Support Vector Regression Based Multi-Objective Parameter Estimation of Interval Type-2 Fuzzy Systems

Mojtaba Ahmadieh Khanesar[1[0000-0001-5583-7295]] , David Branson[2[0000-0001-5818-666X]]

[1,2] Faculty of Engineering, Mechanical, Materials and Manufacturing Department, University of Nottingham, NG7 2RD Nottingham, UK
{[1]ezzma5,[2]ezzdtb}@exmail.nottingham.ac.uk

Abstract. This paper presents a support vector regression based multi-objective parameter estimation method for interval type-2 fuzzy systems, which deals with prediction interval rather than its crisp output value. Such a prediction interval covers future values of data which is quite useful in some tasks. A narrower yet inclusive prediction interval is more desirable and contains more information, as it avoids conservative lower and upper limits for data. Earlier support vector regression based estimation approaches for the parameters of interval type-2 fuzzy systems do not have control over this width and instead focus on prediction accuracy. In this work, to control such a prediction interval a multi-objective cost function is introduced that other than a term corresponding to prediction accuracy, includes a weighted term corresponding to width of prediction interval. The weight used for the width of prediction interval provides a tradeoff between prediction accuracy and width of prediction interval. The cost function is formulated in terms of a constrained quadratic objective function problem which can be solved using well established quadratic programming approaches. The proposed method is successfully applied to the prediction of the chaotic Mackey-Glass time series, where it can be observed that the proposed method is capable of controling prediction interval through appropriate selection of weighting parameter. For instance, the prediction of the chaotic Mackey-Glass time series is done with probable 70% decrease in sum of absolute value of prediction interval with respect to existing support vector regression estimation algorithm while maintaining the prediction accuracy. This is the main benefit of the current approach over previous approaches in literature.

Keywords: Identification, interval type-2 fuzzy logic systems, quadratic programming, regression model

1 Introduction

Since not all experts in a field agree on the same membership grades for a given input to the fuzzy systems, interval type-2 fuzzy systems (IT2FLSs) are introduced which benefit from fuzzy membership functions [1]. Other than an ability to cover different meaning of words for experts, using such fuzzy membership functions, IT2FLSs are capable of dealing with uncertainty and noise in a much more effective way than their type-1 counterparts [2]. The fact that uncertainty and noise exist in almost all real-time applications gives rise to widespread use of IT2FLSs in such applications [1].

Although, IT2FLSs benefit from interval valued membership functions and consequent part parameters, the final output of a fuzzy system after defuzzification and type-reduction is a crisp value. However, interval valued outputs for an identifier are also beneficial and gives more information about future values. Such prediction lowers the risks of decision made in various applications [3].

Various training methods are used to estimate the parameters of an IT2FLS. The interval fuzzy logic models investigated in [4] and [5] benefit from type-1 fuzzy logic systems with its parameters being estimated using a l_∞ approach. Recursive least square method has been applied to estimate the parameters of two independent type-1 fuzzy logic systems acting as the lower and upper limits for data [6]. These approaches are used in cases when data values to be identified are interval.

As the consequent part parameters of IT2FLSs appear linearly in the output, computational methods such as least squares are more appreciated as they are capable of tuning these parameters in two stages without iteration and does not require any design parameters [7]. Support vector regression (SVR) is an alternative non-iterative training method derived from support vector machine and successfully applied to IT2FLSs resulting in high generalisation properties [8].

In this paper, a novel training method for the estimation of the parameters of IT2FLS is introduced. In the proposed estimation algorithm, the output of IT2FLS is represented in the form of an interval. Such an interval may be useful to describe processes such as temperature, stock price, skill transfer of various human motions to robot and etc. The mentioned covering interval needs to be inclusive of the data and be as narrow as possible. To obtain such an interval, an appropriate cost function is proposed and the parameters of IT2FLS are estimated to minimize such cost function in two stages similar to the approaches studied in [8] and [9]. Another similarity of the proposed approach to [8] and [9] is that the proposed algorithm does not require any iterations and does not necessitate the choise of any design parameters. However, the superiority of the proposed approach over [8], [9] and [10], is that it provides means to control the width of uncertainty. As it is mentioned earlier, it is highly desired for the width of prediction interval to be as narrow as possible. Considering the width of prediction interval as well as prediction accuracy, parameter estimation for IT2FLSs is a multiobjective optimization problem. Corresponding cost function is designed, in which an appropriate gain is given to width of prediction interval. Quadratic programming algorithm is then used to minimize this cost function. The proposed algorithm is then tested against chaotic Mackey-Glass time series, where it is shown that the proposed algorithm can predict an appropriate interval for such time-series. It is further observed that using the weight in the cost function, makes it possible to obtain a narrower prediction interval while maintaining prediction accuracy. The proposed approach is compared to number of previous approaches in literature which shows superior performance for the proposed approach.

This paper is organized as: Section 2 provides overview of the basic structure of IT2FLS. The proposed methodology for the training of IT2FLSs are presented in

Section III. The experimental results of the proposed IT2FLSs are illustrated in Section IV. Finally, in Section V the concluding marks are presented.

2 General structure of Interval Type-2 Fuzzy Logic System

Several structure for IT2FLSs and its type-reducers are investigated [11] [12]. The structure used in this paper benefits from interval type-2 fuzzy membership functions in the antecedent part and interval values for the consequent part parameters. A typical fuzzy IF-THEN rule for such a structure is

$IF\ x_1\ is\ \tilde{A}_{j1}\ and\ x_2\ is\ \tilde{A}_{j2}\ and\ \dots and\ x_n\ is\ \tilde{A}_{jn}$

$$THEN \quad y_j = \sum_{i=1}^{n} \tilde{\alpha}_{ij} x_i + \tilde{\beta}_j \tag{1}$$

where x_1, x_2, ... ,x_n are the input variables, y is the single output variable. Moreover $\tilde{A}_{ij}$'s are interval type-2 fuzzy membership functions for j^{th} rule of the i^{th} input. $\tilde{\alpha}_{ij}$ and $\tilde{\beta}_j$ $(i = 1, \dots, n, j = 1, \dots, M)$ are the interval parameters in the consequent part of the rules which satisfy the following equation.

$$\tilde{\alpha}_{ij} \in [\underline{\alpha}_{ij},\ \overline{\alpha}_{ij}],\ \tilde{\beta}_j \in [\underline{\beta}_j,\ \overline{\beta}_j] \tag{2}$$

The following definitions are made.

$$\underline{F}_j = \sum_{i=1}^{n} \underline{\alpha}_{ij} x_i + \underline{\beta}_j \tag{3}$$

$$\overline{F}_j = \sum_{i=1}^{n} \overline{\alpha}_{ij} x_i + \overline{\beta}_j \tag{4}$$

$$\underline{w}^j(x) = \underline{\mu}_{\tilde{F}_1^j}(x_1) * \dots * \underline{\mu}_{\tilde{F}_n^j}(x_n) \tag{5}$$

$$\overline{w}^j(x) = \overline{\mu}_{\tilde{F}_1^j}(x_1) * \dots * \overline{\mu}_{\tilde{F}_n^j}(x_n) \tag{6}$$

where $\underline{\mu}_{\tilde{F}_k^j}(x_k)$ are $\overline{\mu}_{\tilde{F}_k^j}(x_k)$ are the lower and upper membership function corresponding to j^{th} rule for x_k and " $*$ " is a t-norm operator. The output value of IT2FLS is given as

$$Y(x) = [y_l(x),\ y_r(x)] \tag{7}$$

$$= \int_{w^1 \in [\underline{w}^1, \overline{w}^1]} \cdots \int_{w^M \in [\underline{w}^M, \overline{w}^M]} 1 \Big/ \frac{\sum_{j=1}^{M} w^j y_j}{\sum_{j=1}^{M} w^j} \tag{8}$$

where $x \in R^n$ is a vector of inputs of system. There exist various defuzzifiers+type reducers to calculate $Y(x)$ which is an interval type-1 set represented by its right and left bounds among which the Maclauren based first order approximate output of the fuzzy system is chosen [12]. It is shown that the accuracy of this method is less than accurate model of Enhanced Karnik-Mendel model [13] and higher than Biglarbegian-Melek-Mendel [14] and Nie-Tan models [15]. However, this algorithm is faster than

Enhanced Karnik-Mendel model and it does not require the sorting procedure as is required by Enhanced Karnik-Mendel model [13]. The Maclauren based first order approximate output of the fuzzy system is as [12]:

$$y \in [y_r,\ y_l] \tag{9}$$

where y_l and y_r are being the lower and upper bounds of the output of type-2 fuzzy system which are calculated as:

$$y_r \approx \frac{\sum_{j=1}^{M}(\overline{w}^j+\underline{w}^j)\overline{F}^j+\sum_{j=1}^{M}(sign(\overline{m}^j)\Delta w^j\overline{F}^j)}{\sum_{j=1}^{M}(\overline{w}^j+\underline{w}^j)+\sum_{j=1}^{M}(sign(\overline{m}^j)\Delta w^j)} \tag{10}$$

where:

$$\overline{m}^j = \overline{F}^j - \frac{\sum_{j=1}^{M}\overline{w}^j\overline{F}^j}{\sum_{j=1}^{M}\overline{w}^j} \tag{11}$$

and $\Delta w^j = \overline{w}^j - \underline{w}^j$. Furthermore, y_l is calculated as

$$y_l \approx \frac{\sum_{j=1}^{M}(\overline{w}^j+\underline{w}^j)\underline{F}^j-\sum_{j=1}^{M}(sign(\underline{m}^j)\Delta w^j\underline{F}_j)}{\sum_{j=1}^{M}(\overline{w}^j+\underline{w}^j)-\sum_{j=1}^{M}(sign(\underline{m}^j)\Delta w^j)} \tag{12}$$

where:

$$\underline{m}^j = \underline{F}^j - \frac{\sum_{j=1}^{M}\underline{w}^j\underline{F}^j}{\sum_{j=1}^{M}\underline{w}^j} \tag{13}$$

The final crisp output value of IT2FLS is obtained as

$$Y(x) = \frac{y_l+y_r}{2} \tag{14}$$

It is then possible to rewrite (8) as

$$y_r = \sum_{j=1}^{M} v_R^j\overline{F}_R^j \tag{15}$$

where:

$$v_R^j = \frac{\overline{w}^j+\underline{w}^j+sign(\overline{m}^j)\Delta w^j}{\sum_{j=1}^{M}(\overline{w}^j+\underline{w}^j)+\sum_{j=1}^{M}(sign(\overline{m}^j)\Delta w^j)} \tag{16}$$

The parameter y_r in a vector form is obtained as:

$$y_r = \phi_R\overline{\theta} \tag{17}$$

where

$$\phi_R = [\vec{v}_R^T, \vec{v}_R^T x_1, \ldots, \vec{v}_R^T x_n]^T \tag{18}$$

and $\vec{\alpha}_R$ is defined as

$$\vec{v}_R = [v_R^1, \dots, v_R^M]^T \tag{19}$$

Furthermore, $\overline{\theta}$ is defined as

$$\overline{\theta}^T_{(n+1).M} = [\overline{\beta}_1, \dots, \overline{\beta}_M, \overline{\alpha}_{11}, \dots, \overline{\alpha}_{1M}, \dots, \overline{\alpha}_{n1}, \dots, \overline{\alpha}_{nM}]$$

Similarly, it is possible to rewire the equation corresponding to y_l (10) in a vector form as

$$y_l = \sum_{j=1}^{M} v_l^j \overline{F}_l^j \tag{20}$$

where:

$$v_l^j = \frac{(\overline{w}^j + \underline{w}^j) - (sign(\underline{m}^j)\Delta w^j)}{\sum_{j=1}^{M}(\overline{w}^j + \underline{w}^j) - \sum_{j=1}^{M}(sign(\underline{m}^j)\Delta w^j)} \tag{21}$$

The parameter y_l in a vector form is obtained as.

$$y_l = \phi_L \underline{\theta} \tag{22}$$

where

$$\phi_L = [\vec{v}_L^T, \vec{v}_L^T x_1, \dots, \vec{v}_L^T x_n]^T \tag{23}$$

and $\vec{\alpha}_L$ is defined as.

$$\vec{v}_L = [v_L^1, \dots, v_L^M]^T \tag{24}$$

Furthermore, $\underline{\theta}$ is defined as.

$$\underline{\theta}^T_{(n+1).M} = [\underline{\beta}_1, \dots, \underline{\beta}_M, \alpha_{11}, \dots, \underline{\alpha}_{1M}, \dots, \underline{\alpha}_{n1}, \dots, \underline{\alpha}_{nM}]$$

3 Support Vector Regression Model

Let the training set for the fuzzy system be $\{(x_1, t_1), \dots, (x_N, t_N)\}$, with $t_k, k = 1, \dots, N$ being the target values for the fuzzy system. Support vector regression model guarantees that training error is kept less than ε [9]. In this case, the system is penalized if the absolute value of error is larger than ε and is not penalized otherwise. The dead zone cost function can describe the penalty function assigned for the training of fuzzy system.

$$D\varepsilon_k = \begin{cases} 0 & if \ |e_k| < \varepsilon \\ |e_k| - \varepsilon & otherwise \end{cases}, \ k = 1, \dots, N \tag{23}$$

where e_k is identification error. To solve this problem, the following constrained optimization problem is defined.

$$\min_{\underline{\theta},\overline{\theta},\xi,\xi^*} \frac{1}{2}\theta_l^T\theta_l + \frac{1}{2}\theta_r^T\theta_r + C\sum_{k=1}^{N}(\xi_k + \xi_k^*) \tag{25}$$

$$s.t.\quad t_k - \frac{1}{2}(\phi_{L,k}\underline{\theta} + \phi_{R,k}\overline{\theta}) \le \varepsilon + \xi_k,\ \ k = 1,\dots,N \tag{26}$$

$$\frac{1}{2}(\phi_{L,k}\underline{\theta} + \phi_{R,k}\overline{\theta}) - t_k \le \varepsilon + \xi_k^*,\ \ k = 1,\dots,N \tag{27}$$

$$\xi_k, \xi_k^* \ge 0\ \ \forall k \tag{28}$$

While absolute values of errors less than ε are tolerated, for an absolute value of error larger than ε, the positive slack parameters ξ_k and ξ_k^* penalize the cost function. The parameter C forms a trade off between the complexity of model and its accuracy. Figure 2 illustrates this concept.

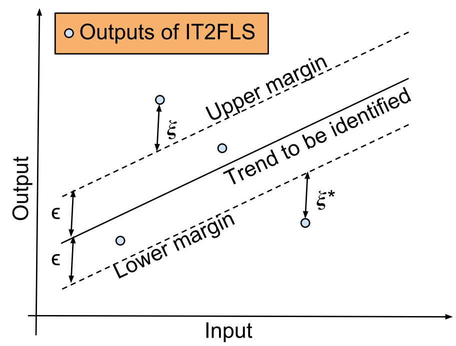

Fig. 1. Dead zone cost function for a linear SVM.

This cost function and its solution for IT2FLSs have been considered widely in literature [8] [9]. Although such approach results in high performance identifier, there is no control over the width of uncertainty $(y_r - y_l)$. This width of uncertainty bounds data and gives valuable information about it. Narrower width of uncertainty contains more information about data and can be used more effectively in decision making applications. Motivated by this fact a modified cost function for support vector regression model is proposed.

4 Proposed Multi-objective Support Vector Regression Based Parameter Estimation Method

As it is mentioned earlier current support vector regression based parameter estimation methods do not have any control over the width of interval of output $(y_r - y_l)$. Therefore, to control such parameter, it is required to include it in the cost function. The proposed cost function in this case is as.

$$\min_{\underline{\theta},\overline{\theta},\xi,\xi^*} \frac{1}{2}\underline{\theta}^T\underline{\theta} + \frac{1}{2}\overline{\theta}^T\overline{\theta} + C\sum_{k=1}^{N} (\xi_k + \xi_k^* + \xi_k^- + \xi_k^+) + \underbrace{\gamma\sum_{k=1}^{N} (t_k - \phi_{l,k}\underline{\theta})^2}_{I_1}$$

$$+\underbrace{\gamma \sum_{k=1}^{N} (t_k - \phi_{r,k}\overline{\theta})^2}_{I_2} \tag{29}$$

$$s.t. \quad t_k - \tfrac{1}{2}(\phi_{L,k}\underline{\theta} + \phi_{R,k}\overline{\theta}) \le \varepsilon + \xi_k,\ k = 1,\dots,N \tag{30}$$

$$\tfrac{1}{2}(\phi_{L,k}\underline{\theta} + \phi_{R,k}\overline{\theta}) - t_k \le \varepsilon + \xi_k^*,\ k = 1,\dots,N \tag{31}$$

$$\phi_{R,k}\overline{\theta} - t_k \le \varepsilon + \xi_k^+,\ k = 1,\dots,N \tag{32}$$

$$t_k - \phi_{L,k}\underline{\theta} \le \varepsilon + \xi_k^-,\ k = 1,\dots,N \tag{33}$$

$$\xi_k, \xi_k^*, \xi_k^+, \xi_k^- \ge 0\ \forall k \tag{34}$$

In comparison to existing SVR algorithm presented in (25)-(28), in the proposed algorithm, the cost function is changed and two more non-equalties of (32) and (33) are added. In the proposed cost function while the condition (33) including ξ^- guarantees that y_l never becomes greater than target value, the condition (32) including ξ^+ guarantees that y_r never falls below target value. Furthermore, the terms I_1 and I_2 in (29) provide means for making the width of interval $[y_l,\ y_r]$ as small as desired by pushing y_r and y_l towards t_k. The parameter γ is responsible for penalizing the prediction interval which makes the interval as narrow as desired.

As the regressor values $\phi_{L,k}$ and $\phi_{R,k}$ depend on values of consequent part parameters, it is required to do the optimization twice. First since the consequent part parameters are unknown, the regressor values ϕ_L and ϕ_R are chosen as.

$$\phi_R = [\vec{v}_R^T, \vec{v}_R^T x_1, \dots, \vec{v}_R^T x_n]^T \tag{35}$$

and $\vec{v}_R$ is defined as.

$$\vec{v}_R = [v_R^1, \dots, v_R^M]^T \tag{36}$$

where:

$$v_R^j = \frac{\overline{w}^j}{\sum_{j=1}^{M} \overline{w}^j} \tag{37}$$

and

$$\phi_L = [\vec{v}_L^T, \vec{v}_L^T x_1, \dots, \vec{v}_L^T x_n]^T \tag{38}$$

and $\vec{v}_L$ is defined as.

$$\vec{v}_L = [v_L^1, \dots, v_L^M]^T \tag{39}$$

where:

$$v_L^j = \frac{\underline{w}^j}{\sum_{j=1}^{M} \underline{w}^j} \tag{40}$$

On the second step, based on the initial conditions found for $\underline{\theta}$ and $\overline{\theta}$, the regressor values of IT2FLS are calculated based on equations (5)-(22) developed in Section II. The pseudo code of the proposed algorithm is as.

1. Input Selection and data processing
2. Present data in terms of maximum and minimum
3. Split data to test and train dataset
4. MF Generation for IT2FLS
5. Obtain regressors using updated consequent part parameters using (35)- (40)
6. Tune the consequent part parameters (first stage) to obtain the regressors using quadratic programming approach to solve optimization problem (29)-(34)
7. Obtain regressors using updated consequent part parameters using (16)- (19) and (20)- (24)
8. Tune the consequent part parameters (second stage) using quadratic programming approach to solve (29)-(34)
9. Evaluate the performance for train and test data. If error is satisfactory STOP, otherwise GOTO 4).

5 Simulation Results

To investigate the efficacy of the proposed algorithm and its ability to control the prediction interval width, it is applied to the prediction of Mackey-glass chaotic system. Accuracy performance of the proposed method is compared with existing methods in literature.

5.1 Chaotic Time-Series Prediction

Mackey-Glass time series is a well-known time-series that models the blood cell regulation and has a dynamic behavior. The delayed differential equation dynamic of this chaotic system is described as [16].

$$\dot{x}(t) = \alpha \frac{x(t-\tau)}{1+x^{10}(t-\tau)} - \beta x(t) \tag{41}$$

where $\alpha = 0.2$, $\beta = 0.1$ and $\tau = 17$. To implement the system in discrete time, first order Euler method is used with its time sample T_s being equal to $T_s = 1s$. The aim is to predict six-steps ahead $x(t+6)$ using current and past data of $x(t)$, $x(t-6)$, $x(t-12)$ and $x(t-18)$. The dataset consists of 1000 data sample with 80% being considered for training and 20% being used for test data. The membership functions used in the paper are Gaussian type-2 fuzzy membership functions with crisp centre and interval σ values.

The pseudo-code of the algorithm is as of Section 4. Table 1. presents the comparison results of the proposed algorithm with some of approaches in literature. As can be seen from this table, the proposed approach is capable of obtaining superior performance than some previous neural network [17] and neuro-fuzzy [8], [18]- [19] based approaches in literature. Furthermore, to investigate the effect of the parameter γ, simulations are done with various γ values. As can be seen from Fig. 2, the width of prediction interval decreases as the parameter γ increases from *zero* to *one* while the prediction accuracy remains close. Width of prediction interval (WPE) is calculated as:

$$WPE = \sum_{i=1}^{N}(y_r - y_l) \tag{42}$$

Table 1. Comparison results between the proposed approach and other results in literature for prediction of chaotic time series Mackey-Glass system

Method	RMSE of error
FALCON-ART [18]	0.04
GA-Ensemble [20]	0.026
SONFIN [21]	0.018
SVR fuzzy [10]	0.013
WNN-HLAs [17]	0.006
ANFIS [22]	0.007
SVD [23]	0.012
SA-T1FLS [19]	0.016
SA-T2FLS [19]	0.009
TSK-SVR I [8]	0.008
TSK-SVR II [8]	0.007
Proposed approach ($\varepsilon = 0.01$, $C = 40$ and $\gamma = 0.1$)	**0.007**
Proposed approach ($\varepsilon = 0.01$, $C = 40$ and $\gamma = 1$)	**0.006**

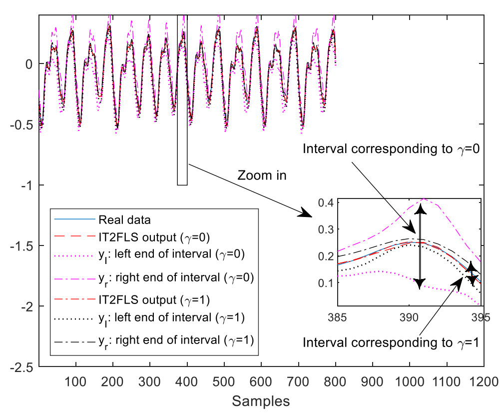

Fig. 2. The influence of γ on width of prediction interval

The index term WPE for the case when $\gamma = 0$ is obtained as equal to 208.43 while it is equal to 52.34 when $\gamma = 1$. This shows over 74% improvement in the results.

6 Conclusions and Future Works

6.1 Conclusions

In this paper, a novel approach for parameter estimation of IT2FLSs based on a multi-objective cost function that includes a term penalizing the width of prediction interval and another term corresponding to prediction accuracy is proposed. Quadratic programming method is used to minimize the cost function and estimate the parameters of IT2FLS. The proposed approach is tested on the prediction of chaotic Mackey-Glass system. It is shown through simulation that with appropriate choice of this parameter it is possible to obtain a narrower yet covering prediction interval while maintaining the prediction accuracy. The comparison between the prediction accuracy of the proposed method with other approaches in literature also shows superior performance.

6.2 Future Work

In future, the proposed method will be implemented on a robot being used to learn human motion when accomplishing simple manufacturing tasks such as sealant path production and assembly operations. In this work, the motion is recorded from several demonstrations made by one or more humans. As all experts do not agree on the same way to perform a task, and they may even change their preferred method over time. The movements result in a histogram. It is highly appreciated to obtain the narrowest possible prediction interval for human motion while obtaining highly accurate crisp output value for IT2FLS.

7 Acknowledgement

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.

8 References

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

[2] O. Castillo, P. Melin, J. Kacprzyk, and W. Pedrycz, "Type-2 fuzzy logic: theory and applications," in *Granular Computing, 2007. GRC 2007. IEEE International Conference on*, 2007, p. 145.

[3] M. H. F. Zarandi, B. Rezaee, I. B. Turksen, and E. Neshat, "A type-2 fuzzy

rule-based expert system model for stock price analysis," *Expert Syst. Appl.*, vol. 36, no. 1, pp. 139–154, 2009.

[4] I. ŠKrjanc, S. Blažič, and O. Agamennoni, "Identification of dynamical systems with a robust interval fuzzy model," *Automatica*, vol. 41, no. 2, pp. 327–332, 2005.

[5] I. Skrjanc, S. Blazic, and O. Agamennoni, "Interval Fuzzy Model Identification Using$ l_\infty$-Norm," *IEEE Trans. Fuzzy Syst.*, vol. 13, no. 5, pp. 561–568, Oct. 2005.

[6] M. A. Khanesar, M. Teshnehlab, and O. Kaynak, "Identification of interval fuzzy models using recursive least square method," in *Conference Proceedings - IEEE International Conference on Systems, Man and Cybernetics*, 2010.

[7] Z. Deng, K.-S. Choi, L. Cao, and S. Wang, "T2FELA: Type-2 Fuzzy Extreme Learning Algorithm for Fast Training of Interval Type-2 TSK Fuzzy Logic System," *Neural Networks Learn. Syst. IEEE Trans.*, vol. 25, no. 4, pp. 664–676, Apr. 2014.

[8] V. Uslan, H. Seker, and R. John, "A support vector-based interval type-2 fuzzy system," in *2014 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*, 2014, pp. 2396–2401.

[9] 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 Trans. fuzzy Syst.*, vol. 18, no. 4, pp. 686–699, 2010.

[10] J.-H. Chiang and P.-Y. Hao, "Support vector learning mechanism for fuzzy rule-based modeling: a new approach," *IEEE Trans. Fuzzy Syst.*, vol. 12, no. 1, pp. 1–12, 2004.

[11] A. Khosravi and S. Nahavandi, "Effects of type reduction algorithms on forecasting accuracy of IT2FLS models," *Appl. Soft Comput.*, vol. 17, pp. 32–38, 2014.

[12] M. A. Khanesar and J. M. Mendel, "Maclaurin series expansion complexity-reduced center of sets type-reduction + defuzzification for interval type-2 fuzzy systems," in *2016 IEEE International Conference on Fuzzy Systems, FUZZ-IEEE 2016*, 2016.

[13] D. Wu and J. M. Mendel, "Enhanced karnik--mendel algorithms," *Fuzzy Syst. IEEE Trans.*, vol. 17, no. 4, pp. 923–934, 2009.

[14] M. Biglarbegian *et al.*, "On the Stability of Interval Type-2 TSK Fuzzy Logic Control Systems," *IEEE Trans. Syst. Man, Cybern. Part B*, vol. 40, no. 3, pp. 798–818, 2010.

[15] M. Nie and W. W. Tan, "Towards an efficient type-reduction method for interval type-2 fuzzy logic systems," in *2008 IEEE International Conference on Fuzzy Systems (IEEE World Congress on Computational Intelligence)*, 2008, pp. 1425–1432.

[16] M. C. Mackey and L. Glass, "Oscillation and chaos in physiological control systems," *Science (80-.).*, vol. 197, no. 4300, pp. 287–289, 1977.

[17] C.-J. Lin, "Wavelet neural networks with a hybrid learning approach," *J. Inf. Sci. Eng.*, vol. 22, no. 6, pp. 1367–1387, 2006.

[18] C.-J. Lin and C.-T. Lin, "An ART-based fuzzy adaptive learning control

network," *IEEE Trans. Fuzzy Syst.*, vol. 5, no. 4, pp. 477–496, 1997.

[19] M. Almaraashi, R. John, S. Coupland, and A. Hopgood, "Time series forecasting using a TSK fuzzy system tuned with simulated annealing," in *International Conference on Fuzzy Systems*, 2010, pp. 1–6.

[20] D. Kim and C. Kim, "Forecasting time series with genetic fuzzy predictor ensemble," *IEEE Trans. Fuzzy Syst.*, vol. 5, no. 4, pp. 523–535, 1997.

[21] C.-F. Juang and C.-T. Lin, "An online self-constructing neural fuzzy inference network and its applications," *IEEE Trans. Fuzzy Syst.*, vol. 6, no. 1, pp. 12–32, 1998.

[22] J.-S. Jang, "ANFIS: adaptive-network-based fuzzy inference system," *IEEE Trans. Syst. Man. Cybern.*, vol. 23, no. 3, pp. 665–685, 1993.

[23] H. Gu and H. Wang, "Fuzzy prediction of chaotic time series based on singular value decomposition," *Appl. Math. Comput.*, vol. 185, no. 2, pp. 1171–1185, 2007.