ICIC Express Letters ICIC International ©2016 ISSN 1881-803X
Volume **10**, Number **7**, July **2016** pp. 1649–1655

PREDICTION OF INTERVAL-VALUED TIME SERIES BASED ON BERNSTEIN FUZZY SYSTEM

Wenyan Song[1,2], Degang Wang[3], Pan Fang[3] and Dehai Liu[2]

[1]School of Mathematics
[2]Center for Econometric Analysis and Forecasting
Dongbei University of Finance and Economics
No. 217, Jianshan Street, Shahekou District, Dalian 116025, P. R. China

[3]School of Control Science and Engineering
Dalian University of Technology
No. 2, Linggong Road, Ganjingzi District, Dalian 116024, P. R. China
wangdg@dlut.edu.cn

Received December 2015; accepted March 2016

Abstract. *In this paper, Bernstein fuzzy system with polynomial consequents is established for predicting interval-valued time series. Prediction of interval-valued time series can be transferred into forecasting the mid-value and half range value of interval values. Accordingly, parameters identification of Bernstein fuzzy system is investigated. Fuzzy C-means method is used to determine the center of membership function. The width of membership function is optimized by genetic algorithm. And, partial least square algorithm is applied to determine the parameters of generalized Bernstein polynomials. Some numerical simulations show that Bernstein fuzzy system can forecast the interval-valued time series with high accuracy.*
Keywords: Bernstein fuzzy system, Interval-valued time series, Prediction

1. **Introduction.** As is well known, time series is an effective tool to describe dynamic system. In the classical time series theory, the state at each time usually takes the real number. However, in some application areas, such as stock market, temperature model and production process, some variables need to be represented as interval-valued data. Accordingly, how to use interval-valued data to establish and analyze the mathematical model of the objective system becomes a hot topic.

Some authors establish various neural networks to handle the interval-valued data. In [1], the forecasting model combined with autoregressive model and artificial neural network is proposed for analyzing the interval-valued time series. In [2], a complex-valued radial basis function neural network is applied to predict the interval-valued time series. In [3], a type-2 fuzzy neural network is given to forecast the dynamic model. In [4], the model of multilayer perceptron is established to handle interval-valued data. In [5], interval-valued fuzzy set is utilized to design recurrent neural network for predicting dynamic model. In addition, some authors attempt to apply various fuzzy modeling methods to process interval-valued data. In [6], when state variable weights of rules are taken as interval value, the approximation properties of the corresponding fuzzy system are investigated. In [7], fuzzy reasoning method which can handle the interval-valued data is established. In [8], an interval-valued intuitionistic fuzzy weighted algorithm is applied for the medical diagnosis. In [9], fuzzy correlation based on interval-valued data is investigated. In order to analyze the approximation properties of fuzzy system to nonlinear function, in [10], the identification algorithm of interval fuzzy model is given.

From the existing results, we find that many fuzzy systems and fuzzy neural networks are obtained based on Mamdani-type model and TS-type model. The consequent of each fuzzy rule is usually represented by linear model. However, for many complex systems, linear model could not fully reflect the nonlinear feature of the objective system in each local region. Hence, in this paper, we will establish a novel fuzzy system with Bernstein polynomial consequents to investigate interval-valued time series. The merit of the proposed model is that in each local region, Bernstein polynomial is used to model the objective system which can effectively describe the nonlinear feature. Since Bernstein polynomials have been widely used in function approximation, fuzzy system generalized by Bernstein polynomials will enrich the structure of fuzzy model and improve the forecasting accuracy for interval-valued time series.

This paper is organized as follows. In Section 2, the representation of Bernstein fuzzy system and some preliminaries are introduced. In Section 3, the identification method for the Bernstein fuzzy system is investigated. In Section 4, some numerical examples are provided to illustrate the proposed method. In Section 5, some main conclusions are summarized.

2. Bernstein Fuzzy System.

In this section, we will introduce the mathematical expression of Bernstein fuzzy system and some notations which are used in this paper. Denote $C[a, b]$ as the space of continuous function on $[a, b]$. Firstly, we introduce two definitions.

Definition 2.1 ([11]). *Suppose that $f \in C[0, 1]$, then the generalized Bernstein polynomials of degree n $(n > 1)$ is defined as*

$$\left(\bar{B}_n f\right)(x) = \sum_{k=0}^{n} \tfrac{n!}{k!(n-k)!} x^k (1-x)^{n-k} f(x_k)$$

where $0 = x_0 < x_1 < \cdots < x_n = 1$.

Different from traditional Bernstein polynomials, the partition point x_k may not be equidistant partition. Then, we give the definition of multi-dimensional generalized Bernstein polynomials.

Definition 2.2. *Suppose that $f \in C([0, 1] \times \cdots \times [0, 1])$, then the multi-dimensional generalized Bernstein polynomial is defined as*

$$\left(\bar{B}_{m_1,\cdots,m_n} f\right)(x) = \sum_{k_1=0}^{m_1} \cdots \sum_{k_n=0}^{m_n} \left(\prod_{j=1}^{n} \begin{pmatrix} m_j \\ k_j \end{pmatrix} \cdot x_j^{k_j} \cdot (1-x_j)^{m_j-k_j} \right) \cdot f(x_{1k_1}, \cdots, x_{nk_n}) \quad (1)$$

where $x = (x_1, \cdots, x_n)$, x_{ik_i} $(k_i = 1, \ldots, m_i;\ i = 1, \ldots, n)$ are the partition points in $[0, 1]$.

Further, we will give the mathematical expression of Bernstein fuzzy system. Fuzzy rules of Bernstein fuzzy system can be represented as follows.

$$\text{If } x_1 \text{ is } A_{i1} \text{ and } \cdots \text{ and } x_n \text{ is } A_{in}, \text{ then } y = H_i(x_1, \cdots, x_n) \quad (2)$$

where $i = 1, \ldots, N$ and

$$H_i(x_1, \cdots, x_n) = \sum_{k_1=0}^{m_{i1}} \cdots \sum_{k_n=0}^{m_{in}} \left(\prod_{j=1}^{n} \begin{pmatrix} m_{ij} \\ k_j \end{pmatrix} x_j^{k_j} (1-x_j)^{m_{ij}-k_j} \right) \cdot y^{(i)}_{k_1\cdots k_n}.$$

Based on fuzzy rule (2), Bernstein fuzzy system determined by singleton fuzzifier, product inference engine and centroid defuzzifier can be expressed as

$$y = \sum_{i=1}^{N} A_i(x) \cdot H_i(x) \quad (3)$$

where $A_i(x) = \prod_{j=1}^{n} A_{ij}(x_j)$ and $A_{ij}(x_j) = \exp\left(-(x_j - c_{ij})^2 / \sigma_{ij}^2\right)$.

It can be seen that in each local region the proposed model is represented as generalized Bernstein polynomials which means that the weighted sum of some generalized Bernstein polynomials is established to approximate the objective model.

3. **Identification of Bernstein Fuzzy System under the Interval-valued Data.** In this section, we will propose the identification algorithm to determine parameters of the Bernstein fuzzy system and use the proposed model to forecast interval-valued time series.

An interval-valued time series is denoted by $\{[X_t](t = 1, 2, \ldots)\}$. X_t is taken as interval value, i.e., $X_t = [X_t^L, X_t^R]$, where X_t^L and X_t^R respectively denote the lower bound and upper bound of the time series at time t. The interval-valued time series can be divided into two time series models with crisp values, i.e., $\{X_t^1, t = 1, 2, \ldots\}$ and $\{X_t^2, t = 1, 2, \ldots\}$, where $X_t^1 = \frac{X_t^L + X_t^R}{2}$ is the mid-value of the interval and $X_t^2 = \frac{X_t^R - X_t^L}{2}$ is the half range of the interval. Accordingly, prediction of interval-valued time series can be transferred into prediction of above two time series models. Without loss of generality, we take $\{X_t^1, t = 1, 2, \ldots\}$ as an example to introduce how to design Bernstein fuzzy system to predict it. A group of training samples $\left\{\left(x_{t-(n-1)}^1, \ldots, x_t^1, x_{t+1}^1\right), t = n, n+1, \ldots\right\}$ could be generated by time series $\{X_t^1, t = 1, 2, \ldots\}$, where n is the order, $\left(x_{t-(n-1)}^1, \ldots, x_t^1\right)$ is the input value and x_{t+1}^1 is the output value.

Based on above samples, we introduce the identification method for the Bernstein fuzzy system. Parameters to be identified include $c_i = (c_{i1}, \cdots, c_{in})$, $\sigma_i = (\sigma_{i1}, \cdots, \sigma_{in})$ and $y_{k_1 \cdots k_n}^{(i)}$.

Firstly, for above samples, fuzzy C-means (FCM) method is used to determine the center c_i $(i = 1, \ldots, N)$ of input variables.

Then, a type of nonlinear iterative partial least squares (NIPALS) algorithm [12,13] is applied to calculate the weights $y_{k_1 \cdots k_n}^{(i)}$ in (3). $\prod_{j=1}^n m_{ij}$ is set to be the upper bound number of components for $A_i(x) \cdot \prod_{j=1}^n \begin{pmatrix} m_{ij} \\ k_j \end{pmatrix} x_j^{k_j} (1 - x_j)^{m_{ij} - k_j}$ $(k_j = 0, \ldots, m_{ij},\ j = 1, \ldots, n,\ i = 1, \ldots, N)$. Then by progressively trying to find relations between latent variables of $A_i(x) \cdot \prod_{j=1}^{n} \begin{pmatrix} m_{ij} \\ k_j \end{pmatrix} x_j^{k_j} (1 - x_j)^{m_{ij} - k_j}$ and the output y, the coefficients $y_{k_1 \cdots k_n}^{(i)}$ can be obtained.

Further, genetic algorithm (GA) is proposed to tune the width σ_i $(i = 1, \ldots, N)$. The fitness value is the output of Bernstein fuzzy system. A production with r chromosomes is randomly generated, denoted by x_{ij} $(i = 1, \ldots, r;\ j = 1, \ldots, N)$, where $x_{ij} = \left(x_{ij}^{(1)}, \cdots, x_{ij}^{(n)}\right)$ denotes the width of fuzzy sets in (3). By (3), we can compute the fitness value of each chromosome. By selection, crossover and mutation new population can be updated. In this way, we can obtain the width σ_i $(i = 1, \ldots, N)$. In the following, we summarize how to design Bernstein fuzzy system to predict interval-valued time series.

Step 1. Divide interval-valued time series $\{[X_t](t = 1, 2, \ldots)\}$ into two models: mid-value time series $\{X_t^1, t = 1, 2, \ldots\}$ and half range value time series $\{X_t^2, t = 1, 2, \ldots\}$.

Step 2. Initialize some parameters, including the order of time series, the number of clustering centers, the number of population and the maximum number of iteration.

Step 3. Use fuzzy C-means clustering method to determine the center of clusters, denoted by c_i $(i = 1, \ldots, N)$. Let c_i $(i = 1, \ldots, N)$ be the center value of fuzzy set $A_i(x)$.

Step 4. Randomly generate the width σ_i $(i = 1, \ldots, N)$ of fuzzy set $A_i(x)$.

Step 5. Use NIPALS algorithm to compute parameters $y_{k_1 \cdots k_n}^{(i)}$ in Equation (3).

Step 6. Carry out selection, crossover and mutation to produce new production and update the width σ_i $(i = 1, \ldots, N)$.

Step 7. If the number of iteration is less than the maximum number, then go to step 5; else stop.

4. **Numerical Examples.** In this section, some numerical simulations are provided to illustrate the validity of the proposed method for the interval-valued time series. In the simulation, some notations are defined as follows. The orders of time series and the generalized Bernstein fuzzy system are denoted by n and m_i $(i = 1, \ldots, n)$. The numbers of fuzzy rules and the population size are denoted by N and r. The generation of GA is M. The duplication rate, crossover rate and mutation rate are denoted by P_d, P_c and P_m.

Example 4.1. *Two synthetic interval-valued time series models are given below:*

(I): $X_t^1 = 1.5X_{t-1}^1 \cdot \exp\left(-\left(X_{t-2}^1/2\right)^2\right) + \varepsilon_t,\ \varepsilon_t \sim N(0, 0.5),\ X_t^2 \sim U[0, 2]$

(II): $X_t^1 = 4X_{t-1}^1\left(1 - X_{t-1}^1\right),\ X_t^2 \sim U[1, 2].$

We will respectively design two Bernstein fuzzy systems to predict above interval-valued time series models. In the simulation, 300 observation data with 200 training samples and 100 testing samples are obtained. For model (I), parameters of Bernstein fuzzy system are chosen as $n = 2$, $m_1 = 3$, $m_2 = 3$, $N = 15$, $M = 100$, $r = 30$, $P_d = 0.67$, $P_c = 0.33$, $P_m = 0.05$. For model (II), parameters of Bernstein fuzzy system are chosen as $n = 1$, $m_1 = 3$, $N = 7$, $M = 100$, $r = 30$, $P_d = 0.67$, $P_c = 0.33$, $P_m = 0.05$.

All simulations are performed 20 trials. The simulation results about testing samples are shown in Figures 1-8, where vertical line segment denotes the interval-valued data on each time, dotted lines denote the actual lower bound series and upper bound series and solid lines denote the corresponding forecasted series. From these simulation results, we can know that Bernstein fuzzy systems can effectively model the lower bound and upper bound of the objective system. These facts illustrate the validity of Bernstein fuzzy system in forecasting the synthetic interval-valued time series.

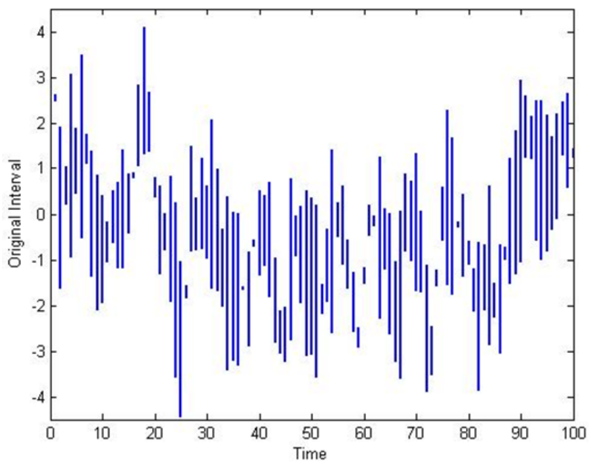

FIGURE 1. The actual value of model (I)

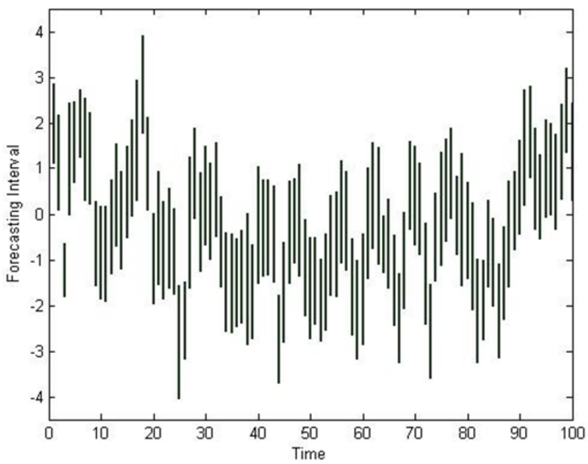

FIGURE 2. The forecasted value of model (I)

Example 4.2. *Experimental results on New York Stock Exchange.*[1]

We will use Bernstein fuzzy system to predict ten stock prices from the New York Stock Exchange. The lowest and highest trading prices of the day are used to generate intervalvalued time series. The basic information of these samples is shown in Table 1.

[1]New York Stock Exchange is available at http://finance.yahoo.com/.

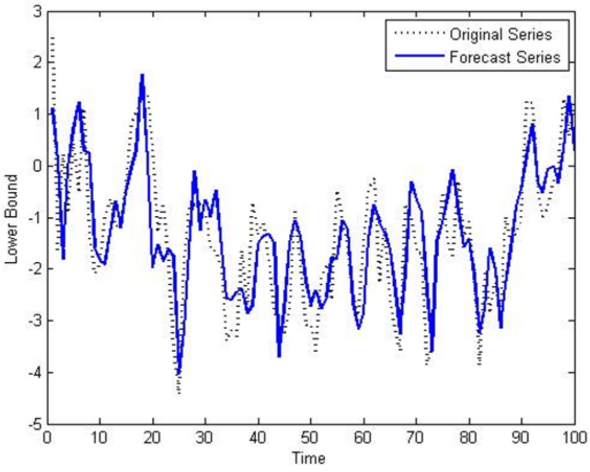

Figure 3. Lower bound of model (I)

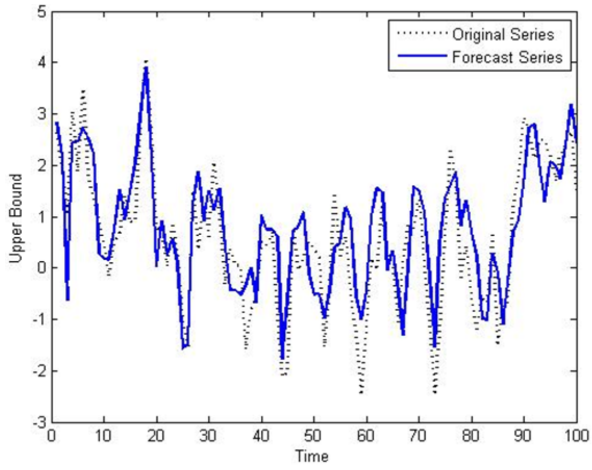

Figure 4. Upper bound of model (I)

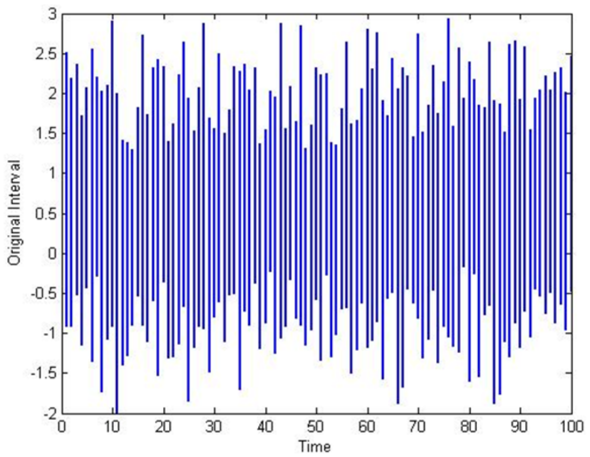

Figure 5. The actual value of model (II)

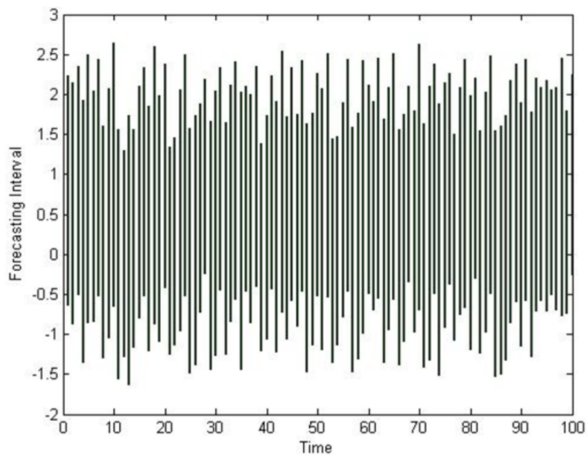

Figure 6. The forecasted value of model (II)

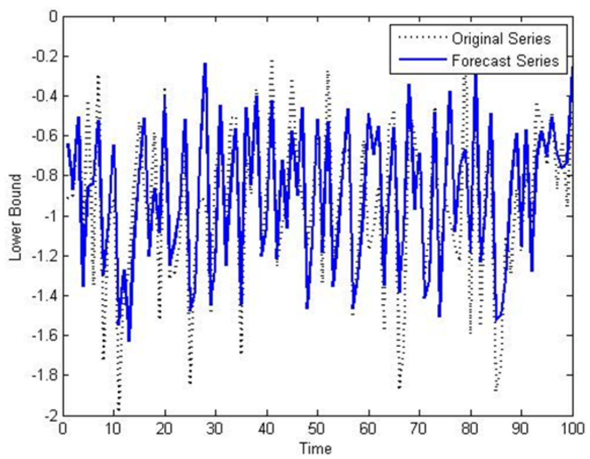

Figure 7. Lower bound of model (II)

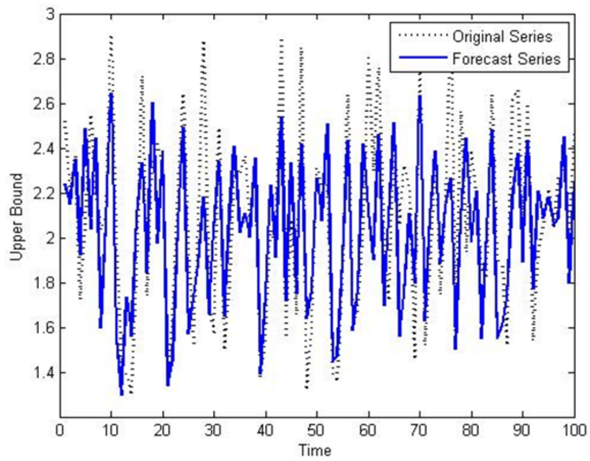

Figure 8. Upper bound of model (II)

Table 1. Basic information about the data sets: symbols and sample sizes

Symbols	AA	BA	CSCO	DD	DIS	INTC	JNJ	MSFT	T	XOM
size	383	540	872	1038	1208	2238	2487	3620	4171	6375

The interval average relative variance (ARV) [14,16] is used to test the forecast errors, i.e.,

$$\mathrm{ARV} = \frac{\sum_{t=1}^{K}\left(X_{t+1}^{L} - \hat{X}_{t+1}^{L}\right)^2 + \sum_{t=1}^{K}\left(X_{t+1}^{R} - \hat{X}_{t+1}^{R}\right)^2}{\sum_{t=1}^{K}\left(X_{t+1}^{L} - \bar{X}^{L}\right)^2 + \sum_{t=1}^{K}\left(X_{t+1}^{R} - \bar{X}^{R}\right)^2}$$

where K is the number of samples, $\left[X_t^L, X_t^R\right]$ is the actual value, $\left[\hat{X}_t^L, \hat{X}_t^R\right]$ is the forecasted value, $\bar{X}_t^L$ and $\bar{X}_t^R$ are respectively the mean values of lower bound and upper bound.

The first two-thirds of above samples are chosen as the training samples and the remaining samples are used for testing. Let $n = 1$, $m_1 = 3$, $N = 5$, $M = 100$, $r = 30$, $P_d = 0.67$, $P_c = 0.33$, $P_m = 0.05$. The simulation results are shown in Table 2.

TABLE 2. ARV for interval stock price time series

Symbols	**Holt's ([14,16])**	**iMLP ([15,16])**	**DPSO/ PSO-FCRBF ([16])**	**The proposed method**
AA	0.487	0.458	0.165	0.0921
BA	0.354	0.508	0.274	0.0448
CSCO	0.632	0.385	0.412	0.0251
DD	0.702	0.231	0.396	0.0369
DIS	0.468	0.524	0.135	0.0337
INTC	0.356	0.308	0.038	0.0132
JNJ	0.529	0.462	0.254	0.0186
MSFT	0.481	0.418	0.214	0.1777
T	0.467	0.418	0.217	0.0075
XOM	0.628	0.519	0.118	0.0062

By the expression of ARV, we can find that the lower ARV value leads to the higher forecasting accuracy. It is shown in Table 2 that the proposed method can achieve lower ARV values for some financial time series models than other methods. These facts mean that Bernstein fuzzy system is suitable for handling interval-valued financial time series.

5. **Conclusions.** In this paper, generalized Bernstein polynomials are applied to establish fuzzy system. And parameter identification method is proposed for the Bernstein fuzzy system to predict the intervalvalued time series. Since Bernstein polynomials possess good approximation accuracy for nonlinear function, in the future, we will apply it to performing pattern classification and solve differential equation.

Acknowledgment. This work is supported by the National Natural Science Foundation of China (71201019, 61374118, 71571035), the Program for Liaoning Excellent Talents in University (WJQ2014036), the Fundamental Research Funds for the Central Universities (DUT14QY30) and China Scholarship Council.

REFERENCES

[1] A. L. S. Maia, F. A. T. D. Carvalho and T. B. Ludermir, Forecasting models for interval-valued time series, *Neurocomputing*, vol.71, pp.3344-3352, 2008.

[2] T. Xiong, Y. K. Bao, Z. Y. Hu and R. Chiong, Forecasting interval time series using a fully complex-valued RBF neural network with DPSO and PSO algorithms, *Information Sciences*, vol.305, pp.77-92, 2015.

[3] T. Wang, C. Han and X. Jin, Design of interval type-2 fuzzy neural network system and application, *ICIC Express Letters, Part B: Applications*, vol.6, no.4, pp.1041-1047, 2015.

[4] A. M. S. Roque, C. Mate, J. Arroyo and A. Sarabia, iMLP: Applying multi-layer perceptrons to interval-valued data, *Neural Processing Letters*, vol.25, no.2, pp.157-169, 2007.

[5] C. F. Juang, Y. Y. Lin and R. B. Huang, Dynamic system modeling using a recurrent interval-valued fuzzy neural network and its hardware implementation, *Fuzzy Sets and Systems*, vol.179, pp.83-99, 2011.

[6] Y. Zhang, Interval-valued variable weighted synthesis inference method for fuzzy reasoning and fuzzy systems, *ICIC Express Letters, Part B: Applications*, vol.4, no.3, pp.511-518, 2013.

[7] W. Zeng and Y. Zhao, Multiple rules interval-valued approximate reasoning based on interval-valued similarity measure set, *ICIC Express Letters*, vol.7, no.4, pp.1265-1271, 2013.

[8] J. Y. Ahn, K. S. Han, S. Y. Oh and C. D. Lee, An application of interval-valued intuitionistic fuzzy sets for medical diagnosis of headache, *International Journal of Innovative Computing, Information and Control*, vol.7, no.5(B), pp.2755-2762, 2011.

[9] H.-L. Hsu and B. Wu, An innovative approach on fuzzy correlation coefficient with interval data, *International Journal of Innovative Computing, Information and Control*, vol.6, no.3(A), pp.1049-1058, 2010.

[10] I. Skrjanc, S. Blazi and O. Agamennoni, Interval fuzzy model identification using l_∞ norm, *IEEE Trans. Fuzzy Systems*, vol.13, no.5, pp.561-568, 2005.

[11] H. X. Li, X. H. Yuan and J. Y. Wang, The normal numbers of the fuzzy systems and their classes, *Sciences in China (Series F)*, vol.53, no.11, pp.2215-2229, 2010.

[12] H. Wold, Soft modeling: The basic design and some extensions, *Systems under Indirect Observations*, vol.2, pp.1-54, 1982.

[13] B. S. Dayal and J. F. MacGregor, Improved PLS algorithms, *Journal of Chemometrics*, vol.11, no.1, pp.73-85, 1997.

[14] A. L. S. Maia and F. A. T. de Carvalho, Holt's exponential smoothing and neural network models for forecasting interval-valued time series, *International Journal of Forecast*, vol.27, no.3, pp.740-759, 2011.

[15] A. M. S. Roque, C. Maté, J. Arroyo and Á. Sarabia, iMLP: Applying multi-layer perceptrons to interval-valued data, *Neural Processing Letters*, vol.25, no.2, pp.157-169, 2007.

[16] T. Xiong, Y. Bao, Z. Hu and R. Chiong, Forecasting interval time series using a fully complex-valued RBF neural network with DPSO and PSO algorithms, *Information Sciences*, vol.305, pp.77-92, 2015.