ICIC Express Letters
Volume **17**, Number **8**, August **2023**

 ISSN 1881-803X
pp. **899–906**

AN INFORMATION GRANULE COMPENSATION MODEL BASED ON BROAD LEARNING SYSTEM FOR MEDICAL TIME SERIES

Degang Wang[1,*], Shuhui Fan[1] and Wenyan Song[2]

[1]School of Control Science and Engineering
Dalian University of Technology
No. 2, Linggong Road, Ganjingzi District, Dalian 116024, P. R. China
13840808402@163.com
*Corresponding author: wangdg@dlut.edu.cn

[2]School of Economics
Dongbei University of Finance and Economics
No. 217, Jianshan Street, Shahekou District, Dalian 116025, P. R. China
songwydufe@dufe.edu.cn

Received December 2022; accepted February 2023

Abstract. *The interval-valued time series models have attracted researchers' attention recently because of their ability to make good use of the information contained in data and the correlation between variables. In this paper, a series interval information granules are generated based on the principle of justifiable granularity, and a threshold fragment model is established through clustering, which is composed of several functional coefficient autoregressive models. Then the volatility characteristics of the modeling object reflected by the results of the first stage model output are compensated. Furthermore, a differential compensation model is established by using the broad learning system to modify the prediction results. Finally, the effectiveness of the proposed method is demonstrated through the application of two medical datasets.*

Keywords: Information granule, Time series, Functional coefficient autoregressive, Differential compensation

1. **Introduction.** In time series, the value of data is changing dynamically all the time. The correlation between data at different times represents the development trend of the observed object. Discovering the patterns contained in time series and predicting its trend can help people to carry out problems such as traffic flow control [1], air pollution control [2], medical monitoring and analysis [3], and stock price prediction [4]. However, due to the influence of noise in reality, the predicted value of time series often has deviation. Due to the fact that the interval data can effectively describe the changing law of the object, and can provide a certain tolerance for prediction results, many researchers have turned to the interval-valued time series modeling.

On the other hand, since the granular computing theory can process data representation efficiently, and the information granules have certain interpretability, the prediction of time series combined with granular computing theory has become a hot topic in recent years. In [5], the original data sequence is granulated as some fuzzy information granules, which simplifies the dimension of the data. In [6], the Bayesian granularity calculation method for time series is proposed. In [7], the input data is granulated firstly, and then the interval boundary in the input space is optimized by an improved genetic algorithm. In the end, the multi-layer perceptron is adopted to realize the prediction of the interval time series.

DOI: 10.24507/icicel.17.08.899

In addition, in some complex forecasting scenarios, a single forecasting model cannot fully describe the objects. In order to solve this type of problem, a strategy of multiple models' combination for nonlinear time series forecasting has been considered. And the commonly used combining methods include weight distribution, error compensation, parameter and structure optimization [8]. Based on the weighted combination method, a differential autoregressive moving average-back propagation (BP) neural network prediction model is proposed to deal with clothing sales data in [9]. A two-stage error compensation model is proposed in [10], in which the first stage uses a single hidden layer feedforward neural network for rough compensation, and the second stage uses a semi-mechanical model for precise compensation. Besides, according to the idea of ensemble learning, the random forest method combined with a fuzzy Takagi-Sugeno model is used to predict time series in [11].

It can be seen that although there have been some researches to discuss the interval-valued time series modeling, how to mine the dynamic association between data and how to improve the performance of predictions still need to be considered. With the help of the information granule theory, which is an efficient data representation method, and the broad learning system, which is a powerful approach in the model identification, in this paper we design an information granule compensation model to construct an interval time series prediction model. Besides, the strategies of differential compensation and combined prediction are integrated for modeling and analyzing the interval time series in order to improve the accuracy of the model.

The structure of the paper is arranged as follows. In Section 2, we introduce the building process of information granule by using the principle of justifiable granularity and the form of functional coefficient autoregressive model. In Section 3, the construction of segmental compensation model based on the broad learning system is discussed in detail. In Section 4, two experimental studies on some medical datasets are carried out to verify the performance of the proposed model. Section 5 gives the conclusion.

2. **Preliminary.**

2.1. **Information granule design.**

The principle of justifiable granularity is a general guideline for constructing information granule which has been proposed in [12]. The meaning of this granularity criterion is outlined as follows.

1) The coverage degree *cov* measures the amount of data contained in the information granule. The more data the information granule contains, the more sufficient the information granule describes. The indicator coverage can be calculated as

$$cov = \max\left(0, 1/N \sum_{x_i : x_i \in D} f(x_i)\right),$$

where D is the dataset, N is the number of data samples in D, and $f(x_i)$ is the membership function for data x_i.

For an interval-type information granule $[a, b]$, suppose that $\chi(x_i)$ is the indicative function. When x_i is within the constructed interval $[a, b]$, $\chi(x_i)$ takes the value 1; otherwise it is 0. In other words, the coverage realizes the counting of the data, i.e., $cov = \frac{\sum_{i=1}^{N} \chi(x_i)}{N}$, in which $\chi(x_i) = \begin{cases} 1, & a \le x_i \le b, \\ 0, & x_i < a \text{ or } x_i > b. \end{cases}$

2) The specificity degree *sp* reflects the information expression of the granule about sample data. The smaller the information granule is, the less data it contains, and the more accurate the feature description of the dataset is. On the contrary, the more data the information granule contains, i.e., the wider the granule's range is, the weaker its ability of data features description is.

As to an interval-type granule, the specificity degree is expressed as

$$sp = \int_0^1 \left(1 - \frac{|m - b_\alpha|}{range}\right) d\alpha,$$

where $b_\alpha = f^{-1}(\alpha)$, and f is the membership function, which is monotonically non-decreasing. $range = |x_{\max} - m|$, where $x_{\max}$ is the maximum data value in the interval, and m is the mean value of the sample data.

It is not difficult to find that there is a contradiction between high coverage and high specificity. In order to take these two indicators into account, an objective evaluation function of information granularity has been introduced as

$$Q = \max_{\sigma}(cov * sp), \tag{1}$$

in which σ are parameters to be estimated, for instance, the end points a and b of an interval granule. The optimal information granule can be obtained by maximizing the above information granulation index Q.

2.2. **Functional coefficient autoregressive model.** As a nonlinear regression model, the functional coefficient autoregressive model [13] plays an important role in nonlinear time series prediction. The basic form of it is as follows:

$$x_t = \varphi_1(\boldsymbol{x}^*_{t-1})x_{t-1} + \varphi_2(\boldsymbol{x}^*_{t-1})x_{t-2} + \cdots + \varphi_p(\boldsymbol{x}^*_{t-1})x_{t-p} + \varepsilon_t, \tag{2}$$

$$\boldsymbol{x}^*_{t-1} = (x_{t-1}, x_{t-2}, \ldots, x_{t-p})^T, \tag{3}$$

where $\varphi_j(\boldsymbol{x}^*_{t-1}) \in \mathbb{R}\ (j = 1, 2, \ldots, p)$ is a measurable function. $\{\varepsilon_t\}$ is a sequence of random variables, which is a white noise only related to the current moment t and independent of x_{t-j}. Equation (2) is the generalized form of a functional coefficient autoregressive model, and different functional coefficient autoregressive models can be constructed by selecting different $\varphi_j(\cdot)$.

It is not hard to see that how to identify the coefficients is the most important process in establishing a functional coefficient autoregressive model. In general, the identification includes the linear part and the nonlinear part. For the linear part, the least squares method can be used to determine the parameters, and then the linear part is shaped as $\boldsymbol{Y} = \boldsymbol{\Phi X} + \varepsilon_t$. The parameter matrix $\boldsymbol{\Phi}$ can be calculated by $\boldsymbol{\Phi} = \left(\boldsymbol{X^T X}\right)^{-1} \boldsymbol{X^T Y}$, in which $\boldsymbol{X}$ is the matrix of input data and $\boldsymbol{Y}$ is the vector of output data. As to the nonlinear part, the genetic algorithm or other optimization algorithm can be chosen to determine parameters.

3. Interval Time Series Forecasting Model Based on Segmented Compensation.

3.1. **Segmented interval time series modeling.** In this paper, the expression of center-radius data pair is adopted to describe the interval data, such as $[\boldsymbol{x}]_t = (x^c_t, x^r_t)^T$. And in order to predict interval time series more accurately, the fuzzy C-means (FCM) clustering algorithm is used to divide the prediction model into n sub-models.

First, the input of a clustering algorithm is settled as $\boldsymbol{W} = \{[\boldsymbol{x}]_t, [\boldsymbol{x}]_{t+1}, \ldots, [\boldsymbol{x}]_{t+p}\}$, where $[\boldsymbol{x}]_j = \left(x^c_j, x^r_j\right)^T\ (j = t, t+1, \ldots, t+p)$ represents a set of center sequence and radius sequence values. p is the time lag order, which can be determined by the Akaike's information criterion (AIC). By using the FCM algorithm, the center of each class and the membership matrix for all the training data can be obtained. Assume that the number of clusters is n, in other words, the number of sub-models is n. According to the clustering results, the matrices $W_1, W_2, \ldots, W_n$ are used for every sub-model construction. The first p columns of the input matrix $W_1, W_2, \ldots, W_n$ represent the input data, and the last column represents the real output. Every sub-model is set as a threshold autoregressive

model. The threshold autoregressive model [14] is a type of functional coefficient autoregressive model, which has shown good results in fitting nonlinear and periodic problems. The staged threshold autoregressive model is as follows:

$$x_t = \sum_{k=1}^{n} \left(\varphi_0 + \sum_{i=1}^{p} \varphi_i^{(k)} x_{t-i} + \varepsilon_t \right) \cdot \chi(\lambda_{k-1} \le z_{t-d} < \lambda_k), \tag{4}$$

$$\chi(\lambda_{k-1} \le z_{t-d} < \lambda_k) = \begin{cases} 1, & \lambda_{k-1} \le z_{t-d} < \lambda_k \\ 0, & \text{otherwise} \end{cases} \tag{5}$$

where d is the lag order of the threshold term, and z_{t-d} is a mapping function of x_{t-d} $(d < p)$, which is usually settled as $z_{t-d} = x_{t-d}$ for simplicity. ε_t is a random noise that obeys the normal distribution. $\chi(\cdot)$ is the indicative function that acts as a switch for the sub-model selection. Equations (4) and (5) give the definition of the numerical threshold autoregressive model.

Next, in order to deal with the interval time series prediction, the interval sequence is used as the input to establish an extended interval threshold autoregressive model as follows:

$$x_t^c = \sum_{k=1}^{n} \left(a_0 + \sum_{i=1}^{p} \left(a_1^{(ik)} x_{t-i}^c + a_2^{(ik)} x_{t-i}^r \right) + \varepsilon_{ct} \right) \cdot \chi(\lambda_{k-1} \le z_{t-d} < \lambda_k), \tag{6}$$

$$x_t^r = \sum_{k=1}^{n} \left(b_0 + \sum_{i=1}^{p} \left(b_1^{(ik)} x_{t-i}^c + b_2^{(ik)} x_{t-i}^r \right) + \varepsilon_{rt} \right) \cdot \chi(\lambda_{k-1} \le z_{t-d} < \lambda_k), \tag{7}$$

in which $a_1^{(ik)}$ and $a_2^{(ik)}$ represent the i-th order parameter values of the k-th sub-model, and the interval boundary $[\boldsymbol{x}]_t$ at the t-time is optimized by the center and radius values at previous $t-1$ to $t-p$ times.

After the threshold autoregressive model is established, the parameters need to be identified. As it can be seen that the threshold autoregressive model essentially contains several piece-wise linear methods, the parameters can be estimated by the least squares method. Set the input matrix $\boldsymbol{X}_{in}^{(k)}$ as

$$\begin{bmatrix} 1 & x_{t-1}^c & x_{t-1}^r & \cdots & x_{t-p}^c & x_{t-p}^r & \cdots & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & & 0 & 0 & & 1 & x_{t-1}^c & x_{t-1}^r & x_{t-p}^c & x_{t-p}^r \\ & & & & & & \vdots & & & & & \\ 1 & x_{lk,t-1}^c & x_{lk,t-1}^r & \cdots & x_{lk,t-p}^c & x_{lk,t-p}^r & \cdots & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & & 0 & 0 & & 1 & x_{lk,t-1}^c & x_{lk,t-1}^r & x_{lk,t-p}^c & x_{lk,t-p}^r \end{bmatrix}, \tag{8}$$

and the functional coefficient matrix as

$$\boldsymbol{\phi}^{(k)} = \left(a_0^k, a_1^{k1}, a_2^{k1}, \ldots, a_1^{kp}, a_2^{kp}, b_0^k, b_1^{k1}, b_2^{k1}, \ldots, b_1^{kp}, b_2^{kp} \right)^T. \tag{9}$$

Accordingly, the output vector for the training data is denoted as

$$\boldsymbol{X}_{out}^{(k)} = \left(x_t^c, x_t^r, \ldots, x_{lk,t}^c, x_{lk,t}^r \right)^T, \tag{10}$$

where lk is the number of training samples in the k-th sub-model, and $\boldsymbol{\varepsilon}_t = (\varepsilon_{ct}, \varepsilon_{rt})^T$ is the white noise that obeys the normal distribution. According to Equations (8)-(10), the generalized k-th sub-model can be established as

$$\widetilde{\boldsymbol{X}}^{(k)} = \boldsymbol{\phi}^{(k)} \boldsymbol{X}_{in}^{(k)} + \boldsymbol{\varepsilon}_t.$$

The model parameters can be obtained by a numerical algorithm or an optimization algorithm. In this paper, the least squares method is used to determine the parameters.

The objective function is taken as the residual sum of squares, i.e.,

$$E\left(\boldsymbol{\phi}^{(k)}\right) = \arg\min_{\boldsymbol{\phi}^{(k)}} \left(\widetilde{\boldsymbol{X}}^{(k)} - \boldsymbol{X}_{out}^{(k)}\right)^2.$$

The parameter matrix $\boldsymbol{\phi}^{(k)}$ can be expressed by the following formula:

$$\boldsymbol{\phi}^{(k)} = \left(\left(\boldsymbol{X}_{in}^{(k)}\right)^T \boldsymbol{X}_{in}^{(k)}\right)^{-1} \left(\boldsymbol{X}_{in}^{(k)}\right)^T \boldsymbol{X}_{out}^{(k)}.$$

3.2. **Compensation combination model based on the broad learning system.** The segmented threshold autoregressive model is the first stage of the combined modeling, which is used to achieve an approximate output to the real data and to obtain a volatility sequence. The second stage of the complete modeling is the combined compensation model based on the broad learning system, which realizes the variable-weight compensation of the volatility sequence.

According to the piece-wise combined compensation model, a volatility sequence, i.e., a difference sequence $\left\{\left(\delta_t^c, \delta_t^r\right)^T, t = 1, 2, \ldots, N-1\right\}$ is established.

$$\delta_t^c = \tilde{x}_t^c - x_{t-1}^c, \quad \delta_t^r = \tilde{x}_t^r - x_{t-1}^r,$$

where $\tilde{x}_t^c$ and $\tilde{x}_t^r$ represent the center sequence value and the radius sequence value of the t-th interval segment. They can be obtained from the first stage model discussed in the previous section. Here x_{t-1}^c and x_{t-1}^r represent the center sequence value and the radius sequence value of the $(t-1)$-th interval segment, which are the actual recorded values. δ_t^c and δ_t^r represent the center and radius volatility values of the segmented threshold autoregressive model. Since the neural network has a good ability of processing nonlinear data, the compensated difference sequence is modeled by the broad neural network.

The specific implementation process is as follows.

Step 1. The lag order q of the volatility sequence is determined by the AIC criterion. Thereupon, the establishment time vector $\left\{\left(\delta_t^c, \delta_t^r\right)^T\right\}$ is obtained, and the vector matrix $\boldsymbol{X}_r$ is established as

$$\boldsymbol{X}_r = \left[[\boldsymbol{\delta}]_1, [\boldsymbol{\delta}]_2, \ldots, [\boldsymbol{\delta}]_{N-q}\right]^T = \begin{bmatrix} \delta_{1,t}^c & \delta_{1,t}^r & \cdots & \delta_{1,t-q+1}^c & \delta_{1,t-q+1}^r \\ \delta_{2,t}^c & \delta_{2,t}^r & & \delta_{2,t-q+1}^c & \delta_{2,t-q+1}^r \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \delta_{N-q,t}^c & \delta_{N-q,t}^r & \cdots & \delta_{N-q,t-q+1}^c & \delta_{N-q,t-q+1}^r \end{bmatrix}$$

where $[\boldsymbol{\delta}]_j = \left(\delta_{1,j}^c, \delta_{1,j}^r, \ldots, \delta_{N-q,j}^c, \delta_{N-q,j}^r\right)$, $j = t, t+1, \ldots, t-q+1$.

Step 2. Each row of the volatility matrix $\boldsymbol{X}_r$ is an input from a group of samples, which contains $N-q$ sample data in total. The volatility matrix for sample data is divided into the training set $\boldsymbol{X}_{train}$ and the testing set $\boldsymbol{X}_{test}$. Next, the training set data are clustered into n clusters by using the FCM algorithm. And suppose that n sub-classes $\left\{\boldsymbol{X}_{train}^{(1)}, \boldsymbol{X}_{train}^{(2)}, \ldots, \boldsymbol{X}_{train}^{(n)}\right\}$ are obtained, where the vector matrix $\boldsymbol{X}_{train}^{(k)}$ $(k = 1, 2, \ldots, n)$ is taken as the input of the k-th sub-model.

Step 3. Initialize the parameters in the broad structure sub-model, such as the number of feature layer nodes, and the number of enhancement layer nodes.

Step 4. Initialize the randomization feature layer weight $\boldsymbol{W}_{im}$ and the enhancement layer weight $\boldsymbol{W}_{ef}$. Determine the feature output $\boldsymbol{Z}^{(k)}$ and $\boldsymbol{H}^{(k)}$ of the k-th sub-model feature layer and enhancement layer as $\boldsymbol{Z}^{(k)} = \varphi\left([\boldsymbol{\delta}]_t^{(k)} \cdot \boldsymbol{W}_{im}^{(k)}\right)$, $\boldsymbol{H}^{(k)} = \xi\left(\boldsymbol{Z}^{(k)} \cdot \boldsymbol{W}_{ef}^{(k)}\right)$, in which $[\boldsymbol{\delta}]_t^{(k)}$ is the input of the volatility sequence for the k-th sub-model. φ and ξ are some presupposed transformation functions. For the convenience of calculation, the last columns of $\boldsymbol{W}_{im}^{(k)}$ and $\boldsymbol{W}_{ef}^{(k)}$ matrices are recorded as the offsets.

Step 5. By the grid searching, the optimal values of parameters in every broad sub-model can be obtained, which are used as the structural parameters of the next incremental learning.

Step 6. Update the feature nodes by the incremental broad learning system to obtain the new feature layer output $\widetilde{\boldsymbol{Z}}^{(k)} = \left[\boldsymbol{Z}^{(k)} \middle| \boldsymbol{Z}_{new}^{(k)}\right]$ and the enhancement layer output $\widetilde{\boldsymbol{H}}^{(k)} = \xi\left(\boldsymbol{Z}^{(k)} \middle| \boldsymbol{Z}_{new}^{(k)}, \boldsymbol{W}_{ef}^{(k)} \middle| \boldsymbol{W}_{ef\_new}^{(k)}\right)$. By using the gradient descent method, the updated weight matrix $\boldsymbol{W}_{im}^{(k)}$, $\boldsymbol{W}_{ef}^{(k)}$ and $\boldsymbol{W}_{m}^{(k)}$ can also be obtained. Hence, the volatility output value of the k-th sub-model is

$$\left[\tilde{\delta}\right]_{t+1}^{(k)} = \left[\widetilde{\boldsymbol{Z}}^{(k)} \middle| \widetilde{\boldsymbol{H}}^{(k)}\right] \boldsymbol{W}_{m}^{(k)},$$

in which $\left[\tilde{\delta}\right]_{t}^{(k)} = \left(\tilde{\delta}_{kt}^{c}, \tilde{\delta}_{kt}^{r}\right)^{T}$ is the center-radius data pair at the time t.

Step 7. When the combined model is adopted to predict new input data, the distance between the current prediction and the cluster center of each sub-model is calculated respectively. The weight of combination is assigned according to this distance. The farther the distance is, the smaller the weight is. Set d_k as the distance from the volatility to the k-th class center, and the corresponding weight is $w_k = \frac{1}{d_k} \Big/ \sum_{k=1}^{n} \frac{1}{d_k}$. Then the correction result of volatility compensation is calculated by the weighted combination:

$$\tilde{\delta}_t^c = \sum_{k=1}^{n} w_k \tilde{\delta}_{kt}^c, \quad \tilde{\delta}_t^r = \sum_{k=1}^{n} w_k \tilde{\delta}_{kt}^r.$$

Step 8. The final results $\hat{x}_t^c = \tilde{x}_t^c + \tilde{\delta}_t^c$ and $\hat{x}_t^r = \tilde{x}_t^r + \tilde{\delta}_t^r$ are predicted by the threshold autoregressive sub-model and corrected by the compensation value, where $(\tilde{x}_t^c, \tilde{x}_t^r)$ is the prediction result of the threshold autoregression, and $(\hat{x}_t^c, \hat{x}_t^r)$ is the prediction result corrected by the segmentation combination compensation.

3.3. **Evaluation indicators.** In this paper, the following two indicators are considered, which are suitable for evaluating interval values confirmed in some existing literature.

1) Average relative variance (ARV):

$$ARV = \frac{\sum_{i=1}^{N}\left(y_i^l - \hat{y}_i^l\right)^2 + \sum_{i=1}^{N}\left(y_i^u - \hat{y}_i^u\right)^2}{\sum_{i=1}^{N}\left(y_i^l - \bar{y}_i^l\right)^2 + \sum_{i=1}^{N}\left(y_i^u - \bar{y}_i^u\right)^2},$$

in which y_i^u and y_i^l represent the real upper and lower bounds of the interval. $\hat{y}_i^u$ and $\hat{y}_i^l$ represent the upper and lower bounds of the predicted interval. $\bar{y}_i^u$ and $\bar{y}_i^l$ are the mean values of the upper and lower bounds of the interval. N is the number of the data.

2) Prediction accuracy (PA):

$$PA = \sum_{t=1}^{N} \frac{\chi(t)}{N} \times 100\%,$$

where $\chi(t)$ is the indicative function, i.e.,

$$\chi(t) = \begin{cases} 1, & \text{if } \hat{y}_t = y_t, \\ 0, & \text{if } \hat{y}_t \neq y_t. \end{cases}$$

4. Simulation Experiments.

4.1. **MIMIC-III public dataset.** The MIMIC-III public dataset is a large-scale database of critically ill patients, which records a number of patients' indicators including vital signs, medication, and medical records [15]. As the heart rate is often used to evaluate

the state of patient's health, and the heart rate abnormal is often a precursor of many diseases, the heart rate fluctuations of the ICU patients are considered to evaluate the validity of the proposed model in this experiment. Besides, in order to obtain an accurate interval prediction result, some related factors that affect the heart rate, such as respiration rate, blood pressure, and oxygen saturation, are considered together in the process of modeling.

Eight patients are randomly selected from the dataset. A model that contains two threshold autoregressive sub-models is established to perform the interval prediction of the heart rate, and then the volatility is input into the broad learning system. For brevity, in this example the number of the threshold autoregressive sub-models is set to 2. The time lag order is 3. The number of sub-models in the broad learning system is 2, and the lag order of the difference sequence model is 2. The evaluation index of the prediction accuracy is calculated. Table 1 shows the results of the proposed model. It can be seen from Table 1 that the segmented combined compensation prediction model accurately predicts the true interval range of the heart rate, and the interval average relative variance also maintains a small error level.

Table 1. Heart rate prediction evaluation

Patient	Prediction accuracy	Patient	Prediction accuracy
#109	87.1428%	#4787	95.8329%
#605	96.1538%	#5242	90.9091%
#711	96.6667%	#8427	82.7586%
#1709	93.5484%	#8492	94.1176%

4.2. **Uremic nephropathy patient sign dataset.** In the process of dialysis treatment for uremia patient, it is necessary to monitor the blood pressure in real time. The fluctuation of blood pressure will not only affect the effect of dialysis, but even will cause serious complications, such as hypotensive shock and cerebral hemorrhage. Under normal circumstance, the systolic blood pressure is maintained at 90 to 139 mmHg. 140 to 179 mmHg is hypertension, and greater than 180 mmHg is severe hypertension. In this example, the combined compensation method of segmented interval modeling is used to predict the blood pressure intervals for some uremic nephropathy patients.

The physical signs of real uremia patients in a hospital are investigated in this example, including some physiological indicators of 374 patients before and after dialysis from July 2019 to July 2020. Physical signs, such as ultrafiltration volume, body weight, systolic blood pressure, diastolic blood pressure, heart rate, are used to predict the blood pressure. The systolic blood pressure after dialysis is selected as the main factor, and the heart rates before and after dialysis and the weights before and after dialysis are used as secondary factors. And the average values of evaluation indicators including the ARV and the PA are calculated. The results are shown in Table 2. Through the comparison to the BP neural network and the support vector regression (SVR), the prediction model designed in this paper achieves satisfactory results.

Table 2. Comprehensive evaluation of patients

Evaluation indicator	BP	SVR	Segment combination compensation model
Average ARV	8.5642	2.8787	**2.5291**
Average PA	59.8391%	51.1245%	**76.8017%**

5. **Conclusions.** In this paper, a segmented combined compensation forecasting method which includes two parts, is designed to carry out the interval time series forecasting. In the first part, an interval time series is constructed through the information granular technique and the threshold autoregressive method. In the second part, the volatility sequence is compensated by a variable-weight incremental broad learning system. Through two simulation comparison experiments on medical datasets, the effectiveness of the proposed model is verified. In the future study, the form of combined model and the optimal algorithm for parameter identification will be investigated and compared further.

Acknowledgment. This work is partially supported by the Humanities and Social Science Project for Ministry of Education of China (21YJA630079), the Fundamental Research Funds for the Central Universities (DUT22YG227) and the National Natural Science Foundation of China under Grant (12071056).

REFERENCES

[1] A. Agafonov and A. Yumaganov, Short-term traffic flow forecasting using a distributed spatial-temporal k nearest neighbors model, *IEEE International Conference on Computational Science and Engineering*, Bucharest, Romania, pp.91-98, 2018.

[2] Z. Qi, T. Wang, G. Song, W. Hu, X. Li and Z. Zhang, Deep air learning: Interpolation, prediction, and feature analysis of fine-grained air quality, *IEEE Transactions on Knowledge and Data Engineering*, vol.30, no.12, pp.2285-2297, 2018.

[3] A. Anguera, J. M. Barreiro, J. A. Lara and D. Lizcano, Applying data mining techniques to medical time series: An empirical case study in electroencephalography and stabilometry, *Computational and Structural Biotechnology Journal*, vol.14, pp.185-199, 2016.

[4] P. Tüfekci, *Predicting the Direction of Movement for Stock Price Index Using Machine Learning Methods*, Springer International Publishing, Switzerland, 2016.

[5] Y. He, Y. Yan and Q. Xu, Wind and solar power probability density prediction via fuzzy information granulation and support vector quantile regression, *International Journal of Electrical Power and Energy Systems*, vol.113, pp.515-527, 2019.

[6] O. Hryniewicz and K. Kaczmarek, Bayesian analysis of time series using granular computing approach, *Applied Soft Computing*, vol.47, pp.644-652, 2016.

[7] M. Qi, Y. Gu and Q. Wang, Internet financial risk management and control based on improved rough set algorithm, *Journal of Computational and Applied Mathematics*, vol.384, 113179, 2021.

[8] J. M. Bates and C. W. J. Granger, The combination of forecasts, *Journal of the Operational Research Society*, vol.20, no.4, pp.451-468, 1969.

[9] S. Yu, H. Dong, Y. Chen et al., Clothing sales forecast based on ARIMA-BP neural network combination model, *IEEE International Conference on Power, Intelligent Computing and Systems*, Shenyang, China, pp.367-372, 2019.

[10] H. Tao, R. Chen, J. Xuan et al., Prioritization analysis and compensation of geometric errors for ultra-precision lathe based on the random forest methodology, *Precision Engineering*, vol.61, pp.23-40, 2020.

[11] X. Qiu, Z. Le, P. N. Suganthan et al., Oblique random forest ensemble via least square estimation for time series forecasting, *Information Sciences*, vol.420, pp.249-262, 2017.

[12] W. Pedrycz and W. Homenda, Building the fundamentals of granular computing: A principle of justifiable granularity, *Applied Soft Computing*, vol.13, no.10, pp.4209-4218, 2013.

[13] R. Chen and R. S. Tsay, Functional coefficient autoregressive models, *Journal of the American Statistical Association*, vol.88, no.421, pp.298-308, 1993.

[14] H. Tong and K. S. Lim, Threshold autoregression, limit cycles and cyclical data, *Journal of the Royal Statistical Society, Series B: Methodological*, vol.42, no.3, pp.245-268, 1980.

[15] A. Johnson, T. J. Pollard, L. Shen et al., MIMIC-III: A freely accessible critical care database, *Scientific Data*, vol.3, 160035, 2016.