energies
MDPI

Article

Electrical Load Prediction using Interval Type-2 Atanassov Intuitionist Fuzzy System: Gravitational Search Algorithm Tuning approach

M. A. Khanesar [1,*], Jingyi Lu [2],Thomas Smith [3], and David Branson [4]

[1,2,3,4] Faculty of Engineering, University of Nottingham, Nottingham NG7 2RD, UK; {ezzma5, eayjl19, epxtas, ezzdtb} @exmail.nottingham.ac.uk
* Correspondence: ezzma5@exmail.nottingham.ac.uk;

Abstract: Establishing accurate electrical load prediction is vital for pricing and power system management. However, the unpredictable behavior of private and industrial users results in uncertainty in these power systems. Furthermore, the utilization of renewable energy sources, which are often variable in their production rates, also increases the complexity making predictions even more difficult. In this paper an interval type-2 intuitionist fuzzy logic systems whose parameters are trained in a hybrid fashion using gravitational search algorithms with the ridge least square algorithm is presented for short term prediction of electrical loading. Simulation results are provided to compare the performance of the proposed approach with that of state-of-the-art electrical load prediction algorithms for Poland, and five regions of Australia. The simulation results demonstrate the superior performance of the proposed approach over seven different current state-of-the-art prediction algorithms in literature, namely: SVR, ANN, ELM, EEMD-ELM-GOA, EEMD-ELM-DA, EEMD-ELM-PSO and EEMD-ELM-GWO.

Keywords: Electrical load prediction, interval type-2 Atanassov intuitionist fuzzy logic system; ridge least square algorithm; gravitational search algorithm

Citation: A. Khanesar, F.; Lu, J.; Smith, T.; Branson, D. Short Term Electrical Load Prediction using Interval Type-2 Atanassov Intuitionist Fuzzy System . *Energies* **xxxx**, *xx*, x.

1. Introduction

Electrical power is vital for our life, it has illuminated our living areas, workplaces, and planet. Electrical load consumption in a region varies with several parameters such as population density, wealth, social factors, climate and distribution of use between home and industry [1]. It is required to have an accurate prediction of electrical load to continue to generate electrical power to meet demand and decrease the cost of electrical power delivery. In the UK, a study has shown that *one percent* load prediction error is equivalent to *10 million* GBP in operating costs per year for the UK power system [2] [3]. Electricity generation companies use load forecasting data from a couple of hours to a week ahead to produce energy at what is perceived to be the correct volume and to plan maintenance [4]. Traditional flat methods of pricing electrical load are not efficient as it results in high peaks corresponding to people's social behavior and the habits of industry [4]. To avoid such peaks, manage the network and reduce bills, newer methods such as variable peak pricing tariff and real time pricing are proposed which heavily depend upon real-time load prediction. The output of the prediction algorithms is further utilized in real time pricing, decision making processes and control algorithms for power system management purposes, improving the performance of predictors is of high interest. Renewable energy sources such as solar energy [5], wind power [6] and offshore energy [7] further increase the complexity and makes predictions even more difficult due to their natural power fluctuations. Soft sensors technology [8] can estimate solar irradiation [5] which may contribute to decreasing the uncertainty introduced by this source of energy to the overall power system. However, along with reducing the uncertainty associated with renewable energy sources, usage of more powerful prediction approaches to deal with these uncertainties is highly desirable.

A variety of machine learning techniques are already used for electrical load prediction. Among classical approaches, the autoregressive integrated moving average method is most commonly used in literature [9][10]. However, more recent approaches to electricity load prediction have explored the use of fuzzy logic due to its power to describe the real world in terms of IF-THEN statements as well as its learning capabilities. Among fuzzy logic approaches, interval type-2 fuzzy logic systems used in [1] and interval type-2 Atanassov fuzzy logic systems (IT2AIFLS) used in [11] are previous approaches which have been used to predict electrical load data in different regions of Australia and Poland, respectively. The IT2AIFLSs benefit from more degrees of freedom compared to interval type-2 fuzzy logic systems, as in addition to membership grades they benefit from non-membership grades. Based on previous studies, the greater degrees of freedom in this type of fuzzy system may improve its overall performance [11], [12] and [13]. However, the success of such fuzzy logic systems is highly dependent on the methods used to solve the complex task of training them. This motivated us to challenge the use of new optimization methods to increase the performance of IT2AIFLSs.

Different methods currently exist to train IT2AIFLSs including gradient descent approaches as well as hybrid computational methods [12] [13]. A full gradient descent method for training this structure is investigated in [14] where statistical analysis supported the null hypothesis ($\alpha = 0.05$) that the IT2AIFLS outperforms type-1 Attanassov intuitionist fuzzy logic systems. Although gradient descent methods provide a computational approach to train IT2AIFLSs, they suffer from falling in a local minima and instability problems.

Hybrid computational methods instead use different training methods in the antecedent and consequent parts of an IT2AIFLS to improve performance. The hybrid training algorithms investigated in [11] use gradient descent for the antecedent part of the IT2AIFLS and Kalman filter for its consequent part. This prediction model is then used for prediction of the Mackey-Glass chaotic time series, the Santa–Fe time series and the Box–Jenkins time series, in addition to the electrical load consumption of Poland in the 1990s [11]. This approach has also been applied for the prediction of energy, stock market and house price datasets [13]. Hybrid training approaches utilizing intelligent optimization algorithms previously implemented on IT2FLSs include particle swarm optimization plus Kalman filter [15], particle swarm optimization plus recursive least square [15] and particle swarm optimization plus gradient descent [16]. However, to the best of the authors' knowledge a hybrid intelligent optimization approach for the antecedent part of IT2AIFLS and a computational method for its conclusion part parameters has not been investigated in literature. Hybrid training methods offer better performance than solely using a computational method as they do not have local minima and instability issues. Since intelligent optimization approaches are more appropriate options for the antecedent part parameters which appear nonlinearly in the output of an IT2AIFLS, they are more appropriate options for training them.

In this paper, a new hybrid training approach is proposed to improve the performance of the hybrid training methods. The gravitational search algorithm (GSA) is chosen for optimizing the antecedent part parameters of the IT2FLS. The GSA is a physics inspired optimization algorithm [17] which defines each solution in terms of an object with mass, position, velocity and acceleration. The mass assigned to each object is proportional to its cost function and gravitational force among various objects absorbs masses with worst cost function to better object while scanning the space between them to find the optimum solution to the problem [18] [19]. The GSA was chosen due to its high performance, benefiting from multiple solutions and less probability to fall in a local minimum. This algorithm has already outperformed several optimization algorithms such as particle swarm optimization, the real genetic algorithm, the differential evolution algorithm and central force optimization in benchmark optimization problems [19][20][21]. Additionally, the ridge least square (R-LS) algorithm is chosen for the consequent part parameter which solves an l_2 cost function as a summation over identification error square plus the l_2-norm of the weights of the prediction system [22]. The R-LS algorithm is chosen as it is widely known to have superior generalization performance than the simple least square algorithm [22]. Comparisons are made between the proposed approach and some of existing approaches in the field including hybrid gradient descent Kalman filter training for an IT2AIFLS [11] for electrical load prediction in Poland. Another comparison is provided between the proposed estimation method and seven different estimation algorithms for electrical load prediction in five regions of Australia, namely: SVR, ANN, ELM, EEMD-ELM-GOA, EEMD-ELM-DA, EEMD-ELM-PSO and EEMD-ELM-GWO. The comparison results show the superior performance of the proposed approach over state-of-the-art approaches in literature. In particular, it outperforms the approach investigated in [11] which uses the IT2AIFLS but hybrid Gradient descent Kalman filter as the training algorithm.

This paper is organized as follows. Section II introduces the general structure of interval type-2 The methodology of this paper is presented in Section III. Simulation results are presented in Section IV. Finally concluding marks are presented in Section V.

2. Attanassov Intuitionist fuzzy system

Atanassov intuitionist fuzzy logic systems are a newer variant of the fuzzy logic family which have been successfully applied for prediction purposes [11] as well as pattern recognition [23]. Non-membership grade (ν) for an input to an ordinary fuzzy MF (μ) is simply calculated as the complement of its membership grade as ($\nu = 1 - \mu$). However, for an intuitionist fuzzy set ($\nu + \mu$) does not necessarily need to be equal to *one* [24]. This introduces a degree of hesitation or intuition for the fuzzy set. Atanassov in 1986 defined $\pi \in [0, 1]$ as the degree of hesitation which complements the membership and non-membership grades of an input such that $\pi + \nu + \mu = 1$.

2.1. Attanassov Intutionist fuzzy set

A fuzzy set is an ordinary fuzzy set if the degrees of membership and non-membership for every single input value add up to one. However, if the degrees of membership and non-membership do not add up to one for some input values, the fuzzy set is an intuitionist one. Let X be the universe of discourse and x be an individual value selected from it. The intuitionist fuzzy set $\tilde{A}^*$ is defined as presented in the followings:

$$\tilde{A}^* = \{< \mu_{\tilde{A}^*}(x), \nu_{\tilde{A}^*}(x) > | x \in X\} \quad (1)$$

where $\mu_{\tilde{A}^*}(x): X \to [0,\ 1]$ is the membership grade and $\nu_{\tilde{A}^*}(x): X \to [0,\ 1]$ is the non-membership grade for the input value $x \in X$, and we have $0 \le \mu_{\tilde{A}^*}(x) + \nu_{\tilde{A}^*}(x) \le 1$ [25]. In the special case, when $\nu_{\tilde{A}^*}(x) = 1 - \mu_{\tilde{A}^*}(x)$ the intutionist fuzzy set reduces to an ordinary fuzzy set. However, if this equality does not hold the intuition index of X in A is represented by $\pi_{\tilde{A}^*}(x)$ and defined by:

$$\pi_{A^*}(x) = max(0, (1 - (\mu_{\tilde{A}^*}(x) + \nu_{\tilde{A}^*}(x)))) \quad (2)$$

2.2 Structure of Attanassov Intutionist fuzzy system

Let T be the total number of inputs for the IT2AIFLS, with each sample containing $n-$dimensional input values $x \in R^n$ and $m-$dimensional output values $y \in R^m$. The membership functions considered for this structure are interval type-2 Gaussian MFs with uncertain σ values. The inference mechanism to calculate the output of the IT2AIFLS is demonstrated in Fig. 1 and is explained as follows:

Layer 1: The input layer is the first layer of this system which consists of n nodes passing input values to the fuzzification layer.

Layer 2: The fuzzification layer is the second layer which uses interval type-2 Atanassov membership functions. The inputs to this layer are the outputs of the previous layer and its outputs are the degrees of membership and non-membership which are themselves interval values. The $\pi_{c,ik}(x_i)$ represents the IF-index or hesitation of center and $\pi_{var,ik}(x_i)$ is the IF-index of variance and are defined by:

$$\pi_c(x) = max\left(0, \left(1 - \left(\mu_{\tilde{A}^*}(x) + \nu_{\tilde{A}^*}(x)\right)\right)\right)$$
$$\underline{\pi}_{var}(x) = max\left(0, (1 - (\underline{\mu}_{\tilde{A}^*}(x) + \overline{\nu}_{\tilde{A}^*}(x)))\right)$$
$$\overline{\pi}_{var}(x) = max\left(0, (1 - (\overline{\mu}_{\tilde{A}^*}(x) + \underline{\nu}_{\tilde{A}^*}(x)))\right) \quad (3)$$

such that $0 \leq \pi_c(x) \leq 1$, $0 \leq \underline{\pi}_{var}(x) \leq 1$ and $0 \leq \overline{\pi}_{var}(x) \leq 1$ and the degrees of membership as well as non-membership are defined as follows:

$$\overline{\mu}_{ik}(x_i) = exp\left(\frac{(x_i - c_{ik})^2}{2\sigma_{2,ik}^2}\right)\left(1 - \pi_{c,ik}(x_i)\right), i = 1, \ldots, n, k = 1, \ldots, M$$
$$\underline{\mu}_{ik}(x_i) = exp\left(\frac{(x_i - c_{ik})^2}{2\sigma_{1,ik}^2}\right)(1 - \pi_{c,ik}(x_i)), i = 1, \ldots, n, k = 1, \ldots, M$$
$$\overline{\nu}_{ik}(x_i) = (1 - \overline{\pi}_{var,ik}(x_i)) - \left[exp\left(\frac{(x_i - c_{ik})^2}{2\sigma_{1,ik}^2}\right)(1 - \pi_{c,ik}(x_i))\right], i = 1, \ldots, n, k = 1, \ldots, M$$
$$\underline{\nu}_{ik}(x_i) = (1 - \underline{\pi}_{var,ik}(x_i)) - \left[exp\left(\frac{(x_i - c_{ik})^2}{2\sigma_{2,ik}^2}\right)(1 - \pi_{c,ik}(x_i))\right], i = 1, \ldots, n, k = 1, \ldots, M \quad (4)$$

The parameters $\overline{\sigma}_{2i,k}$, $\underline{\sigma}_{1i,k}$, $\pi_{c,ik}$, $\pi_{var,ik}$ and c are premise part parameters associated with interval type-2 intuitionistic fuzzy MFs. Furthermore, *n* is the number of inputs to the system and *M* is the total number of rules in the fuzzy system.

<u>Layer 3</u>: The rule layer is the third layer of the IT2AIFLS which calculates the firing values of the rules of the fuzzy system are presented by [26]:

$$\overline{f}_k^{\mu}(x) = \overline{\mu}_{1k}(x_1) * \overline{\mu}_{2k}(x_2) * \ldots * \overline{\mu}_{nk}(x_n)$$
$$\underline{f}_k^{\mu}(x) = \underline{\mu}_{1k}(x_1) * \underline{\mu}_{2k}(x_2) * \ldots * \underline{\mu}_{nk}(x_n)$$
$$\overline{f}_k^{\nu}(x) = \overline{\nu}_{1k}(x_1) * \overline{\nu}_{2k}(x_1) * \ldots * \overline{\nu}_{nk}(x_n)$$
$$\underline{f}_k^{\nu}(x) = \underline{\nu}_{1k}(x_1) * \underline{\nu}_{2k}(x_2) * \ldots * \underline{\nu}_{nk}(x_n) \quad (5)$$

<u>Layer 4</u>: The output layer is the last layer of the system which performs the defuzzification + type reduction and calculates the output of the fuzzy system as follows.

$$y = \frac{\beta \sum_{k=1}^{M} \left(\underline{f}_k^{\mu} + \overline{f}_k^{\mu}\right) F_k^{\mu}}{\sum_{k=1}^{M} \underline{f}_k^{\mu} + \sum_{k=1}^{M} \overline{f}_k^{\mu}} + \frac{(1-\beta) \sum_{k=1}^{M} \left(\underline{f}_k^{\nu} + \overline{f}_k^{\nu}\right) F_k^{\nu}}{\sum_{k=1}^{M} \underline{f}_k^{\nu} + \sum_{k=1}^{M} \overline{f}_k^{\nu}} \quad (6)$$

where the parameter $\beta \in [0,\ 1]$ is the coefficient which determines which determines the weight of its corresponding terms in the output. Moreover, the firing of the rules corresponding to the membership functions are defined as follows:

$$r_k^{\mu} = \frac{\underline{f}_k^{\mu} + \overline{f}_k^{\mu}}{\sum_{k=1}^{M} \underline{f}_k^{\mu} + \sum_{k=1}^{M} \overline{f}_k^{\mu}} \quad (7)$$

and the ones corresponding to non-membership values are defined as follows:

$$r_k^{\nu} = \frac{\underline{f}_k^{\nu} + \overline{f}_k^{\nu}}{\sum_{k=1}^{M} \underline{f}_k^{\nu} + \sum_{k=1}^{M} \overline{f}_k^{\nu}} \quad (8)$$

Furthermore, F_k^{μ} and F_k^{ν} are defined as follows:

$$F_k^{\mu} = \sum_{i=1}^{n} \alpha_{ik}^{\mu} x_i + \beta_k^{\mu}$$
$$F_k^{\nu} = \sum_{i=1}^{n} \alpha_{ik}^{\nu} x_i + \beta_k^{\nu} \quad (9)$$

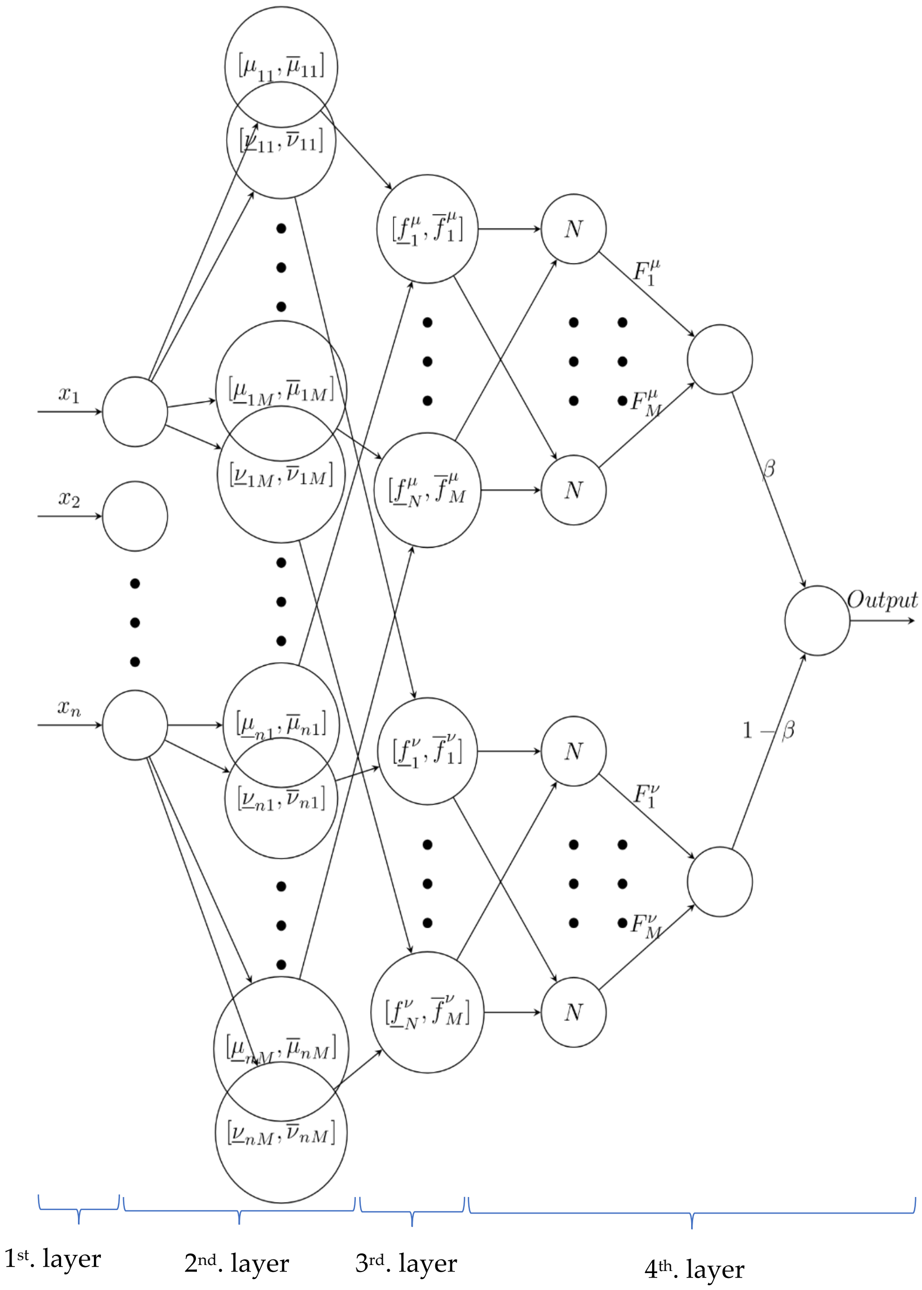


Figure 1. Structure of IT2AIFLS

The output of the IT2AIFLS can be rewritten as follows [27]:

$$y = \frac{\beta \sum_{k=1}^{M} r_k^{\mu} F_k^{\mu}}{\sum_{k=1}^{M} r_k^{\mu}} + \frac{(1-\beta) \sum_{k=1}^{M} r_k^{\nu} F_k^{\nu}}{\sum_{k=1}^{M} r_k^{\nu}} \tag{10}$$

The equation (10) can be further compacted to the following form.

$$y = \beta y^{\mu} + (1-\beta) y^{\nu} \tag{11}$$

To be able to apply ridge least square algorithm to update the parameters of the IT2AIFLS, we need to write it as follows:

$$y = \beta \theta^{\mu T} \phi^{\mu} + (1-\beta) \theta^{\nu T} \phi^{\nu} \tag{12}$$

where the parameters ϕ^{μ} and θ^{μ} are as follows:

$$\phi^{\mu} = [R^{\mu T} \quad R^{\mu T} x_1 \quad \dots \quad R^{\mu T} x_n]^T \tag{13}$$

$$R^{\mu} = \left[\frac{r_1^{\mu}}{\sum_{k=1}^{M} r_k^{\mu}} \quad \frac{r_2^{\mu}}{\sum_{k=1}^{M} r_k^{\mu}} \quad \dots \quad \frac{r_M^{\mu}}{\sum_{k=1}^{M} r_k^{\mu}}\right]$$

$$\theta_{(n+1).M}^{\mu T} = [\beta_1^{\mu}, \ldots, \beta_M^{\mu}, \alpha_{11}^{\mu}, \ldots, \alpha_{1M}^{\mu}, \ldots, \alpha_{n1}^{\mu}, \ldots, \alpha_{nM}^{\mu}]$$

and the parameters ϕ^{ν} and θ^{ν} are as follows:

$$\begin{gathered} \phi^{\nu} = [R^{\nu T} \quad R^{\nu T} x_1 \quad \ldots \quad R^{\nu T} x_n] \\ R^{\nu} = \left[\frac{r_1^{\nu}}{\sum_{k=1}^{M} r_k^{\nu}} \quad \frac{r_2^{\nu}}{\sum_{k=1}^{M} r_k^{\nu}} \quad \cdots \quad \frac{r_M^{\nu}}{\sum_{k=1}^{M} r_k^{\nu}} \right] \\ \theta_{(n+1).M}^{\nu T} = [\beta_1^{\nu}, \ldots, \beta_M^{\nu}, \alpha_{11}^{\nu}, \ldots, \alpha_{1M}^{\nu}, \ldots, \alpha_{n1}^{\nu}, \ldots, \alpha_{nM}^{\nu}] \end{gathered} \tag{14}$$

3. Methodology

The parameters of the IT2AIFLS are estimated using the GSA and the R-LS algorithm. The GSA is used to estimate the antecedent part parameters of the IT2AIFLS which appears nonlinearly in the output of the system. The R-LS algorithm is used to estimate the consequent part parameters of the IT2AIFLS as they appear linearly in the output. Because the GSA is initiated from multiple start points, the probability of falling into a local minimum for this algorithm is less than computational approaches, such as gradient descent, that are initiated from a single point.

3.1 GSA optimization of Antecdent part parameters

The GSA is employed to optimize the antecedent part parameters which include the interval associated with sigma value $[\sigma_{1,ik} \quad \sigma_{2,ik}]$ and the crisp center (c_{ik}) of intutionist Gaussian functions in (4). Using the GSA, the antecedent part parameters are encoded in terms of the positions of particles in the GSA with their velocity vector and acceleration terms are updated using the GSA. Each solution in the GSA is defined as the particle positions in a d −dimensional search space representing the antecedent part parameters. The position vector associated with the GSA is as follows:

$$p^l = \left(\sigma_{1,11}^l, \ldots, \sigma_{1,ik}^l, \ldots, \sigma_{1,nM}^l, \sigma_{2,11}^l, \ldots, \sigma_{2,ik}^l, \ldots, \sigma_{2,nM}^l, c_{11}^l, \ldots, c_{ik}^l, \ldots, c_{nM}^l\right), l = 1, \ldots, N \tag{15}$$

where $\sigma_{1,ik}^l$, $\sigma_{2,ik}^l$ and c_{ik} refer to antecedent part parameters as appeared in (4) which correspond to the *k-th* rule for *i-th* input and *l* represents the *l-th* solution for the particle, $\sigma_{2,ik}^l$ refers to σ_2 value as appeared in (4) corresponding to the *k-th* rule for *i-th* input and *l* is the solution counter. The total number of antecedent part parameters to be estimated in this case is equal to $3 \times n \times M$. To evolve the antecedent part parameters according to the GSA, a mass value is assigned to each particle according to its merit, with a higher mass representing an antecedent part with a lower mean squared error and a lower mass representing a higher mean squared error. This makes the particles with worse performance move towards better particles. The mass of particles is updated and normalized at the t^{th} iteration as [5]:

$$M^l(t) = \frac{m^l(t)}{\sum_{l=1}^{N} m^l(t)} \tag{16}$$

where $m_i(t)$ is a non-normalized mass value corresponding to the i^{th} particle at iteration number t. The values of $m_i(t)$ represent the quality of a solution and are defined as [5]:

$$m^l(t) = \frac{f(\mathrm{p}^l) - f_{worst}(t)}{f_{best}(t) - f_{worst}(t)} \tag{17}$$

where $f(p^l)$ is the mean squared error value corresponding to p^l after estimating the consequent part parameters corresponding to p^l parameters using the R-LS algorithm. The R-LS is summarized in Section 5.3, $f_{worst}(t)$ and $f_{best}(t)$ are updated at every iteration as follows:

$$\begin{gathered} f_{worst}(t) = max\{f(p^l)\}_{l=1,\ldots,N} \\ f_{best}(t) = min\{f(p^l)\}_{l=1,\ldots,N} \end{gathered} \tag{18}$$

In each iteration, the acceleration term and velocity value must be calculated to update the position vector p^l. The position term in each step is updated as follows:

$$p^l(t+1) = p^l(t) + v^l(t+1) \tag{19}$$

where v^l represents the velocity vector and is updated using the acceleration term a^l as follows:

$$v^l(t+1) = r_i\ v^l(t) + a^l(t) \tag{20}$$

where $v^l(t) \in R^d$ represents the d −dimensional velocity of the particles at t^{th} iteration and $r_i \in [0,\ 1]$ is a uniform random number. The acceleration term, $a^l(t)$, is then updated as follows:

$$a^l(t) = \sum_{j\in\{1,\ldots,k_b\}} r_j G(t) \frac{M^l(t)(p^j(t) - p^l(t))}{\| p^l(t) - p^j(t) \|^{r_p} + \varepsilon} \tag{21}$$

where k_b represents the number of best solutions selected, $\|.\|$ stands for the Euclidean norm, ε is a small value added to prevent division by $zero$, r_p is the power considered for the Euclidean distance between two particles, $G(t)$ is the gravitational constant and $r_j \in [0,\ 1]$ is a uniform random value. The gravitational constant is then updated at each iteration using the following equation:

$$G(t) = G_0 exp\left(-\beta \frac{t}{t_{max}}\right) \tag{22}$$

where G_0 has a constant real value and t_{max} is the maximum value of algorithm iteration.

3.2 Ridge Least Square Estimation of Consequent Part Parameters

The consequent part parameters of the IT2AIFLS are trained using the R-LS algorithm also known as Tikhonov regularization. To be able to apply the R-LS algorithm to the consequent part parameters we need to write them in a matrix form as in (23). Then the R-LS algorithm of (27) can be applied to estimate the consequent part parameters. The R-LS algorithm, also known as Tikhonov regularization, is the solution to the following $\boldsymbol{l_2}$-norm cost function [22].

$$\min_{\theta\in R^n} \|\Phi\theta - Y\|_2^2 + \lambda\|\theta\|_2^2 \tag{24}$$

where $\|.\|_\mathbf{2}$ denotes the l_2-norm and λ is a regulation parameter. Furthermore, the parameters Φ, θ and Y are as follows:

$$\Phi = \begin{bmatrix} \phi_1^{\mu T} & \phi_1^{\nu T} \\ \phi_2^{\mu T} & \phi_2^{\nu T} \\ & \vdots & \\ \phi_S^{\mu T} & \phi_S^{\nu T} \end{bmatrix} \tag{25}$$

$$Y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_S \end{bmatrix} \tag{26}$$

$$\theta = \begin{bmatrix} \theta^{\mu T} & \theta^{\nu T} \end{bmatrix} \tag{27}$$

where ϕ^μ and ϕ^ν for each sample are defined in (13) and (14), S is the total number of samples, θ^μ and θ^ν are defined in (13) and (14), $y_s, s = 1,\ldots,S$ are the measured values. This solution to this problem prevents over-fitness and the analytical solution to the problem is obtained by

$$\theta = (\Phi^T\Phi + \lambda \mathrm{I})^{-1}\Phi^T Y \tag{28}$$

where **I** is the identity matrix. It is then necessary to select the parameter λ for the R-LS algorithm. Large values for $\boldsymbol{\lambda}$ need to be avoided to maintain the prediction accuracy of the system. An appropriate value for this parameter was found to be 0.01 in this paper through trial and error.

3.3. Performance measurement

Root mean squared error (RMSE), as the most common performance evaluation method, is used to illustrate the identification performance with its mathematical formula being as follows:

$$RMSE = \sqrt{\frac{1}{N}\sum_{t=1}^{N}(y_{IT2AIFLS}(t) - y_{actual}(t))^2} \tag{29}$$

where N is the number of test samples, $y_{actual}(t)$ is the actual value of the system target and $y_{IT2AIFLS}(t)$ is the output of the IT2AIFLS.

The flow chart of the proposed tuning method for the IT2AIFLS using GSA for the antecedent part parameters and the R-LS algorithm for the consequent part parameters is illustrated in Fig. 2. The first step is to create the random initial population, then the consequent part parameters are estimated using R-LS according to (28). The output of the IT2AIFLS can then be calculated using (2)-(14). The cost function associated with each member of the population is calculated using the RMSE formula (29). The GSA algorithm is then iterated for a single iteration to generate the next positions according to (16) – (22). If the stop conditions are not met, we need to iterate the algorithm until the stop condition is satisfied.

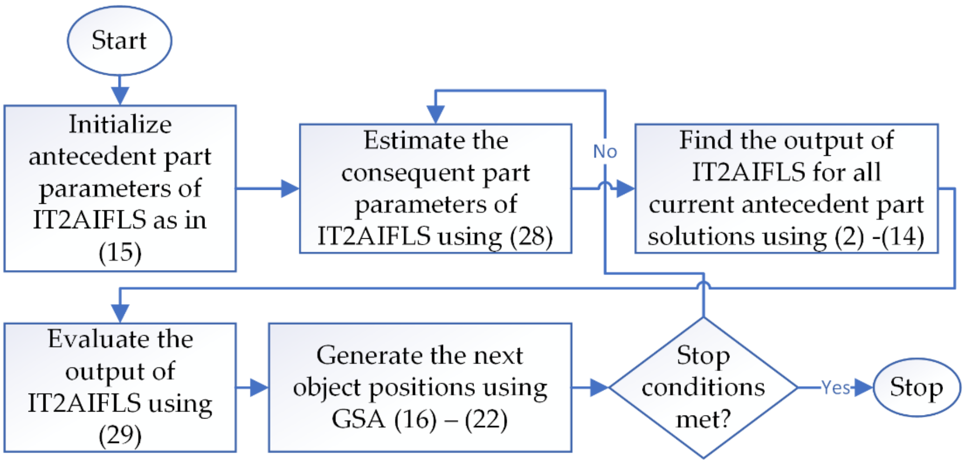


Figure 2. Flow chart of the proposed tuning method for the IT2AIFLS using the GSA for antecedent part parameters and the R-LS algorithm for the consequent part parameters

4. Simulation Results

4.1. Benchmark Identification Problem

Although the proposed prediction method is mainly designed for electricity load prediction, to test its efficacy, it is implemented on a benchmark second order nonlinear dynamic system with time-varying parameters [28]. This nonlinear dynamic system has been previously tested in several papers with other prediction methods to show their efficacy. The output of this dynamic system is a nonlinear time-varying function of inputs, with time delays of input and output as follows [28].

$$y(t+1) = f(y(t), y(t-1), y(t-2), u(t), u(t-1) \tag{30}$$

where the nonlinear function $f(.)$ is defined as follows:

$$f(x_1, x_2, x_3, x_4, x_5) = \frac{x_1 x_2 x_3 x_5 (x_3 - b) + c x_4}{a + x_2^2 + x_3^2} \tag{31}$$

and parameters a, b and c are time-varying parameters defined by:

$$a(t) = 1.2 - 0.2\cos\left(\frac{2\pi t}{T}\right) \tag{32}$$

$$b(t) = 1.0 - 0.4\sin\left(\frac{2\pi t}{T}\right)$$
$$c(t) = 1.0 + 0.4\sin\left(\frac{2\pi t}{T}\right)$$

with T, the total number of samples, taken to be equal to *1000*. The input signal to the system $u(t)$ is taken as follows:

$$u(t) = \begin{cases} \sin\left(\frac{\pi t}{25}\right) & t < 250 \\ 1.0 & 250 \le t < 500 \\ -1.0 & 500 \le t < 750 \\ f(t) & 750 \le t < 1000 \end{cases} \tag{33}$$

where:

$$f(t) = 0.3\sin\left(\frac{\pi t}{25}\right) + 0.1\sin\left(\frac{\pi t}{32}\right) + 0.6\sin\left(\frac{\pi t}{10}\right) \tag{34}$$

The first 80% of the generated data is used for training and the last 20% is chosen for testing purposes. The comparison results with several other methods including Type-1 TSK FNS [29], Type-2 TSK FNS [29], Feedorward Type-2 FNN [11], SIT2FNN [30], SEIT2 FNN [31], TSCIT2FNN [32], IT2 FNN-GD [28], IT2 FNN-SMC [28], IT2 FNNPSO+ SMC [28], IT2 IFLS -DEKF+GD [12], IT2FLS with Modified SVR [33] are presented in Table I. Where results support the idea that the proposed approach is effective at system identification by outperforming the other tested algorithms. The behaviour of the proposed identification system for the training and tesing data is presented in Figs 3 and 4, respectively. As these figures show, the error between the real data and the output of the IT2AIFLS is very low.

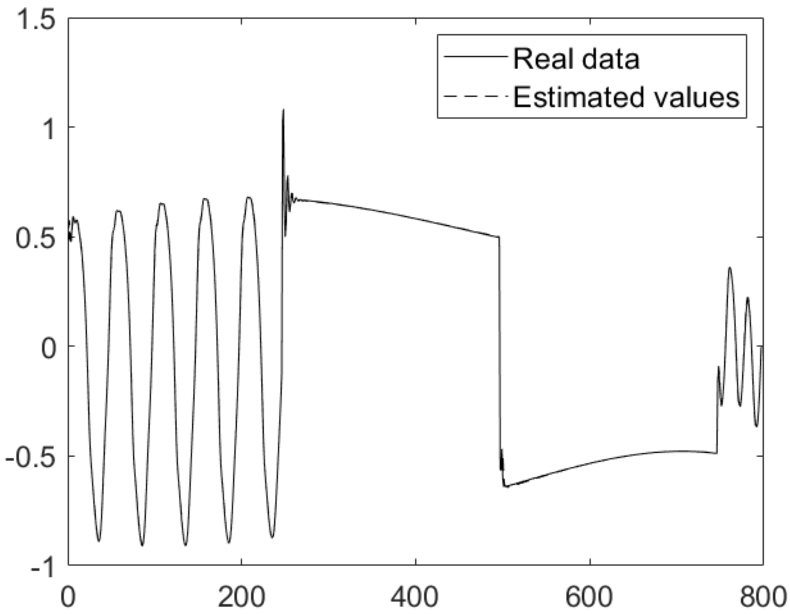


Figure 3 Performance of the proposed prediction method for the time-varying nonlinear system dataset (train set)

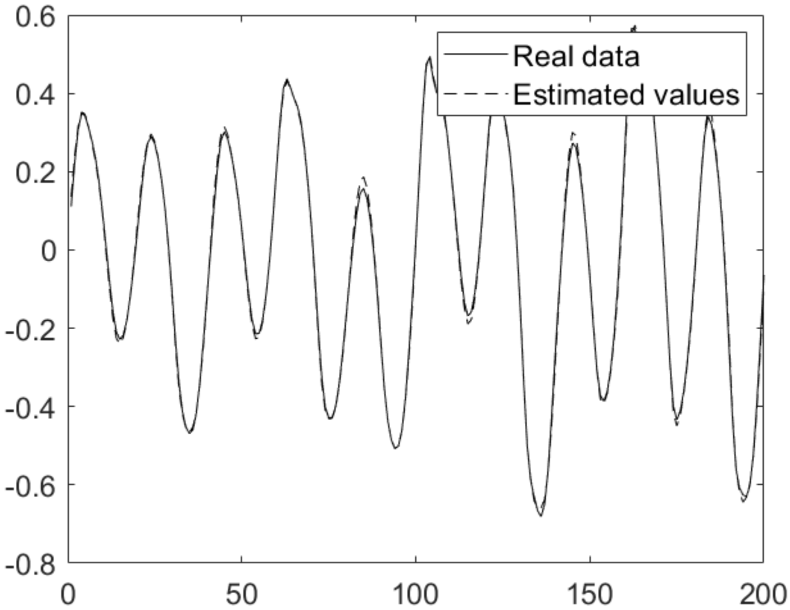


Figure 4 Performance of the proposed prediction method for the time-varying nonlinear system dataset (test set)

Table I. Performance of the Proposed Approach as well as Several Existing Methods in Literature on Time-Varying System Identification. Bold faced results indicate the best results.

	Rules	Epoch	Training RMSE	Testing RMSE
Type-1 TSK FNS [29]	9	100	0.0282	0.0598
Type-2 TSK FNS [29]	4	100	0.0284	0.0601
Feedorward Type-2 FNN [11]	3	100	0.0281	0.0593
SIT2FNN [30]	4	100	0.0351	0.0560
SEIT2 FNN [31]	3	100	0.0274	0.0574
TSCIT2FNN [32]	3	100	0.0279	0.0576
IT2 FNN-GD [28]	-	200	0.0540	0.0613
IT2 FNN-SMC [28]	-	200	0.0360	0.0390
IT2 FNNPSO+ SMC [28]	-	200	0.0199	0.0390
IT2 IFLS -DEKF+GD [12]	4	100	0.0250	0.0310
IT2FLS with Modified SVR [33]	11	Non-iterative	0.0146	0.0348
Proposed approach	**14**	**200**	**0.0095**	**0.0106**

Motivated by the fact that the proposed system identification method can successfully outperform several state-of-the-art system identification approaches in literature. This method was then used for electrical load prediction in Sections 4.2 and 4.3, where the real-time electrical load of Poland and five different regions in Australia are considered, respectively. Comparisons with state-of-the-art prediction models are then presented to illustrate the performance of the proposed prediction method.

4.2. Electrical Load Prediction for Poland

The dataset selected in this part shows the performance of the proposed approach to deal with electrical load prediction using the Poland electricity load dataset available online [34] which presents electricity load values of Poland in the 1990s on a daily basis [35]. A statistical analysis is conducted for the Poland dataset to examine if this time series is stationary or not. To evaluate this property of the Poland electrical load dataset, Augmented Dickey Fuller (ADF) and Kwiatkowski-Phillips-Schmidt-Shin (KPSS) tests are utilized. These two tests are available under the statsmodels Python package. The test statistic value under these two tests as well as their critical vales are given in Table II where it can be seen that the test value for ADF is greater than the critical value with a 5% confidence interval, and the test value for KPSS is less than the critical value with a 5% confidence interval. This means that this type of dataset is a trend stationary one. Although the stationary property of the signal is required in some statistical approaches, it is expected that the IT2AIFLS can handle this trend stationary signal due to its nonlinear nature.

Table II. Results of the ADF and KPSS tests for the Poland electrical load dataset

Test	Value	Critical value (5% confidence interval)
ADF-test	-2.55	-2.86
KPSS-test	0.34	0.46

As separated on the website, 1400 sample data are selected for training and 201 data samples are selected for testing. In this approach the one-step ahead prediction problem is investigated. The inputs taken for this prediction are the current value of electricity load and its time delays as $[y(t),\ y(t-1)\ y(t-2)\ y(t-3)]$. Here 14 rules (number of r_k^{μ}s and r_k^{ν}s) are considered for the fuzzy system and the number of rules was obtained by trial and error to maximize performance. The performance comparison between the proposed approach and previous approaches in literature are presented in Table III. As can be seen from the table, IT2AIFLS with GSA-R-LS outperforms IT2FLS DEFKF+GD, IFLS DEKF+GD, IT2 IFLS DEKF+GD and IT2 AIFLS DEKF+GD (previously studied in [11] on the same dataset) by at least 5.7%. The prediction performance of the proposed algorithm for the training and testing data are presented in Figs 5 and 6, respectively. As can be seen from these figures, the prediction output of the IT2AIFLS replicates the real data with high performance. The enlarged portion of the plot given in Fig 5 shows that the prediction output of the IT2AIFLS closely follows not only the low frequency part of the measured data but also the high frequency oscillations of the measured data present in the train dataset.

Table III. Performance comparison of the proposed hybrid algorithm for IT2AIFLS on Poland electrical load dataset in terms of RMSEs. The best results are highlighted in bold.

Model	*Train/Test*	*RMSE Train*	*RMSE Test*
IT2FLS DEKF+GD [11]	*1395/196*	*0.0564*	*0.0595*
IFLS DEKF+GD [11]	*1395/196*	*0.0589*	*0.0599*
IT2 IFLS DEKF+GD [11]	*1395/196*	*0.0560*	*0.0572*
IT2AIFLS GSA-R-LS *(proposed algorithm)*	*1395/196*	***0.0528***	***0.0501***

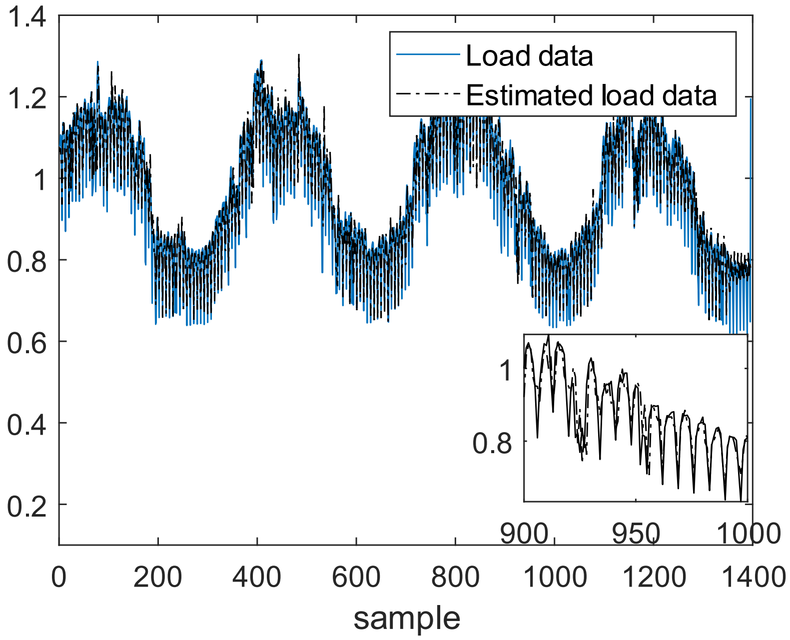


Figure 5 Performance of the proposed prediction method for the Poland electrical load dataset (train set)

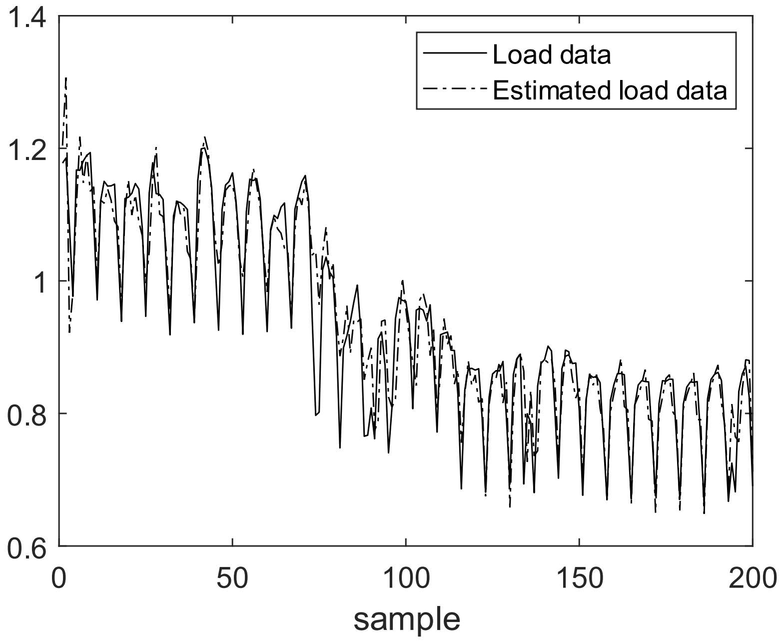


Figure 6 Performance of the proposed prediction method for the Poland electrical load dataset (test set)

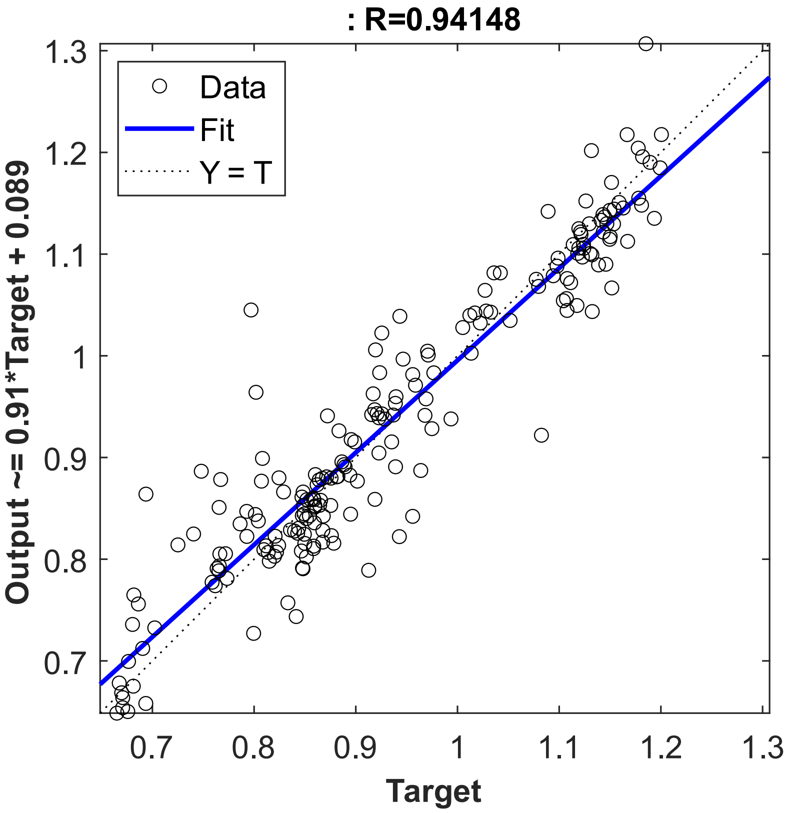


Figure 7 Regression analysis for IT2AIFLS on the Poland electrical load dataset for test data

Figure 7 shows the regression analysis for the IT2AIFLS on the Poland electrical load dataset for test data. It can be seen that the predicted value-target value graph is close to the ideal graph of (Y=T). Furthermore, the R value for this prediction is equal to 0.94 which is close to one. Hence, the prediction is performed with high quality.

4.3. Electrical Load Prediction for Five regions in Australia

In this Section, the proposed hybrid training method is used for the prediction of electrical load (MW) for five different regions in Australia namely: New South Wales (NSW), Queensland (QLD), South Australia (SA), Tasmania (TAS) and Victoria (VIC). The datasets used in this part are retrieved from the Australian Energy Market Operator (AEMO) website at http://www.aemo.com.au. Data available on this website is available at a 30 minute sample time basis. For comparison purposes, the time range of the test and train data are selected as the same as the previous study in [36]. This means that 1152 data samples from 2018/11 0:30 to 2018/ 2125 0:00 are used for training purposes. The data samples available from 201812/25 0:30 to 2018/2/270:00 are used for testing the prediction performance of the IT2AIFLS. The time delayed input values considered for the study are $[y(t),\ y(t-1)\ y(t-2)\ y(t-3)\ y(t-4)]$ and 14 rules are considered for the fuzzy system. The number of the rules were selected via trial and error.

Table IV. Statistical parameters of the electrical load data in five regions of Australia.

	Type of data	Number of samples	min	max	mean	std	skewness
NSW	Train	1152	5809.3	12846	8288	1306.7	0.35
	Test	96	5884.2	8563.3	7618.2	754.5	-0.67
QLD	Train	1152	5127.4	9480.1	6737.8	1020.5	0.73
	Test	96	5337.0	8910.9	6821.9	1060.0	0.22
SA	Train	1152	816.3	2798.0	1447.6	377.8	1.13
	Test	96	778.0	1479.8	1134.5	179.4	0.03
TAS	Train	1152	896.0	1302.1	1082.0	88.38	0.03
	Test	96	902.7	1294.1	1072.1	90.25	0.28
VIC	Train	1152	3601.6	9044.9	5049.3	927.04	1.15
	Test	96	3482.3	5945.2	4466.2	692.3	0.49

Table V the result of ADF-test and KPSS-test for electrical load prediction of five regions of Australia

Dataset	ADF-test-value	Critical value (5% confidence interval) for ADF	KPSS-test	Critical value (5% confidence interval) for KPSS	Result
NSW	-5.05	-2.86	0.45	0.46	Stationary
QLD	-6.07	-2.86	0.69	0.46	Difference stationary
SA	-3.99	-2.86	0.39	0.46	stationary
TAS	-5.95	-2.86	0.15	0.46	stationary
VIC	-4.46	-2.86	0.25	0.46	stationary

Table VI. Performance criteria in terms of RMSEs for electrical load prediction data in five regions of Australia (test data). The best results are highlighted in bold and MPI represents the minimum percentage improvement.

	SVR[36]	ANN [36]	ELM [36]	EEMD-ELM_GOA [36]	EEMD-ELM-DA[36]	EEMD-ELM-PSO [36]	EEMD-ELM-GWO [36]	Proposed approach	MPI
NSW	2351	1468	1651	603	684	839	766	**89**	**85%**
QLD	905	717	705	564	336	941	558	**76**	**77%**
SA	552	380	381	154	155	261	207	**41**	**73%**
TAS	168	192	162	61	81	257	139	**20**	**67%**
VIC	1323	975	961	419	499	877	975	**103**	**75%**

Table IV presents the statistical parameters associated with data samples which show that this dataset has high levels of variations with a large standard deviation. Furthermore, there exist some cases in Table IV for which the Skewness value is larger than 0.5 which means that data is moderately right skewed in these cases. Example moderately right-skewed cases are the train data for QLD, the train data for SA, and the train and test data for VIC. Furthermore, the skewness for the test data in the case of the NSW region is -0.69 which is less than -0.5. This means that test data in the NSW region is moderately left-skewed. In the remaining cases, the absolute value of skewness belongs to the interval of [-0.5 0.5] which means that those datasets are fairly symmetrical. Table V presents the results of the ADF and KPSS tests for the electrical load datasets of five regions of Australia. As Table V shows, the datasets associated with the electrical load for the NWS, SA, TAS and VIC regions of Australia are stationary, however, the dataset associated with the QLD region is difference stationary. Usually, it is easier to deal with stationary signals. Despite this, it is expected that the IT2AIFLS can handle difference stationary signals, where the difference of the signal is stationary, due to its nonlinear nature.

The comparison results of the proposed training algorithm for the IT2AIFLS with seven different approaches using the RMSE performance index on the test data are presented in Table VI. As Table VI shows, the IT2AIFLS with the proposed training method outperforms seven other previously studied algorithms in [36] namely: SVR, ANN, ELM, EEMD-ELM-GOA, EEMD-PSO-ELM, and EEMD-ELM-GWO. The percentage improvement over these algorithms are significant and is at least 67% of improvement. Fig 8 presents plots of the prediction output of the proposed algorithm against the measured data for the electrical load data in each of the five regions of Australia. These plots show that the prediction output of the proposed algorithm closely follows the measured data which indicated that the performance of the proposed algorithm is satisfactory.

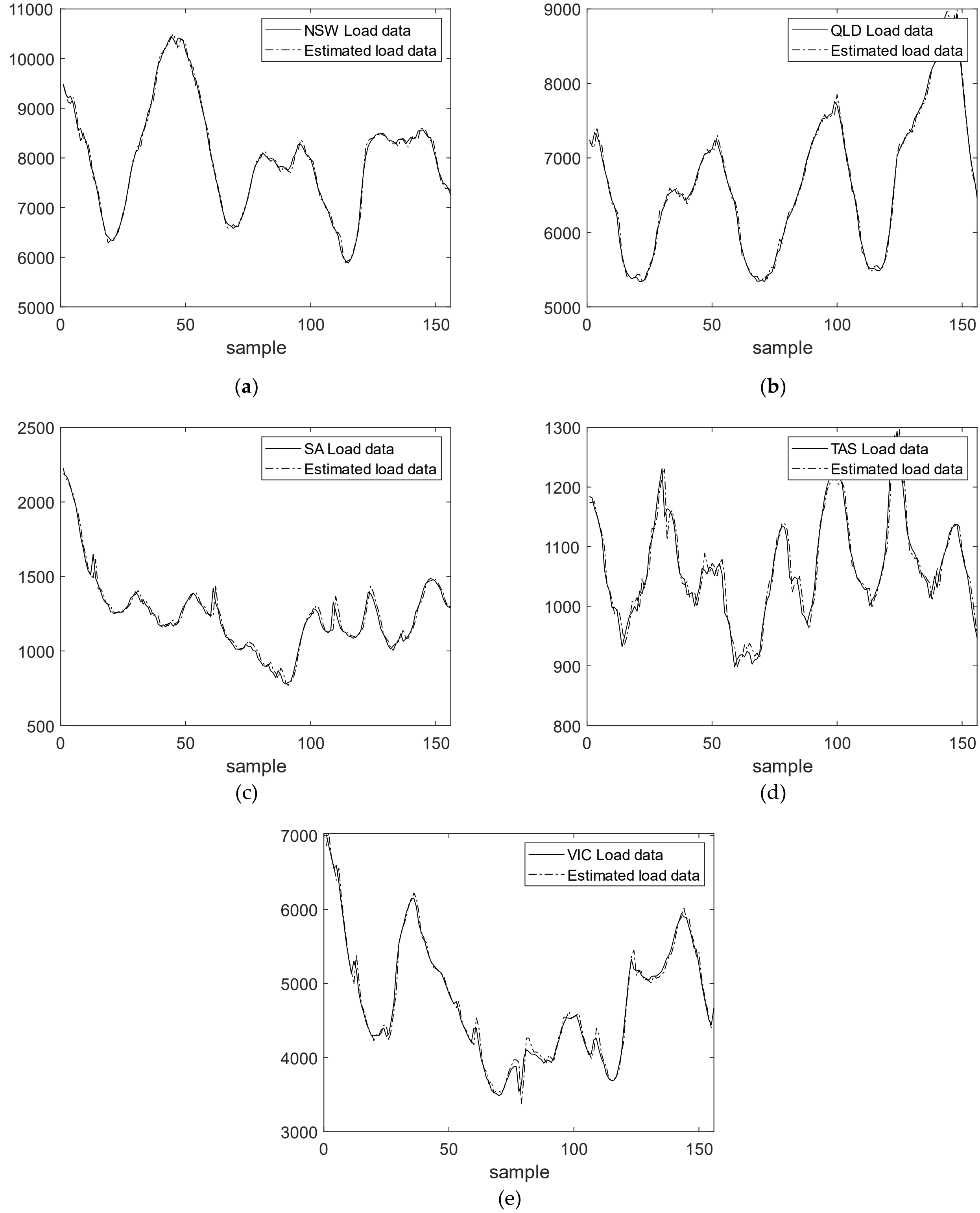
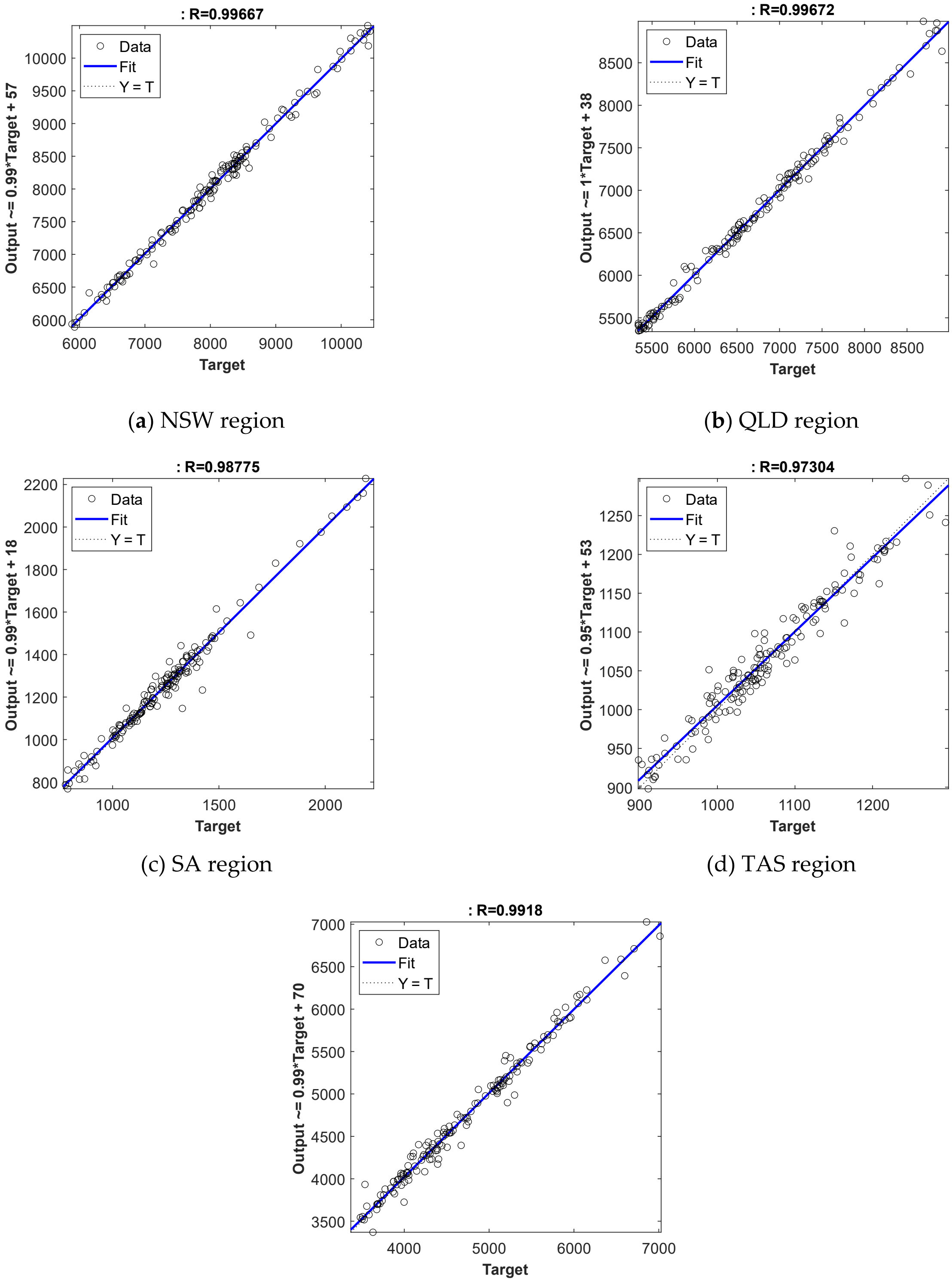


Figure 8 Estimation performance of the proposed prediction method for five regions of Australia electrical load dataset (train set): a) NSW b) QLD c) SA d) TAS e) VIC



(**a**) NSW region (**b**) QLD region

(c) SA region (d) TAS region

(e) VIC region

Figure 9 Regression analysis of the proposed prediction method for five regions of Australia electrical load dataset (train set): a) NSW b) QLD c) SA d) TAS e) VIC

To provide greater insight on the obtained results, regression analysis is performed on the results. Fig. 9 provides regression analysis for all 5 regions of Australia for the test data. As Fig. 9 shows, the R value in all cases is very close to 1 and the bias term is reasonable. This indicates that the electrical load forecasting for 5 regions of Australia is performed with high quality. Moreover, although the electrical load dataset of the QLD region is a difference stationary signal, the R value obtained for it is 0.9967 which is very close to one and supports the idea that an IT2AIFLS, because of its non-linear nature, can handle this class of data as well.

5. Conclusions

In this paper, a IT2AIFLS is trained using a hybrid method containing the GSA for the antecedent part parameters and the R-LS algorithm for consequent part parameters. A benchmark system identification problem is studied to investigate the efficacy of the proposed system parameter tuning approach on previous identification benchmark problems. The comparisons with several other methods including Type-1 TSK FNS, Type-2 TSK FNS, Feedorward Type-2 FNN, SIT2FNN, SEIT2 FNN, TSCIT2FNN, IT2 FNN-GD, IT2 FNN-SMC, IT2 FNNPSO+ SMC, IT2 IFLS -DEKF+GD, and IT2FLS with Modified SVR support the idea that the proposed approach is an efficient approach in system identification problems. The proposed approach is then investigated on electrical load prediction for five regions of Australia and Poland in the presence of noise and uncertainty which inherently exist in these datasets. Statistical properties of these datasets are presented that show they can be either stationary, trend stationary or difference stationary. In the case of the Poland electrical load dataset, the inputs to the IT2AIFLS are considered to be current values as well as three consecutive time delays of data. In the case of the five regions in Australia, current data values as well as four consecutive lags are used as the inputs to the fuzzy logic system. In both cases, one-step ahead prediction is considered. For the Poland dataset, the obtained prediction results are compared with several other algorithms including IT2FLS DEKF+GD, IFLS DEKF+GD and IT2IFLS DEKF+GD. The comparisons made in this paper show that the proposed algorithm results in superior performance when it is compared to these methods. For the case of the five regions of Australia, it is observed that the proposed prediction can perform with a much higher performance as compared with SVR, ANN, ELM, EEMD-ELM-GOA, EEMD-ELM-DA, EEMD-ELM-PSO and EEMD-ELM-GWO methods of prediction. Hence, the proposed approach is an automated solution to develop the antecedent part parameters of an IT2AIFLS, which successfully outperforms current state-of-the-art approaches in literature when applied to an electrical load prediction problem.

A possible direction of future work is the development of an interpretable IT2AIFLS which would lead to IF-THEN rules to describe the overall behavior of the electrical load system. Such IF-THEN rules are easy to communicate and may help experts to make more efficient decisions using their knowledge and experience. An additional direction of future work would be to apply the model to other time series prediction or system identification problems to test its applicability to enhance prediction in other industries.

Supplementary Materials: No supplementary materials.

Author Contributions: For research articles with several authors, a short paragraph specifying their individual contributions must be provided. The following statements should be used "Conceptualization, M. A. Khanesar and Jingyui Lu; methodology, M. A. Khanesar; software, M. A. Khanesar; validation, M. A. Khanesar; formal analysis, M. A. Khanesar; investigation, M. A. Khanesar; resources, M. A. Khanesar and Jingyui Lu; data curation, M. A. Khanesar; writing—original draft preparation, M. A. Khanesar, Thomas Smith, Jingyui Lu and David Branson; writing—review and editing, M. A. Khanesar, Thomas Smith, Jingyui Lu and David Branson; visualization, M. A. Khanesar; supervision, David Branson; project administration, David Branson; funding acquisition, David Branson. All authors have read and agreed to the published version of the manuscript." Please turn to the CRediT taxonomy for the term explanation. Authorship must be limited to those who have contributed substantially to the work reported.

Funding: This work is funded and supported by the Engineering and Physical Sciences Research Council (EPSRC) under grant number: EP/R021031/1—New Industrial Systems: Chatty Factories.

Data Availability Statement: The electrical load dataset of Poland is available from https://research.cs.aalto.fi/aml/datasets.shtml and the electrical load datasets for five regions of Australia are downloaded from http://www.aemo.com.au.

Acknowledgments: This work is funded and supported by the Engineering and Physical Sciences Research Council (EPSRC) under grant number: EP/R021031/1—New Industrial Systems: Chatty Factories.

Conflicts of Interest: Declare conflicts of interest or state "The authors declare no conflict of interest."

References

[1] S. Hassan, A. Khosravi, J. Jaafar, and M. A. Khanesar, "A systematic design of interval type-2 fuzzy logic system using extreme learning machine for electricity load demand forecasting," *Int. J. Electr. Power Energy Syst.*, vol. 82, 2016, doi: 10.1016/j.ijepes.2016.03.001.

[2] D. Wu, B. Wang, D. Precup, and B. Boulet, "Multiple Kernel Learning-Based Transfer Regression for Electric Load Forecasting," *IEEE Trans. Smart Grid*, vol. 11, no. 2, pp. 1183–1192, 2020, doi: 10.1109/TSG.2019.2933413.

[3] D. Bunn and E. D. Farmer, "Comparative models for electrical load forecasting," 1985.

[4] G. Dutta and K. Mitra, "A literature review on dynamic pricing of electricity," *J. Oper. Res. Soc.*, vol. 68, no. 10, pp. 1131–1145, 2017, doi: 10.1057/s41274-016-0149-4.

[5] J. Ma, H. Jiang, K. Huang, Z. Bi, and K. L. Man, "Novel Field-Support Vector Regression-Based Soft Sensor for Accurate Estimation of Solar Irradiance," *IEEE Trans. Circuits Syst. I Regul. Pap.*, vol. 64, no. 12, pp. 3183–3191, 2017, doi: 10.1109/TCSI.2017.2746091.

[6] K. Ma, M. Soltani, A. Hajizadeh, J. Zhu, and Z. Chen, "Wind Farm Power Optimization and Fault Ride-Through under Inter-Turn Short-Circuit Fault," *Energies*, pp. 1–16, 2021.

[7] K. Ma, J. Zhu, M. Soltani, A. Hajizadeh, and Z. Chen, "Optimal power dispatch of an offshorewind farm under generator fault," *Appl. Sci.*, vol. 9, no. 6, 2019, doi: 10.3390/app9061184.

[8] Y. Jiang, S. Yin, J. Dong, and O. Kaynak, "A Review on Soft Sensors for Monitoring, Control and Optimization of Industrial Processes," *IEEE Sens. J.*, vol. 21, no. 11, pp. 1–1, 2020, doi: 10.1109/jsen.2020.3033153.

[9] J. Contreras, R. Espínola, F. J. Nogales, and A. J. Conejo, "ARIMA models to predict next-day electricity prices," *IEEE Trans. Power Syst.*, vol. 18, no. 3, pp. 1014–1020, 2003, doi: 10.1109/TPWRS.2002.804943.

[10] W. Li and Z. G. Zhang, "Based on time sequence of ARIMA model in the application of short-term electricity load forecasting," *ICRCCS 2009 - 2009 Int. Conf. Res. Challenges Comput. Sci.*, pp. 11–14, 2009, doi: 10.1109/ICRCCS.2009.12.

[11] I. Eyoh, R. John, G. De Maere, and E. Kayacan, "Hybrid Learning for Interval Type-2 Intuitionistic Fuzzy Logic Systems as Applied to Identification and Prediction Problems," *IEEE Trans. Fuzzy Syst.*, vol. 26, no. 5, pp. 2672–2685, 2018, doi: 10.1109/TFUZZ.2018.2803751.

[12] I. Eyoh, R. John, G. De Maere, and E. Kayacan, "Hybrid Learning for Interval Type-2 Intuitionistic Fuzzy Logic Systems as applied to Identification and Prediction Problems," *IEEE Trans. Fuzzy Syst.*, p. 1, 2018, doi: 10.1109/TFUZZ.2018.2803751.

[13] S. Member and R. John, "Interval Type-2 A-Intuitionistic Fuzzy Logic for Regression Problems," vol. 26, no. 4, pp. 2396–2408, 2018.

[14] I. Eyoh, R. John, and G. De Maere, "Interval Type-2 Intuitionistic Fuzzy Logic Systems-A Comparative Evaluation," in *International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems*, 2018, pp. 687–698.

[15] 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 Syst.*, vol. 160, no. 7, pp. 922–948, 2009.

[16] S. Hassan, M. A. Khanesar, J. Jaafar, and A. Khosravi, "Comparative analysis of three approaches of antecedent part

generation for an IT2 TSK FLS," *Appl. Soft Comput. J.*, vol. 51, 2017, doi: 10.1016/j.asoc.2016.11.015.

[17] F. Olivas, F. Valdez, P. Melin, A. Sombra, and O. Castillo, "Interval type-2 fuzzy logic for dynamic parameter adaptation in a modified gravitational search algorithm," *Inf. Sci. (Ny).*, vol. 476, pp. 159–175, 2019, doi: 10.1016/j.ins.2018.10.025.

[18] M. A. Khanesar and D. Branson, "XOR binary gravitational search algorithm," *Conf. Proc. - IEEE Int. Conf. Syst. Man Cybern.*, vol. 2019-Octob, pp. 3269–3274, 2019, doi: 10.1109/SMC.2019.8914607.

[19] E. Rashedi, H. Nezamabadi-Pour, and S. Saryazdi, "GSA: a gravitational search algorithm," *Inf. Sci. (Ny).*, vol. 179, no. 13, pp. 2232–2248, 2009.

[20] S. Duman, U. Güvenç, Y. Sönmez, and N. Yörükeren, "Optimal power flow using gravitational search algorithm," *Energy Convers. Manag.*, vol. 59, pp. 86–95, 2012, doi: 10.1016/j.enconman.2012.02.024.

[21] C. Purcaru, R. E. Precup, D. Iercan, L. O. Fedorovici, R. C. David, and F. Dragan, "Optimal robot path planning using gravitational search algorithm," *Int. J. Artif. Intell.*, vol. 10, no. 13 S, pp. 1–20, 2013.

[22] L. S. Lls, T. Long, W. Jiao, G. He, and L. Ax, "RPC Estimation via 1 -Norm-Regularized," vol. 53, no. 8, pp. 4554–4567, 2015.

[23] M. M. Khalaf, S. O. Alharbi, and W. Chammam, "Similarity measures between temporal complex intuitionistic fuzzy sets and application in pattern recognition and medical diagnosis," *Discret. Dyn. Nat. Soc.*, vol. 2019, 2019, doi: 10.1155/2019/3246439.

[24] K. T. Atanassov, *On intuitionistic fuzzy sets theory*, vol. 283. Springer, 2012.

[25] O. Castillo and K. Atanassov, *Comments on fuzzy sets, interval type-2 fuzzy sets, general type-2 fuzzy sets and intuitionistic fuzzy sets*, vol. 372. Springer International Publishing, 2019.

[26] I. Eyoh, R. John, and G. De Maere, "Interval Type-2 A-Intuitionistic Fuzzy Logic for Regression Problems," *IEEE Trans. Fuzzy Syst.*, vol. 26, no. 4, pp. 2396–2408, 2018.

[27] I. Eyoh, J. Eyoh, and R. Kalawsky, "Interval Type-2 Intuitionistic Fuzzy Logic System for Time Series and Identification Problems - A Comparative Study," *Int. J. Fuzzy Log. Syst.*, vol. 10, no. 1, pp. 1–17, 2020, doi: 10.5121/ijfls.2020.10101.

[28] E. Kayacan and M. A. Khanesar, *Fuzzy Neural Networks for Real Time Control Applications: Concepts, Modeling and Algorithms for Fast Learning*. 2015.

[29] R. H. Abiyev and O. Kaynak, "Type 2 fuzzy neural structure for identification and control of time-varying plants," *IEEE Trans. Ind. Electron.*, vol. 57, no. 12, pp. 4147–4159, 2010, doi: 10.1109/TIE.2010.2043036.

[30] Y. Y. Lin, S. H. Liao, J. Y. Chang, and C. T. Lin, "Simplified interval type-2 fuzzy neural networks," *IEEE Trans. Neural Networks Learn. Syst.*, vol. 25, no. 5, pp. 959–969, 2014, doi: 10.1109/TNNLS.2013.2284603.

[31] C.-F. Juang and Y.-W. Tsao, "A Self-Evolving Interval Type-2 Fuzzy Neural Network With Online Structure and Parameter Learning," *Fuzzy Syst. IEEE Trans.*, vol. 16, no. 6, pp. 1411–1424, Dec. 2008.

[32] 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 Trans. Ind. Electron.*, vol. 61, no. 1, pp. 447–459, 2014, doi: 10.1109/TIE.2013.2248332.

[33] M. A. Khanesar and D. T. Branson, "Prediction Interval Identification Using Interval Type-2 Fuzzy Logic Systems: Lake Water Level Prediction Using Remote Sensing Data," *IEEE Sens. J.*, vol. XX, no. XX, pp. 1–13, 2021, doi: 10.1109/JSEN.2021.3067841.

[34] "Applications of Machine Learning Group," *https://research.cs.aalto.fi/aml/datasets.shtml*, 2019. .

[35] A. Lendasse, M. Cottrell, V. Wertz, and M. Verleysen, "Prediction of electric load using Kohonen maps - Application to the polish electricity consumption," *Proc. Am. Control Conf.*, vol. 5, pp. 3684–3689, 2002, doi: 10.1109/ACC.2002.1024500.

[36] J. Wu, Z. Cui, Y. Chen, D. Kong, and Y. G. Wang, "A new hybrid model to predict the electrical load in five states of Australia," *Energy*, vol. 166, pp. 598–609, 2019, doi: 10.1016/j.energy.2018.10.076.