International Journal of Intelligent Information and Management Science

Volume 3, Issue 1, February 2014

http://www.hknccp.org

Publisher: HongKong New Century Cultural Publishing House

Address: Unit E79, 3/F., Wing Tat Commercial Building, 97 Bonham Strand East, Sheung Wan, HK.

International Journal of Intelligent Information and Management Science
ISSN: 2307-0692 Volume 3, Issue 1, February 2014

Contents

International Journal of Intelligent Information and Management Science
ISSN: 2307-0692 Volume 3, Issue 1, February 2014

Construction and Application of Compacted Wavelet Neural Network Model

Zhike KUANG
Hunan City University, Hunan Yiyang 413000, China

Abstract: Because of several problems, such as randomness of neural network model in predicting the presence, lack of transparency in mechanism, difficulty to determine the initial parameters, the phenomenon of over-fitting and easiness to fall into local minima, this paper presents a compacted wavelet neural network model. The model will transplant the wavelet function to the hidden layer of neural network in place of sigmoid activation function, and uses a randomly determined state command to obtain certain predictions. Finally, on the gas emission wavelet packet-Wavelet network forecasting experimental results, it showed that: compacted wavelet neural network model is possessed of higher speed of training, easier operation, applicability to large quantities of data for training and processing, data adaptability and robustness. It is more convenient than optimized algorithms such as genetic algorithms, particle swarm algorithm.

Keywords: Compactness; Convergence; Wavelet Function; Reconstruction

1. Introduction

Wavelet neural network is a transferring function, a kind of BP neural network topology based on the wavelet function as the hidden layer, a neural network when signal is propagated, the forward error back propagate. Wavelet neural network in the experiments use 3 layers network: input layer, hidden layer and output layer [1-3].
With Introduction into the field of neural network, prediction theories and methods of forecasting produced a qualitative leap [4]. The traditional linear prediction methods, such as Auto Regressive model, Moving Average in solving the problems of nonlinear models prediction encountered great difficulties, and neural network in nonlinear prediction has its unique advantages, it does not need to build complex nonlinear systems and mathematical model of explicit relationships, it can extracts data characteristics and internal rules through the training data samples. It makes the information distributed storage come true, resulting in associative memory. Thus the untrained samples can be extrapolated to predict the effect so that it provides a powerful tool for nonlinear prediction [5-8].
It was for the first time that carrying out forecasting by using neural network for nonlinear time series, pioneering the field of neural networks used to predict in 1987 [9-13]. After that, neural network had a rapid development in the application of prediction. Wavelet analysis, a mathematical theory developed in recent years, is considered a major breakthrough since the Fourier analysis. Based on wavelet analysis, Wavelet network is a class of network to be constructed that combines time-frequency localization properties of wavelet transformation and self-learning ability of neural network [14-17]. Wavelet neural network has received growing concern as a novel neural network. It is both possessed of time-frequency localization properties of wavelet and neural network's function approximation and generalization ability. And it has wined a strong advantage in the field of predicting [18].
Currently, as for the neural network to predict, there are two main forms: Trend forecasting and regression-based causality, corresponding time series prediction and multiple regression prediction. Neural networks, with distributed, associative, memorial and strong generalization ability, as well as self-learning ability and fault tolerance, can approximately approach nonlinear functions with arbitrary precision [19-20]. It can not be matched by the linear prediction method. For most predictive objects, especially data with non-linear relationship, using the neural network will get higher prediction accuracy. However, there are the following problems to come forth when the neural network is used to predict: randomness of predicting results, lack of transparency in mechanism, difficulty to determine the initial parameters, the phenomenon of over-fitting and easiness to fall into local minima and so on. Most of those problems need to be determined that based on the experimental results, by using statistical methods to evaluate the predicted results, or

using trial and error to find the optimal parameter that is convenient for further prediction.

The more prominent problem of above-mentioned problems is the randomness of neural network's forecasting result, wavelet neural network has no exception, that is, many predicted results are different, sometimes dispersive, namely the prediction accuracy of neural networks has uncontrollable nature. In this regard, there is little introduction in the current literature. This paper presents a simple and practical predicting method for wavelet neural networks, and can get a stable prediction results.

This paper mainly made a work in the following areas expansively and innovatively:

(a)As for randomness of neural network model in predicting the presence, lack of transparency in mechanism, difficulty to determine the initial parameters, the phenomenon of over-fitting and easiness to fall into local minima, this paper presents a compacted model of wavelet neural network. The model is a simple and practical method for determining the prediction, which transplant the wavelet function to hidden layer of neural network in place of Sigmoid activation function, using a command in randomly determined state to obtain certain predictions. And compared with the wavelet neural network of programming and BP network, this method is suitable for large quantities of data's training, has its adaptability of data sample and ability of robustness, especially for time series with high frequency and randomness having a better ability to adapt, with features of identifying predicted results and practicability. And it can significantly improve training's speed, prediction accuracy and prediction efficiency of the model.

(b) In order to further validate the correctness and validity of the proposed compacted model of wavelet neural network, a experiment was carried on in the environment of MATLAB R2006b which wad based on wavelet packet transformation and wavelet neural network's gas emission. Network structure is 7-15-1; training precision is set to 0.0001; function of hidden layer is wavelet network toolbox " $wavenet\_tool$ "; the output layer selects " $purelin$ " function; it was trained by using the " $trainlm$ " algorithm. The initial condition in the experiment is set to Q = 9, which uses the command rand ('state', 9) initialized wavelet network parameters. The simulation results show that: the compacted model of wavelet neural network has the characteristics of strong robustness, high speed of training, easy operation.

2. Uncertainty of Simulation Prediction's Result of Model

A large number of experiments of neural network's prediction model show that: the initial parameters of the network have a great influence on the predicted results. When the network structure is determined, namely the number of neurons of the network's input layer, hidden layer and output layer, and the learning rate, training accuracy are determined, the predicted results depend on initial parameters of the network. The initial parameters include network weights and threshold limit values and as for wavelet neural network it also comprises translation factor and scale parameters.

The initial parameters of neural networks are usually set as random number [-1, 1], and it is the leading cause of predictions' uncertainty. In the premise of determining the network structure, if the initial parameter set a determined value, predictions must be unique. Experiments show that as to the commonly used three-layer neural network, the initial value of network parameters has the greatest impact, followed by training's accuracy, the number of hidden layer's neurons, learning rate and momentum factor and so on.

Whether it is for BP network or for wavelet neural network, the network's initialization process is very important for whether the network's subsequent study is converged or not and for the speed of convergence; initial weights are well chosen, they can greatly accelerate convergence; initial weights are set incorrectly, then the times of learning will be greatly increased, even contributing to non-convergence; as for wavelet network, if the scale parameter and the initialization of displacement parameter are inappropriate, it will cause the learning process of the entire network does not converge.

Chen Guo had proposed several initialization methods of network parameters, which revealed the links among initial parameters, network structure and learning samples along with some limitations [21]. Currently the methods to optimize the network initial parameters such as genetic algorithms, particle swarm optimization, chaos optimization and so on, are very impressive for more complex networks. Furthermore, optimization algorithm itself is not just one time to determine the optimal values.

3. Construction of Compacted WNN

3.1. Compacted WNN

Compacted wavelet neural network is a product of combination between wavelet function and neural network, referring to a wavelet function or scaling function as activation function of neural network's hidden layer forming neurons, like Morlet wavelet function in place of BP neural network's hidden layer, Sigmoid activation function, the structure of the network is shown in Figure 1.

Figure 1 shows the three layers structure of wavelet network. Wherein, xi is the network's input variable, outputting to be a neuron. In this paper the outputting y corresponds to gas emission; $y_i(t)$ is the wavelet function ($i=1,2,...,m$); v_{ni}, w_i are connection weights between input layer / hidden layer, hidden layer / output layers. The currently used mother wavelets have good locality and smoothness like spline wavelet and Morlet wavelet. These functions' flexibility and parallelism can constitute an orthogonal basis of L2 (R), generating the most precise wavelet series at approximation. The hidden layer in this paper adopts Morlet wavelet function, the expression is described as

$$y\ (x) = \cos(1.86x)\exp\left(-x^2/2\right) \tag{1}$$

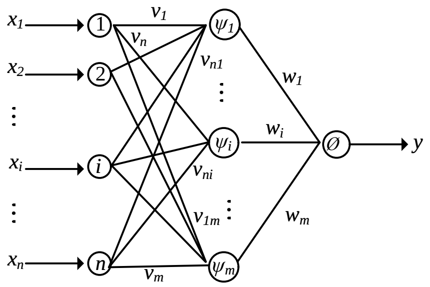

Figure 1. Wavelet Neural Network's Structure

Wavelet transformation corresponds to the Hilbert space of square integral L2 (R), if there exists a function

$$y\ (x) \in k^2(t), \int_t \left|y\ (k)\right|^2 dt\ \mathbf{p}\ +\infty$$

And its Fourier transforms into

$$g_y\ = \int_t \frac{\left|y\ (m)\right|^2}{|m|} jm\ \mathbf{p}\ +\infty \tag{2}$$

The function ψ (t) is a based wavelet, through the based wavelet's flexibility and parallelism, it can obtain a bases wavelet's function family

$$y_{a,b}(k) = \frac{1}{\sqrt{|a|}} y\left(\frac{k-b}{a}\right) \tag{3}$$

Wherein: a, b∈R, a≠0. a and b are factors of flexibility and parallelism. Wavelet transforms into

$$\begin{aligned}(v_m t)(a,b) &= \left\langle t(k), y_{a,b}(k)\right\rangle \\ &= \int_m f(k) y^*_{a,b}(k)\,dt\end{aligned} \tag{4}$$

Ψ (x) is the wavelet function. Formula (4) shows the change of wavelet is similar to the projection of the signal wavelet function on the based wavelet function, or a comparison between the signal and the wavelet in the associated position of a, b to transform into ($wy\ f$) (a, b), which described the degree of similarity in two aspects. The size of the projection reflects of the size of signal on the scale energy.

In Figure 1, the wavelet network structure consists of three layers: the input layer, hidden layer and output layer. Let three neurons, set as $n, m, 1$ respectively, excitation function of each hidden layer's neuron is $y\ a,b(x)$, the excitation function of output layer neuron takes Sigmoid, then the output expression is

$$t^e(k) = f\left(\sum_{k=1}^{n} w_n y\left(\sum_{j=1}^{m} h_{ik} s_i^j\right)(k) - l_i)/g_i\right) \tag{5}$$

Wherein, $p = 1, 2, ..., m$, p = 1, 2, ..., P (P is the number of samples). Let y be the actual value, $\hat{t}$ the predicted value, then the training samples p, taking the error energy function:

$$L = \frac{1}{2}\sum_{j=1}^{P}\sum_{i=1}^{m}\left(\hat{t}_i^j - t_i^j\right)^2 \tag{6}$$

$$J_{ik}(s) = j_{ik}(s-1) - h_m \frac{\partial s}{\partial s_{ik}} + \partial\ \ s_{ik} \tag{7}$$

$$r_i(s) = r_i(s-1) - h_v \frac{\partial s}{\partial r_i} + \partial \Delta r_i \tag{8}$$

$$a(s) = a(s-1) - h_a \frac{\partial r}{\partial a} + \partial\ \ a \tag{9}$$

$$b(s) = b(s-1) - h_b \frac{\partial r}{\partial b} + \partial\ \ b \tag{10}$$

Wherein: h_m, h_v, h_a, h_b are the learning rates, α is the momentum factor, by adjusting the parameters of WNN, make the formula (6) minimize.

3.2. Wavelet Neural Network Toolbox

The new version of MATLAB software offers Wavelet networks order to achieve wavelet neural network, but it is far from easiness to use BP neural network toolbox. Currently the application of wavelet neural networks is mainly achieved by programming. The general program designation is more complex, the programming cycle is long, and their types are different, especially for certain data sets or large quantities of data which are not easy to train, so such that the inherent superiority of wavelet neural network has been confined. The key to achieve the wavelet neural network toolbox is a user-defined transferring function, that is, there is a need to create wavelet function. By replacing BP neural network toolbox $\tan sig, \log sig$ with Morlet wavelet function and its derivatives:

$$n\ = \cos(1.78x) \times \exp\left(-x^2/2\right) \tag{11}$$

$$\begin{aligned}\frac{sy}{sx} &= -x\cos(1.78x)\times\exp\left(-x^2/2\right) \\ &-1.78\sin(1.78x)\times\exp\left(-x^2/2\right)\end{aligned} \tag{12}$$

Above treatments were replaced only for activation functions, not having a function of flexibility and parallelism. The network is lack of a flexible factor and paralleled factor b. the following evidence, which can be later incorporated into the equivalent connection weights and threshold limit values.

$$y_{a,b}\left(\sum_{i=1}^{m} v_{ik} x_i^p + q_{ik}\right) = y\left(\left(\sum_{i=1}^{m} v_{ik} x_i^p + q_{ik} - r_h\right) / a_k\right)$$
$$= y\left(\sum_{i=1}^{m} \frac{t_{ik}}{a_k} s_i + \frac{q_{ik} - s_k}{a_k}\right) = y\left(\sum_{i=1}^{r} m' s_i + q_{ik}'\right) \quad (13)$$

From the formula (13), it shows that the flexible factor ah and the paralleled factor b_h in the formula (5) shift into factors and threshold limit values, which weights and threshold limit values include function of flexible and paralleled factors. Training neural network like ordinary BP neural network's, provides training functions by using the toolbox, such as training function taking "$trainlm$", the activation function of output layer neuron using "$\log sig$" or "$purelin$".

3.3. Initialization Method of neural network parameters

There is the problem when using wavelet toolbox in experiments: Wavelet networks toolbox results is slightly better than BP neural network toolbox from a overall point of view. As predicted results are random numbers, their results are equal in a number of forecasting process. From this perspective, the toolbox is of little significance, and it fails to reflect the superiority of wavelet neural network. Therefore, this article uses the command rand $(state, Q)$ to obtain stable predictions, where Q is a self-defined parameter, which means network forecasting performance. The value of Q is to be determined by experiments, the command can reproduce a random number once produced, preferably by changing the initial Q value to get a better parameter. The evaluation of Q's value adopts average absolute percentage of forecasting errors. In the paper, Q is defined as the control parameters of the neural network prediction accuracy.

In MATLAB software, the function of command rand $(state, Q)$ is: reset generator to its original state, which means the state of generating random number. The effect is that: because of the difference of random number generated in each rand, in order to obtain the same state as the previous, it implements this function to generate the same random number.

Sequence of numbers generated by the rand command is decided by the inner generator of MATLAB. Generator, set to the same fixed state (Q), helps accomplish repetitive calculation; Q, set to be at different states, gets different results. It is the only calculation result that is the value of their application. However, no improvement of its statistical properties is made. As MATLAB restarted, at a fixed Q, rand generated the same sequence of numbers. Expressions and meanings are as follows:

Rand ('state', 0): reset the generator in the initial state;

Rand ('state', S): reset generator to S of its resolution states;

Rand ('state', J): reset generator to J of its first states for integer J

Experiments is carried out under the instruction of rand ('state', Q) which can predict the network's repeatability, comparison among different networks under the same conditions; and this prediction is very important.

Experiment is carried out by taking gas emission data with poor regularity as the object of the research. In order to extract geological features of gas Emission data, the study decomposes the original sequence through using wavelet packet's transformation, uses different wavelet neural networks to predict in sub-sequences, and then receives final prediction results for each sub-sequence prediction results obtained by wavelet packet reconstruction.

Gas Emission is a very complex geological parameter. It is affected by many factors, such as geological structure, coal seam thickness, coal structure, depth of burial and other natural factors, as well as other related mining technology. These factors themselves are random variables; they have mutual restraints among various factors and reinforce each other. Therefore, the gas emission rate is actually a multi-variable, time-varying, gray, highly nonlinear and complex dynamic system. It is often difficult to accurately predict. The size of Methane Emission not only reflects the degree of risk of different coal seams, but also becomes an important indicator of the level of safety technology which decides to develop a new well, the new mining area, the new face size, ventilation.

At present, the prediction of gas emission in BP network is more common, but the presence of BP neural network has the problems of long time convergence, easiness to fall into local minima.

3.4. Wavelet Packet's Transformation Theory

1) Wavelet Packet's Transformation

Let $\{ s_j; j \in z \}$ (Z is the set of integers) constitutes L2 (R) (R is a real number) on the orthogonal multi-resolutive analysis, its scaling function, the mother wavelet function are φ (t) and ψ (t) respectively. They satisfy the following two-scale equation.

$$\begin{cases} j(s) = \sqrt{2}\sum_{j\in z} l(k) j(2s-k) \\ y(s) = \sqrt{2}\sum_{j\in z} f(k) j(2s-k) \end{cases} \quad (14)$$

Wherein: coefficients h (k) and g (k) is the filter coefficients of a multi-resolutive analysis; $l(k) = -1_{ik}(1-k)$, that is, two coefficients have an orthogonal relationship.

The scale space v_j and wavelet subspace y_j are described as a new space m_j^n . Namely, $m_j^0 = v_j, m_j^1 = y_j, j \in z$, then the orthogonal decomposition in the Hilbert space $v_j+1 = v_j \oplus y_j$ can be denoted by m_j^n.

$$m_{j+1}^0 = m_j^0 \oplus m_j^1$$

Define a subspace m_j^n as the closed space of function $s_n(t)$, m_j^{2n} is the closure space of function $s_{2n}(t)$. For the case of the fixed scale, define recursive function:

$$\begin{cases} s_{2n}(t) = \sqrt{2}\sum_{j\in z} k(s) w_n(2t-h) \\ s_{2n+1}(t) = \sqrt{2}\sum_{j\in z} r(s) v_n(2t-h) \end{cases} \tag{15}$$

Wherein $n = 0,1,2,...,n$. When n = 0, the

$$\begin{cases} m_0(s) = \sqrt{2\sum_{j\in z} l(k) m_0(2s-k)} \\ m_1(s) = \sqrt{2\sum_{j\in z} p(k) m_0(2s-k)} \end{cases} \tag{16}$$

2) Wavelet Packet's Decomposition and Reconstruction

Let (s) ∈ (t), g_j^n can be expressed as:

$$g_j^n(s) = \sum_i tj - n_i^{j-n} m(2^j t - p)$$

Wavelet packet decomposition algorithm

Seeking by $g_1^{j+1,n}$ to find $\{g_1^{j,2n}, g_1^{j,2n+1}\}$

$$\begin{cases} f_i^{j,2n} = \sum_k a_{k-1i} f_k^{j+1,n} \\ f_i^{j,2n+1} = \sum_k b_{k-2i} f_k^{j+1,n} \end{cases} \tag{17}$$

Wavelet packet reconstruction algorithm:

Seeking by $g_i^{j,2n}, g_i^{j,2n+1}$ to find $g_i^{j+1,n}\}$

$$\begin{cases} f_i^{j,2n} = \sum_k a_{k-2i} f_k^{j+1,n} \\ f_i^{j,2n+1} = \sum_k b_{k-2i} f_k^{j+1,n} \end{cases} \tag{18}$$

3) Wavelet Packet's Analysis of Gas Emission Data

Using wavelet packet transformation to predict gas emission data series, it is the first step to decompose wavelet packet. The selection of decomposition level may be based on the minimum prediction error. The paper selects three layers. Use the small wavelet d_{b4} to decompose 8 frequency components from low to high, as shown in figure 2.

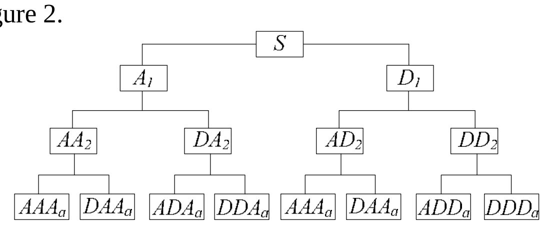

Figure 2. Wavelet packet's decomposition of three layers

In Figure 2, A represents a low frequency, D represents a high frequency, and the subscript indicates the number of layers of wavelet packet decomposition. It is can be seen that the wavelet packet's change not only further decomposition for each lower frequency section, but also further decomposition for the same high-frequency part.

Original sequence S can be expressed as:

$$s = AAA_3 + DAA_3 + ADA_3 + DDA_3 + \\ AAD_3 + DAD_3 + ADD_3 + DDD_3 \tag{19}$$

Original experimental data is: 429 day record data of Gas Emission. Original data is the signal S, using db_4 to transform wavelet packet for getting $s30, s31, ..., s37$ of 8 sub-sequences.

4) Wavelet Packet-wavelet neural network prediction model

Configuration method of prediction model: original data sequence of gas emission, the wavelet packet can be decomposed to $s30, s31, ..., s37$. Each sub-sequences using different wavelet neural network toolboxes to predict which adopts $WNN1, WNN2, ..., WNN8$. The respective predictions wavelet packet will be reconstructed to get the final prediction.

Currently there are many literature using wavelet packet analysis for fault diagnosis, but rarely used for prediction, mainly due to wavelet packet sequences, especially for high-frequency sequences generally difficult to adapt neural networks, and large forecasting workload. If the high-frequency component is uncertain, forecasting results will be greatly affected. In this paper, using wavelet neural network to predict, it can effectively extract the characteristics of high frequency sub-sequence information through wavelet neural network toolbox, so that training speed is greatly improved, and has strong robustness. The timing sequence prediction of subspace: time series forecasting estimates future observations according to the use of the past and the present value. Let the sequence of one-dimensional {x (i)} ($i = 1,2,...n$) with n before neighboring m of the time values (m <n), the first $n+1$ predictive value of time ($n+1$), the future and past values of a function to determine the existence of a relation:

$$\hat{x}(n+1) = f\left[x(k+1), x(k+2), ..., x(k+m)\right]$$

Similarly, with the numerical prediction of $x(k+2), x(k+3), ...x(k+m+1)$, the corresponding values of $\hat{x}(n+2)$ in the time of $n+2$, carry on the process of recursion. If the sequence {x (i)} can be divided into h group, then can get a sample set of a neural network's training ($\hat{x}h, f_h$). According to predicting effects taking $m = 7$, namely input layer neurons of wavelet neural network is 7, the output of a neuron is 1. Hidden layer neurons is $q = 2n+1, n = 7$, n is the number of network input variable, so q = 15, that is, the network structure is Ⅴ. Training error precision is set to 0.0001; normalization uses command norm; anti-normalization use norm. Such data processing method avoids data appearing 0 and 1, in order to improve the model prediction performance and input normalized data into network for training.

4. Experimental Results

As for subsequence S30, in the 429 series data, using the data before seven days to predict he eighth day, and carry on recursion. Afterwards, it constitutes 422 sample set, take the first 300 samples to train neural network, the left 122 samples for test.

WNN1's designation: commence on the experiments in the environment of MATLAB R2006b. Network structure: 7-15-1; training precision is set to 0.000 1; the wavelet network toolbox function of the network hidden layer takes " *wavenet tool* ", selects output layer " *purelin* " function, uses the " *trainlm* " algorithm for training. Initial conditions in the experiments is set as Q = 9, which uses the command rand ('state', 9) to initialize wavelet network parameters. The predictions of S30 are shown in Figure 3; details can be seen in Table 1. Table 1 also gives the prediction results of S31 ~ S37 and the predicted results of BP network under the same conditions.

Table1. Prediction Results of Wavelet Packet's Decomposed Subsequence

Wavelet Packet Sequence	Norm	WNN too Case	BP too Case	Selection of Q Value, Accuracy Set
S30	Prediction accuracy%	0.3742	3.2321	7-15-1
	Maximum precision%	2.1231	8.3201	Q=9
	The training and prediction time	84.021	1.3786	0.0001
S31	Prediction accuracy%	83.0963	455.02	7-15-1
	Maximum precision%	2309	25218	Q=5
	The training and prediction time	1.4092	1.2871	0.0001
S32	Prediction accuracy%	32.053	257.03	7-15-1
	Maximum precision%	318.01	3502.01	Q=19
	The training and prediction time	1.335	1.2901	0.0001
S33	Prediction accuracy%	87.031	277.04	7-15-1
	Maximum precision%	1678.02	4267.03	Q=26
	The training and prediction time	1.4123	1.2536	0.0001
S34	Prediction accuracy%	19.386	415.03	7-15-1
	Maximum precision%	320.09	9258.01	Q=30
	The training and prediction time	1.294	1.467	0.0001
S35	Prediction accuracy%	56.632	230.032	7-15-1
	Maximum precision%	760.432	3564.021	Q=11
	The training and prediction time	1.254	1.223	0.0001
S36	Prediction accuracy%	45.076	469	7-15-1
	Maximum precision%	1530.87	24267	Q=13
	The training and prediction time	1.253	1.478	0.0001
S37	Prediction accuracy%	95.021	138.02	7-15-1
	Maximum precision%	1474.04	3781.01	Q=2
	The training and prediction time	1.3480	1.2503	0.0001
reconsitution	Prediction accuracy%	3.1792	15.6351	

Similar to S30, the sub-sequences of S31 ~ S37 have the same treatment. Q value determination method: range generally from $0 \le Q \le 20$ can meets the requirements, Q value, as a integer,can pass the test individually, or predict the accuracy of performance indicators for programming optimization. The forecasting results are demonstrated in Figure 4 to 10, each of FIG is the comparison between actual and predicted values.

It can be seen from the results, the subsequence, 0.3740%, prediction accuracy of S30, is the highest; the maximum prediction accuracy of S30 is 2.111 2%; the prediction results are reproducible. Figure 3 illustrates that the wavelet network toolbox has a strong ability for fitting with the generalization, and the remaining sequences graphics show WNN toolbox method for high frequency signal is also very strong ability to adapt, and is much better than the BP network. In S30, the present method prediction accuracy is important; it is the predominant component of signal reconstruction.

When the sub-series are predicted, you can reconstruct wavelet packet, the results are shown in Figure 11, which is the contrast of the total predicted actual and predicted values. Final prediction accuracy is 3.1454%; the maximum prediction accuracy is -14.3672%; reconstruction time required is $0.8742s$. Overall prediction accuracy is 15.6315% after BP network's reconstruction; the maximum prediction accuracy is 40.0132%.

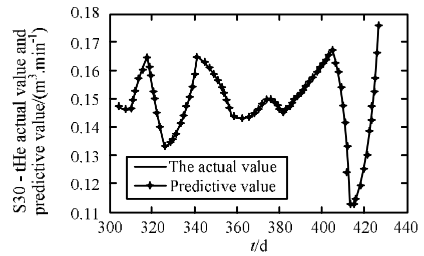

Figure 3. S30 Predicting Results

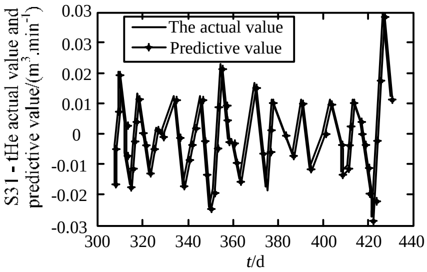

Figure 4. S31 Predicting Results

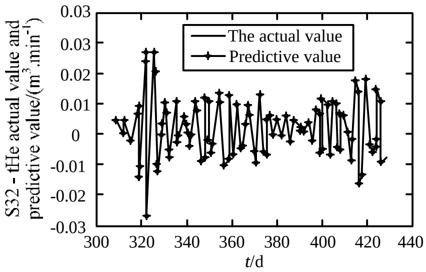

Figure 5. S32 Predicting Results

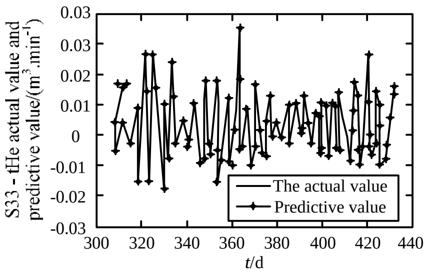

Figure 6. S33 Predicting Results

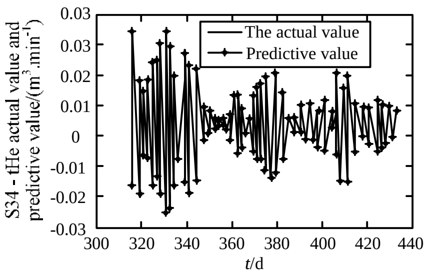

Figure 7. S34 Predicting Sesults

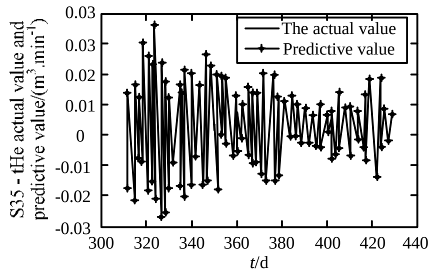

Figure 8. S35 Predicting Results

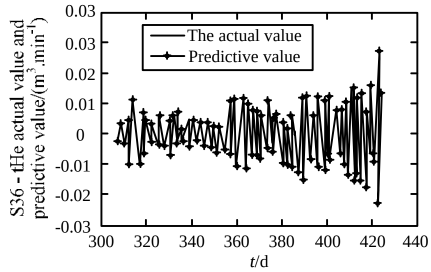

Figure 9. S36 Predicting Results

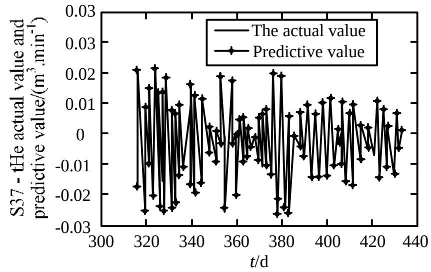

Figure 10. S37 Predicting Results

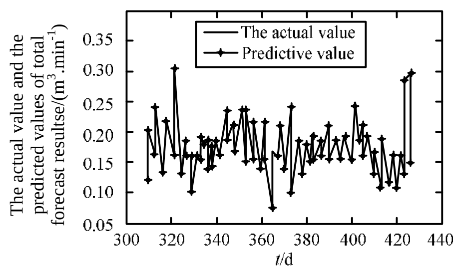

Figure 11. Predicting Results after Reconstruction

In the last column of Table 1, it gives the structure of the network, Q value (BP, WNN) and the setting of training prediction accuracy.

Some explanations on the Table 1 are shown below:

a) The results of the prediction accuracy in Table 1 can be repeated in experiments; the training time can not be repeated. However, its volatility is so small that its influence can be neglected.

b) BP network prediction of Table 1 is obtained under the same experimental conditions in connection with WNN , therefore the predicted result is determined. Q value is the optimal values for WNN, but it is not necessarily optimal for the BP.

c) In Table 1, it seems difficult to understand that some high-frequency sub-sequence prediction accuracy is poor. But it does not mean that overall prediction accuracy is low. The higher the accuracy is, their overall prediction accuracy can be improved.

d) Neural networks are data-driven model, a greater dependence on the data. Similarly, the prediction results of the wavelet network have a great relationship with the quality of data sample; different data may get different predictions.

Training and prediction time's calculation: training and forecasting of 8 Sub-sequence, is $10.723s$ totally; reconstructed forecasting time is 0.8731s; consuming time is $11.58734s$ totally, which does not contain operating time. It is much higher than the programmed BP's, WNN'S neural network speed, and it is sometimes difficult to train which are based on programming BP, WNN local minima due to other reasons .

It is accomplished on the experiments of S30 by using WNN1 programming prediction. Programming of WNN1 network structure: 7-15-1; learning rate: 0.3; momentum factor: 0.02. When the training accuracy is set to 0.0013 , the absolute average training accuracy is 1.5078%; the maximum precision is 6.4080%, time required for training and prediction is $12.3388s$. All sequences are difficult to train when training accuracy is set to 0.0001. When training accuracy of S31 ~ S37 is set to 0.06 , it can be achieved to training accuracy requirements, but the prediction accuracy is poor, indicating that the programming of the WNN has a larger difference with wavelet network toolbox's performance.

Conclusion

Randomness is a natural characteristics of neural network's forecasting methods. Based on the research of wavelet neural network toolbox, the initialization commands of Q-value, makes random transformation of predicting models affirmative. It is an expansion of the prediction theory for wavelet neural network. Through gas emission wavelet packet-Wavelet network forecasting examples, it shows that the method has not only both wavelet feature extraction capabilities, but also a series of advantages of BP neural network toolbox, such as higher speed of training, easier operation, applicability to large quantities of data for training and processing, data adaptability and robustness with a flexible, practical features. By controlling parameter Q to get the best predictive value, it is more convenient than optimized algorithms such as genetic algorithms, particle swarm algorithm. This method of wavelet neural network is of practical significance to promote the application of wavelet neural network.

References

[1] J. He, Y. Geng and K. Pahlavan, Modeling Indoor TOA Ranging Error for Body Mounted Sensors, 2012 IEEE 23nd International Symposium on Personal Indoor and Mobile Radio Communications (PIMRC), Sydney, Australia Sep. 2012 (page 682-686)

[2] S. Li, Y. Geng, J. He, K. Pahlavan,Analysis of Three-dimensional Maximum Likelihood Algorithm for Capsule Endoscopy Localization, 2012 5th International Conference on Biomedical Engineering and Informatics (BMEI), Chongqing, China Oct. 2012 (page 721-725)

[3] D. Xu, Z. Y. Feng, Y. Z. Li, et al. Fair Channel allocation and power control for uplink and downlink cognitive radio networks. IEEE., Workshop on mobile computing and emerging communication networks, 2011:591-596

[4] W. Q. Yao, Y. Wang, T. Wang. Joint optimization for downlink resource allocation in cognitive radio cellular networks. IEEE., 8th Annual IEEE consumer communications and networking conference, 2011:664-668

[5] S. H. Tang, M. C. Chen, Y. S. Sun, et al. A spectral efficient and fair user-centric spectrum allocation approach for downlink transmissions.IEEE., Globecom.,2011:1-6

[6] D. L. Sun, X. N. Zhu, Z. M. Zeng, et al. Downlink power control in cognitive femtocell networks. IEEE., International conference on wireless communications and signal processing, 2011: 1-5

[7] Muhammad J. Mirza, Nadeem Anjum.Association of Moving Objects across Visual Sensor Networks. Journal of Multimedia, Vol 7, No 1 (2012), 2-8

[8] Kasman Suhairi, Ford Lumban Gaol, The Measurement of Optimization Performance of Managed Service Division with ITIL Framework using Statistical Process Control. Journal of Networks, Vol 8, No 3 (2013), 518-529

[9] Guang Yan, Zhu Yue-Fei, Gu Chun-Xiang, Fei Jin-long, He Xin-Zheng, A Framework for Automated Security Proof and its Application to OAEP. Journal of Networks, Vol 8, No 3 (2013), 552-558

[10] R. Berangi, S. Saleem, M. Faulkner, et al. TDD cognitive radio femtocell network (CRFN) operation in FDD downlink spectrum. IEEE, 22nd International Symposium on Personal, Indoor and Mobile Radio Communications, 2011: 482-486

[11] Pearson S. Taking account of privacy when designing cloud computing services. In CLOUD '09: Proceedings of the 2009 ICSE workshop on software engineering challenges of cloud computing, IEEE Computer Society, Washington, DC, USA, 2009. pp. 44–52.

[12] http://dx.doi.org/10.1109/CLOUD.2009.5071532

[13] Mondol, J.-A.M. Cloud security solutions using FPGA. In Communications, Computers and Signal Processing (PacRim), 2011 IEEE Pacific Rim Conference on, 2011, pp. 747-752.

[14] Wang Lina, Gao Hanjn, Liuwei, Peng yang. Detection and management of virtual machine monitor. Research and development process of Computer, 2011, pp: 1534-1541.

[15] Yinghua Xue, Hongpeng Liu, Intelligent Storage and Retrieval Systems Based on RFID and Vision in Automated Warehouse. Journal of Networks, Vol. 7, No. 2 (2012), pp: 365-369

[16] Huan Zhao, Kai Zhao, He Liu, Fei Yu, Improved MFCC Feature Extraction Combining Symmetric ICA Algorithm for Robust

Speech Recognition, Journal of multimedia, Vol. 7, No. 1, 2012. pp: 74-81

[17] LI La yuan, LI Chun lin, "A multicast routing protocol with multiple QoS constraints," Journal of Software, vol. 15, No. 2, 2004, pp. 286-291.

[18] Q. He, C. Han, "Satellite Constellation Design with Adaptively Continuous Ant System Algorithm," Chinese Journal of Aeronautics. vol. 6, No. 4, 2007, pp. 297-303. http://dx.doi.org/10.1016/S1000-9361(07)60047-8

[19] L. Y. Ren, "Study on Scheduling Optimization in Construction Project of Lagerstroemia Hope City," Xi'an University of architecture & techinology. vol. 6, No. 2, 2011, pp. 12-17.

[20] Y. Yona, M. Feder, "Efficient parametric decoder of low-density lattice codes," IEEE International Symposium on Information Theory: June 28-July 3, 2009, Seoul, Korea. New York, NY, USA: IEEE, 2009, 8: 744-748.

[21] B. Kurkoski, J. Dauwels, "Message-passing decoding of lattices using Gaussian mixtures," IEEE International Symposium on Information Theory: June 6-11, 2008, Toronto, Canada. New York, NY, USA: IEEE, 2008, 8: 2489-2493.

[22] BICKSON D, IHLER A, AVISSAR H et al. A low-density lattice decoder via non-parametric belief propagation. Forty-Seventh Annual Allerton Conference on Communication, Control and Computing: Sep 30-Oct 2, 2009, Illinois, USA. Monticello, IL, USA: IEEE, 2010, 1: 439-446.