFLAT conf.*, 2009, pp. 478–482.

[34] 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*, 2016, pp. 1063–1068.

[35] 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.

[36] 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.

[37] 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.

[38] G. M. Mendez, A. Hernández, A. Cavazos, and M.-T. Mata-Jiménez, "Type-1 non-singleton type-2 takagi-sugeno-kang fuzzy logic systems using the hybrid mechanism composed by a kalman type filter and back propagation methods," in *International Conference on Hybrid Artificial Intelligence Systems*. Springer, 2010, pp. 429–437.

[39] 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.

[40] G. M. Mendez and M. Hernandez, "Interval type-1 non-singleton type-2 tsk fuzzy logic systems using the hybrid training method rls-bp," in *Foundations of Computational Intelligence, 2007. FOCI 2007. IEEE Symposium on*. IEEE, 2007, pp. 370–374.

[41] G. Mendez and M. De Los Angeles Hernandez, "Interval type-2 anfis," *Innovations in Hybrid Intelligent Systems*, pp. 64–71, 2007.

[42] G. M. Mendez and O. Castillo, "Interval type-2 tsk fuzzy logic systems using hybrid learning algorithm," in *Fuzzy Systems, 2005. FUZZ'05. The 14th IEEE International Conference on*. IEEE, 2005, pp. 230–235.

[43] 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.

[44] 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.

[45] D. Simon, "Training fuzzy systems with the extended kalman filter," *Fuzzy sets and systems*, vol. 132, no. 2, pp. 189–199, 2002.

[46] 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, 2011, pp. 262–265.

[47] P. Hájek and V. Olej, "Intuitionistic fuzzy neural network: The case of credit scoring using text information," in *Engineering Applications of Neural Networks*. Springer, 2015, pp. 337–346.

[48] E. Kayacan and M. Khanesar, *Fuzzy Neural Networks for Real Time Control Applications: Concepts, Modeling and Algorithms for Fast Learning*. Elsevier Science, 2015. [Online]. Available: https://books.google.co.uk/books?id=8X4VBgAAQBAJ

[49] 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. IEEE, 2004, pp. 4146–4150.

[50] 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.

[51] D. Simon, "Training radial basis neural networks with the extended kalman filter," *Neurocomputing*, vol. 48, no. 1, pp. 455–475, 2002.

[52] 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.

[53] 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.

[54] 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, 2011, pp. 976–980.

[55] 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)*, 2014, pp. 2396–2401.

[56] 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.

[57] 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.

[58] 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.

[59] 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.

[60] 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.

[61] 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.

[62] 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. IEEE, 1998, pp. 1106–1111.

[63] 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*. IEEE, 2009, pp. 80–85.

[64] 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.

[65] M. Russo, "Genetic fuzzy learning," *IEEE transactions on evolutionary computation*, vol. 4, no. 3, pp. 259–273, 2000.

[66] 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.

[67] F. Gaxiola, P. Melin, F. Valdez, J. R. Castro, and O. Castillo, "Optimization of type-2 fuzzy weights in backpropagation learning for neural networks using gas and pso," *Applied Soft Computing*, vol. 38, pp. 860–871, 2016.

[68] 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.

[69] 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.

[70] 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.

[71] 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.

[72] 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.

[73] 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.

[74] G. E. Box and G. M. Jenkins, *Time series analysis: forecasting and control, revised ed*. Holden-Day, 1976.

[75] R. Tong, "The evaluation of fuzzy models derived from experimental data," *Fuzzy sets and systems*, vol. 4, no. 1, pp. 1–12, 1980.

[76] W. Pedrycz, "An identification algorithm in fuzzy relational systems," *Fuzzy sets and systems*, vol. 13, no. 2, pp. 153–167, 1984.

[77] 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.

[78] M. Sugeno and T. Yasukawa, "Linguistic modeling based on numerical data," in *IFSA*, vol. 91, 1991, pp. 264–267.

[79] 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.

[80] 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.

[81] 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.

[82] 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.

[83] J.-S. R. Jang, C.-T. Sun, and E. Mizutani, "Neuro-fuzzy and soft computing, a computational approach to learning and machine intelligence," 1997.

[84] 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.

[85] S. Singh, "Noise impact on time-series forecasting using an intelligent pattern matching technique," *Pattern Recognition*, vol. 32, no. 8, pp. 1389–1398, 1999.

[86] 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.

[87] 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.

[88] 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.

[89] 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.

REFERENCES

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

[2] 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.

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

[4] A. K. T. Cornelis, Chris and E. Kerre, "Intuitionistic fuzzy sets and intervalvalued fuzzy sets: a critical comparison," in *Proc. Eusflat03*. Citeseer, 2003.

[5] K. Atanassov and G. Gargov, "Interval valued intuitionistic fuzzy sets," *Fuzzy sets and systems*, vol. 31, no. 3, pp. 343–349, 1989.

[6] 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.

[7] 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.

[8] D. Dubois and H. Prade, "Interval-valued fuzzy sets, possibility theory and imprecise probability." in *EUSFLAT Conf.*, 2005, pp. 314–319.

[9] 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.

[10] 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.

[11] O. Castillo, P. Melin, R. Tsvetkov, and K. Atanassov, "Short remark on interval type-2 fuzzy sets and intuitionistic fuzzy sets," in *Proceedings of the 18th International Conference on IFSs*, 2014, pp. 1–5.

[12] O. Castillo, P. Melin, R. Tsvetkov, and K. T. Atanassov, "Short remark on fuzzy sets, interval type-2 fuzzy sets, general type-2 fuzzy sets and intuitionistic fuzzy sets," in *Intelligent Systems' 2014*. Springer, 2015, pp. 183–190.

[13] 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.

[14] 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.

[15] 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.

[16] 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)*, 2013, pp. 299–304.

[17] 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)*, 2013, pp. 207–212.

[18] 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*. IEEE, 2012, pp. 1–8.

[19] 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*. IEEE, 2013, pp. 1–8.

[20] 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.

[21] J. M. Mendel, "Uncertain rule-based fuzzy logic system: introduction and new directions," 2001.

[22] 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.

[23] 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.

[24] H. Hagras, "Type-2 FLCs: a new generation of fuzzy controllers," *Computational Intelligence Magazine, IEEE*, vol. 2, no. 1, pp. 30–43, 2007.

[25] J. M. Mendel, "Computing with words: Zadeh, turing, popper and occam," *IEEE computational intelligence magazine*, vol. 2, no. 4, pp. 10–17, 2007.

[26] 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.

[27] 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.

[28] E. Kayacan and R. Maslim, "Type-2 fuzzy logic trajectory tracking control of quadrotor vtol aircraft with elliptic membership functions," *IEEE/ASME Transactions on Mechatronics*, vol. 22, no. 1, pp. 339–348, 2017.

[29] 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.

[30] 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.

[31] 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.

[32] 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.

[33] E. Barrenechea, M. Pagola, J. Fernandez, and J. Sanz, "Generalized atanassovs intuitionistic fuzzy index. construction method," in *IFSA-EUSFLAT conf.*, 2009, pp. 478–482.

[34] 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*, 2016, pp. 1063–1068.

[35] 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.

[36] 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.

[37] 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.

[38] G. M. Mendez, A. Hernández, A. Cavazos, and M.-T. Mata-Jiménez, "Type-1 non-singleton type-2 takagi-sugeno-kang fuzzy logic systems using the hybrid mechanism composed by a kalman type filter and back propagation methods," in *International Conference on Hybrid Artificial Intelligence Systems*. Springer, 2010, pp. 429–437.

[39] 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.

[40] G. M. Mendez and M. Hernandez, "Interval type-1 non-singleton type-2 tsk fuzzy logic systems using the hybrid training method rls-bp," in *Foundations of Computational Intelligence, 2007. FOCI 2007. IEEE Symposium on*. IEEE, 2007, pp. 370–374.

[41] G. Mendez and M. De Los Angeles Hernandez, "Interval type-2 anfis," *Innovations in Hybrid Intelligent Systems*, pp. 64–71, 2007.

[42] G. M. Mendez and O. Castillo, "Interval type-2 tsk fuzzy logic systems using hybrid learning algorithm," in *Fuzzy Systems, 2005. FUZZ'05. The 14th IEEE International Conference on*. IEEE, 2005, pp. 230–235.

[43] 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.

[44] 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.

[45] D. Simon, "Training fuzzy systems with the extended kalman filter," *Fuzzy sets and systems*, vol. 132, no. 2, pp. 189–199, 2002.

[46] 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, 2011, pp. 262–265.

[47] P. Hájek and V. Olej, "Intuitionistic fuzzy neural network: The case of credit scoring using text information," in *Engineering Applications of Neural Networks*. Springer, 2015, pp. 337–346.

[48] E. Kayacan and M. Khanesar, *Fuzzy Neural Networks for Real Time Control Applications: Concepts, Modeling and Algorithms for Fast Learning*. Elsevier Science, 2015. [Online]. Available: https://books.google.co.uk/books?id=8X4VBgAAQBAJ

[49] 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. IEEE, 2004, pp. 4146–4150.

[50] 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.

[51] D. Simon, "Training radial basis neural networks with the extended kalman filter," *Neurocomputing*, vol. 48, no. 1, pp. 455–475, 2002.

[52] 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.

[53] 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.

[54] 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, 2011, pp. 976–980.

[55] 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)*, 2014, pp. 2396–2401.

[56] 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.

[57] 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.

[58] 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.

[59] 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.

[60] 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.

[61] 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.

[62] 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. IEEE, 1998, pp. 1106–1111.

[63] 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*. IEEE, 2009, pp. 80–85.

[64] 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.

[65] M. Russo, "Genetic fuzzy learning," *IEEE transactions on evolutionary computation*, vol. 4, no. 3, pp. 259–273, 2000.

[66] 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.

[67] F. Gaxiola, P. Melin, F. Valdez, J. R. Castro, and O. Castillo, "Optimization of type-2 fuzzy weights in backpropagation learning for neural networks using gas and pso," *Applied Soft Computing*, vol. 38, pp. 860–871, 2016.

[68] 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.

[69] 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.

[70] 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.

[71] 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.

[72] 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.
[73] 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.
[74] G. E. Box and G. M. Jenkins, *Time series analysis: forecasting and control, revised ed*. Holden-Day, 1976.
[75] R. Tong, "The evaluation of fuzzy models derived from experimental data," *Fuzzy sets and systems*, vol. 4, no. 1, pp. 1–12, 1980.
[76] W. Pedrycz, "An identification algorithm in fuzzy relational systems," *Fuzzy sets and systems*, vol. 13, no. 2, pp. 153–167, 1984.
[77] 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.
[78] M. Sugeno and T. Yasukawa, "Linguistic modeling based on numerical data," in *IFSA*, vol. 91, 1991, pp. 264–267.
[79] 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.
[80] 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.
[81] 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.
[82] 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.
[83] J.-S. R. Jang, C.-T. Sun, and E. Mizutani, "Neuro-fuzzy and soft computing, a computational approach to learning and machine intelligence," 1997.
[84] 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.
[85] S. Singh, "Noise impact on time-series forecasting using an intelligent pattern matching technique," *Pattern Recognition*, vol. 32, no. 8, pp. 1389–1398, 1999.
[86] 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.
[87] 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.
[88] 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.
[89] 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.