International Journal of Smart Home
Vol. 10, No. 11, (2016), pp. 133-146
http://dx.doi.org/10.14257/ijsh.2016.10.11.12

Research of Short-Term Load Forecasting Using DWT and LSSVM Optimized by QDE

Wei Sun[1] and Jingyi Sun[2*]

[1]*Department of Business Administration, North China Electric Power University, Baoding, China*
[2]*Department of Business Administration, North China Electric Power University, 689 Huadian Road, Baoding 071000, China. Tel:+86 15031282075*
[1]*bdsunwei@163.com,*[2]*sunjingyi0224@126.com*

Abstract

To evaluate short-term power load properly and efficiently, this paper proposes a modified DWT-QDE-LSSVM (Discrete wavelet transform (DWT) and least squares support vector machine (LSSVM) optimized by quantum differential evolution (QDE)) model combined with input selected. The load data series of the previous days are first decomposed into an approximation component and a detail component. Then LSSVM is built to model the approximation component and QDE algorithm is applied to overcome the problems faced by LSSVM in selecting parameters. In order to raise forecasting accuracy, this paper proposes the refinement of related factors. The empirical results show that the proposed DWT-QDE-LSSVM model is feasible and can satisfy the short-term load forecasting requirements in China.

Keywords: *Short-term load forecasting; Quantum differential evolution; Wavelet transform; Least squares support vector machine; Refinement*

1. Introduction

During the Twelfth Five-year Plan period, with large-scale pilots in smart grid and deepening reform in electricity market, the smart grid faces lots of unprecedented challenges and opportunities [1-2]. Therefore, improving short-term load forecasting method and the corresponding forecasting accuracy is very important for power demand analysis, such as power dispatching [3], unit commitment [4], operation and trading activities of market participants [5] and so on.

Currently, the methods in load forecasting are divided into classical mathematical statistical methods and artificial intelligence models. Most of these methods, including vector autoregressive (VAR) model [6] and ARMA model [7-8] are based on time series analysis. Pappas. S. S. *et. al.,* [7] employed ARMA model to predict the short-term load. Time series smoothness prediction methods are criticized by researchers for their weakness of non-linear fitting capability. With the development of the electricity market in the world, high self-learning ability prediction methods are presented for load forecasting, such as neural network prediction technology [9] and support vector machine (SVM) [10-11]. Among them, the application of artificial neural network is the most extensive, Beccali *et. al.,* [9] came up with a combined approach based on unsupervised and supervised neural networks and confirmed its validity to forecast the electric energy demand of a suburban area with a prediction time of 24h. Although load forecasting is the most suitable areas of ANN applications in power system, the errors between input data and actual value lead to defects in ANN prediction accuracy [10]. According to the characteristics of different regions, various parameters and ANN structure should be

*Corresponding Author

ISSN: 1975-4094 IJSH

selected according to its own variation of load and weather. This increases the difficulty of model promotion [11].

Unlike ANN which uses the empirical risk minimization principle to minimize the generalization errors, SVM exploits the structural risk minimization principle to convert the solution process into a convex quadratic programming problem. This overcomes some shortcomings in neural network and achieves a good performance in the practical load forecasting [12]. In paper [13], Sousa et al. successfully proposed SVM which was optimized by simulated annealing to forecast the short-term load. However, the problem of hyperplane parameter selection in SVM leads to a large solving scale [14]. In order to solve this, J.A.K. Suykens and J. Vandewalle proposed LSSVM as a classifier in 1999. Unlike the inequality constraints which were introduced by SVM, LSSVM proposes equality constrains in formulation. Shayeghi, H *et. al.*, [15] proved that the LSSVM based model has good potential for simultaneous forecasting of electricity price and load. However, the selection of regularization parameters and kernel parameters greatly affects the performance of LSSVM. Differential evolution (DE) algorithm [16], particle swarm optimization (PSO) [17] and other optimization models are used in order to select appropriate parameters for power load forecasting. Aiming at solving the problems of local optimum and low efficiency faced by the above optimization models, this paper comes up with QDE to improve global optimization capability.

When using models for short-term load forecasting, the original data is usually directly used in prediction models. However, due to the chaotic nature of short-term load data, describing the movement trend of short-term load data and accurately predicting it become difficult. In order to solve this problem, the application of wavelet transform (WT) in eliminating the irregular fluctuation of the load data has achieved much attention in recently years [18]. Bahrami *et. al.*, [18] developed wavelet transform to eliminate the high frequency components of the previous days load data. Literature survey shows that the combination of WT and other intelligent algorithms achieves good results.

In this paper, a hybrid method composed of the DWT and QDE-LSSVM is proposed for short-term load forecasting. In this method, not only the highest temperature, the lowest temperature, the mean relative humidity, the mean value of the wind speed and the load data of the previous days, but also the refinement of day type are considered as the inputs for QDE-LSSVM. In this regard, the wavelet transform is used to eliminate the high frequency components of the previous days load data. In addition, this paper comes up with QDE algorithm to select and automatically adjust appropriate parameters of LSSVM model. Finally, the proposed model is employed to forecast the load of Yangquan city in China.

This paper is organized as follows: Section 2 shows the brief description of DWT, LSSVM and QDE; Section 3 presents the frame work of the proposed technique; Section 4 analyzes an experiment study, while Section 5 concludes this paper.

2. Methodology

2.1. Discrete Wavelet Transform

As an effective method for signal processing, wavelet transform can be divided into two classifications: DWT and continuous wavelet transform (CWT). Compared with other WTs, DWT, as a kind of WT, whose wavelets are discretely sampled, can capture both frequency and location information in temporal resolution. Thus, DWT has a key advantage over Fourier transforms. In this paper, DWT is used in data filtering stage.

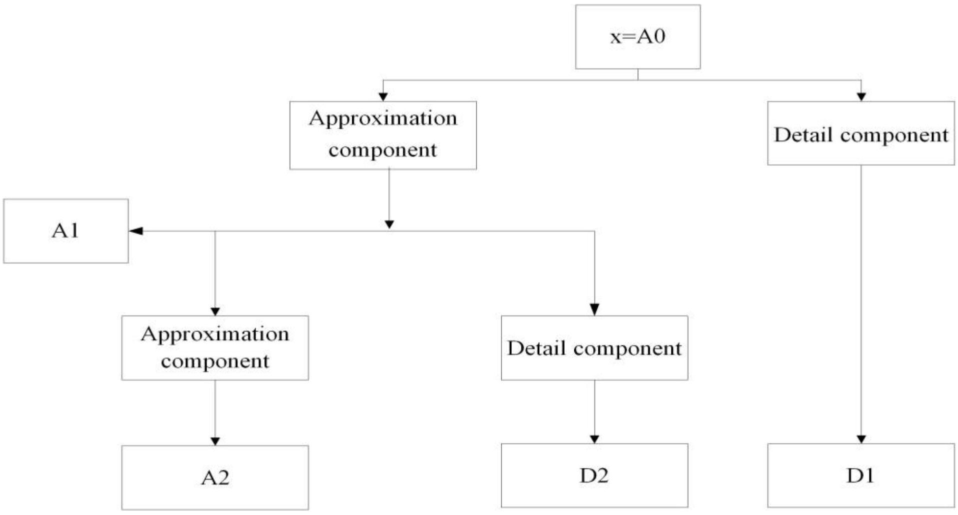

Figure 1. Multi-Resolution Decomposition for Two-Level DWT

This paper uses a three-level DWT to decompose the original load time-series into approximation and detail components as Figure 1, shows. The equation applied for the DWT of a discrete time signal $x(k)$ is listed as follows:

$$D^{c}_{m,n} = \sum_{k} x(k)\psi_{m,n}(k) \tag{1}$$

where $D^{c}_{m,n}$ and $\psi_{m,n}(k)$ are the detailed coefficient and function at location n and scale m. $\psi_{m,n}(k)$ with $m,n=0$ is the detailed and translated form of the mother wavelet $\psi(k)=\psi_{0,0}(k)$. Scaling functions for DWT are given by:

$$\varphi_{m,n}(k) = \left(2^{-\frac{m}{2}}\right)\varphi\left(2^{-m}k-n\right) \tag{2}$$

where $\varphi(k)$ is the scaling function with $m,n=0$.

The approximate coefficient at scale m and location n is:

$$A^{c}_{m,n} = \sum_{k} x(k)\varphi_{m,n}(k) \tag{3}$$

The approximation A_m and detail D_j of the signal at scales m and j are calculated by:

$$A_m = \sum_{n=-\infty}^{n=+\infty} A^{c}_{m,n}\varphi_{m,n} \tag{4}$$

$$D_j = \sum_{n=-\infty}^{n=+\infty} D^{c}_{j,n}\psi_{j,n} \tag{5}$$

For a DWT with M-level decomposition, the sum of the approximation and details up to scale M provides the original signal x as

$$x = A_M + \sum_{j=1}^{M} D_j \tag{6}$$

The approximate and detailed coefficients at any arbitrary scale m are used to calculate the corresponding coefficients at the next scale m+1 by:

$$A^c_{m+1,n} = \sum_i l_i A^c_{m,2n+i} = \sum_i l_{i-2n} A^c_{m,i} \quad (7)$$

$$D^c_{m+1,n} = \sum_i h_i A^c_{m,2n+i} = \sum_i h_{i-2n} A^c_{m,i} \quad (8)$$

where l_i and h_i are the coefficients of the approximation and detail components used for the wavelet decomposition.

2.2. Least Squares Support Vector Machine

LSSVM is a novel SVM method which puts up with by Suykens J.A.K to solve the problems of model decomposition and function estimation. LSSVM adopts least squares linear system as a loss function which replaces quadratic programming in SVM. This method simplifies the computational complexity and improves the running speed. Similar to ANN and other intelligent algorithms, the performance of LSSVM seriously depends on the input and the parameters.

2.3. Quantum Differential Evolution

The detailed working steps of QDE are presented as follows:

1) Population initialization

$|0\rangle$ and $|1\rangle$ are used to represent two basic states of microscopic particle. The quantum bit represents as:

$$|\varphi\rangle = \alpha|0\rangle + \beta|1\rangle \quad (9)$$

among it, α, β are the probability amplitudes which correspond to the quantum bits, and $|\alpha|^2 + |\beta|^2 = 1$. Thus, the quantum bits can also be expressed as $\left[\cos(\theta), \sin(\theta)\right]^T$.

The quantum bit $Q_{i,j}$ is defined as:

$$Q_{i,j} = \begin{bmatrix} \cos(\theta_{i,1})\cos(\theta_{i,2})\cdots\cos(\theta_{i,j}) \\ \sin(\theta_{i,1})\ \sin(\theta_{i,1})\cdots\sin(\theta_{i,j}) \end{bmatrix} \quad (10)$$

where $\theta = 2\pi rand$, $rand \in [0,1]$, $i \in \{1,2,\ldots,\mathrm{N}\}$, $j \in \{1,2,\ldots,D\}$. D represents the dimension of the problem, N represents the population size.

After generating initial population Q, this paper supposes P_m and C as mutation probability and contraction factor, respectively.

2) The solution space transformation and fitness calculation

The quantum chromosomes variables should be mapped from the unit space $I = \left[-1,1\right]$ to the optimization problem solution space, so that each quantum chromosome variables correspond to each optimization variables of optimization problems. Suppose the definition-domain of $X_{i,j}$, the solution space variable of optimization problem is $[a_j, \mathrm{b}_j]$. Thus the corresponding solution space variables are shown as follows:

$$\begin{bmatrix} X^0_{i,j} \\ X^1_{i,j} \end{bmatrix} = \frac{1}{2}\begin{bmatrix} 1+\cos(\theta_{i,j}) & 1-\cos(\theta_{i,j}) \\ 1+\sin(\theta_{i,j}) & 1-\sin(\theta_{i,j}) \end{bmatrix}\begin{bmatrix} b_j \\ a_i \end{bmatrix} \quad (11)$$

Each quantum chromosome Q corresponds to a solution X of optimization problems, and the fitness function $f(\mathrm{X})$ can be calculated and updated based on the greed principle for global optimal solution.

3) The rotation angle adjustments of quantum own position:

For each chromosome of the population, calculate the rotation angle of quantum own position according to the following formulas:

$$\Delta Q_{ij}^{g} = \theta_{\min} + fit(\theta_{\max} - \theta_{\min}) rand_j \exp(\frac{gen}{\max gen}) \tag{12}$$

$$fit = \frac{fit_{gBest} - fit_i}{fit_{gBest}} \tag{13}$$

in which, $\theta_{\min}$, $\theta_{\max}$ are the min 0.001π and max 0.05π of the interval $\Delta\theta$, respectively; fit_{gBest} represents the fitness value of the global optimum individual; fit_i is the fitness value of current individual; gen represents the former iterations of the population while maxgen represents the maximum times of iterations which are limited by population.

4) Quantum mutation

The update of quantum chromosome in QDE algorithm draws lessons from the DE algorithm [19]. The quantum chromosome is randomly selected from the quantum population, and qubit phase is proposed as basis vector with two different qubit phases used as differential vectors.

$$v_{i,j}^{g+1} = \theta_{r1,j}^{g} + C * rand * (\theta_{r2,j}^{g} - \theta_{r3,j}^{g}) \tag{14}$$

among it, $r_1, r_2, r_3 \in \{1, 2, \ldots, \mathrm{N}\}$, and $r_1 \neq r_2 \neq r_3$. The scaling factor C is a random number between [0, 1].

5) Quantum crossover

In order to generate new individuals, quantum individual variation is combined with pre-determined parent individuals by a certain principle.

$$u_{i,j}^{g+1} = \begin{cases} v_{i,j}^{g+1}, if\,(rand_j \leq C_r, j = j_{rand}) \\ x_{i,j}^{g}, otherwise \end{cases} \tag{15}$$

in which, $j \in \{1, 2, \ldots, D\}$. D is dimension of the problem, j_{rand} represents a integer randomly selected in $\{1, 2, \ldots, D\}$, and crossover probability $C_r \in [0,1]$.

6) Quantum selection

To make better fitness individuals into the next generation, one to one greedy algorithm is used to select the operator.

$$\theta_{i,j}^{g+1} = \begin{cases} u_{i,j}^{g+1}, if\,(f(u_{i,j}^{g+1}) \leq f(\theta_{i,j}^{g+1})) \\ \theta_{i,j}^{gt}, otherwise \end{cases} \tag{16}$$

$$\theta_{i,j}^{gt} = \theta_{i,j}^{g} + \Delta\theta_{i,j}^{g} \tag{17}$$

where f represents the fitness function.

After updating, new qubits can be expressed as:

$$\theta_{i,j}^{g+1} = \begin{bmatrix} \cos(\theta_{i,1}^{g+1}) \cdots & \cos(\theta_{i,\mathrm{D}}^{g+1}) \\ \cos(\theta_{i,1}^{g+1}) \cdots & \sin(\theta_{i,\mathrm{D}}^{g+1}) \end{bmatrix} \tag{18}$$

7) The comparison of results

If the algorithm reaches iterations or meets the convergence criteria, the best result can be recorded, otherwise returns to Step 2.

3. Approaches of DWT-QDE-LSSVM

In this section, the short-term load forecasting model incorporating DWT, QDE and LSSVM is constructed as Figure 2, shows.

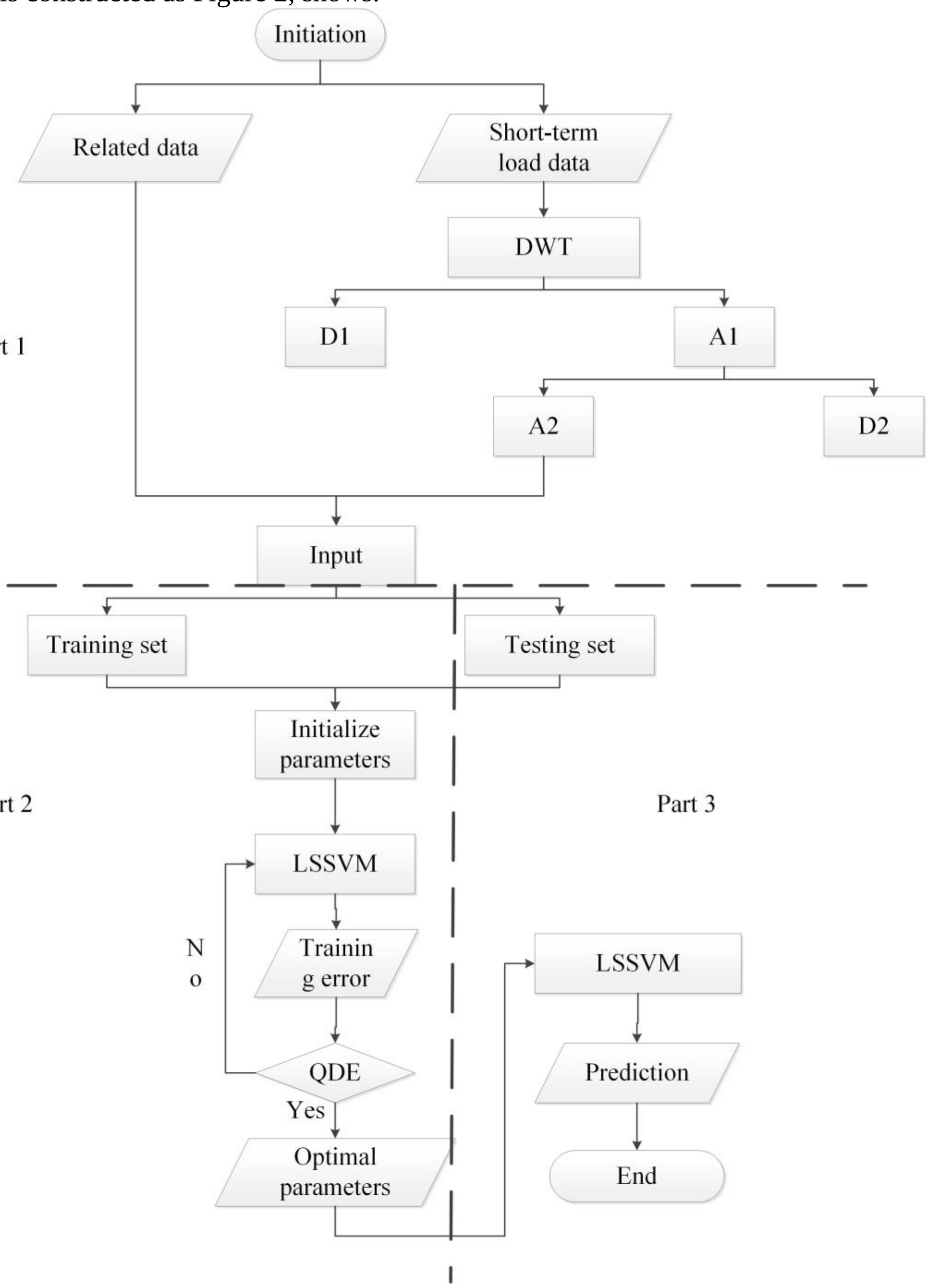

Figure 2. Flowchart of DWT-QDE-LSSVM Algorithm

The modeling flowchart contains three parts. Prior to constructing the model, the short-term load data S are decomposed into an approximation signal A and a detail signal D by DWT. Researchers select the day type, the highest temperature, the lowest temperature, the mean relative humidity and the mean value of the wind speed as related factors. According to the load curve, this paper carries on a more careful classification to day types based on different holidays, the specific classification is described later.

After two-layer DWT decomposition and the selection of input, the original short-term load data are divided into the training set and the test set.

In the second part, LSSVM is proposed to model the sub-series in the first part, so that the tendencies of the sub-series can be predicted. The fitness function of QDE is defined as:

$$\begin{cases} f(\gamma,\sigma^2)=\dfrac{1}{N}\sum_{i=1}^{N}(y_i-y_i^{'}) \\ s.t. \gamma\in[\gamma_{\min},\gamma_{\max}], \sigma^2\in[\sigma^2{}_{\min},\sigma^2{}_{\max}] \end{cases} \tag{19}$$

where y_i is the *i-th* input value of known samples, $y_i^{'}$ is the *i-th* output value predicted by corresponding model, N is the total number of samples while γ and σ^2 represent regularization parameters and kernel parameters, respectively.

In the final part, LSSVM with optimal parameters which obtain by QDE is used to forecast the short-term load data.

Our approach is elaborately presented in Algorithm 1.

Algorithm 1 The proposed DWT-QDE-LSSVM forecasting algorithm

Assumption: $Y=Y^{(Lin)}$ (linear part)+ $Y^{(NLin)}$ (nonlinear part)

Inputs: The training dataset $Y_{tr}=\left[y_1,y_2,\ldots,y_{N_{tr}}\right]^T$ and size N_{tr} of the testing dataset

Outputs: The combined forecast vector $\hat{Y}=\left[\hat{y}_{N_{tr}+1},\hat{y}_{N_{tr}+2},\ldots,\hat{y}_{N_{tr}+N}\right]^T$

Steps:

1. Apply DWT to Y_{tr} to decompose it as follows: $[A,D]=DWT(Y_{tr},'filter')$
2. Initialize population $N=10$; mutation factor $F=0.05$; crossover rate $C=0.5$
3. Set generation $G=0$ and random generate the populations
4. Set quantum coding on input data as Formula (10)
5. Do transform parameters as Formula (11)
6. Use the populations in LSSVM to forecast A, get the load forecasting results and get fitness function value $f(X)$
7. If G <max generation number
8. Do mutation operation Formula (14)
9. Do crossover operation Formula (15)
10. Selection operation in order to generate offspring, set $G=G+1$
11. Return to Step 6
12. Else if
13. End the QDE-LSSVM operation

End

4. A Numeric Example and Results

4.1. Data Preprocessing

The 24h short-term load forecasting has been made on the power system of Yangquan city in China from March 1, 2013 to June 10, 2013. Figure 3, shows the power load of 2280 samples, ranging from around 650 MW to 950 MW. From Figure 3, none apparent regularity of power load can be obtained.

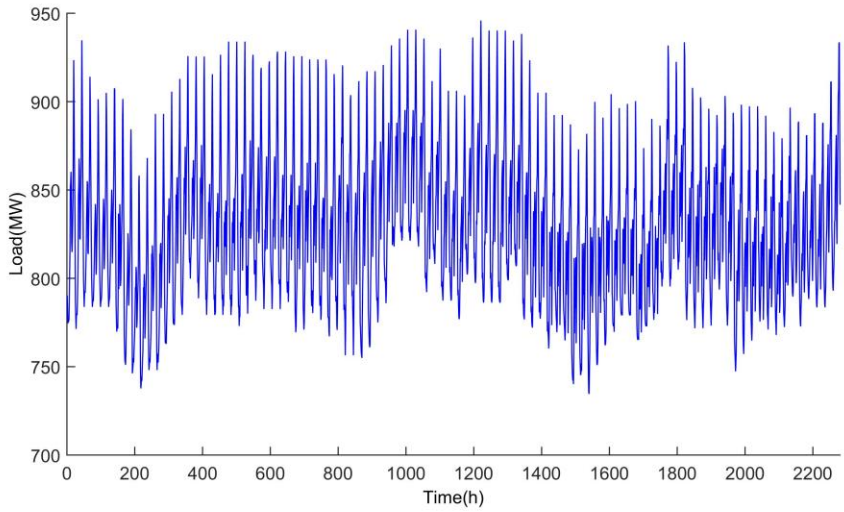

Figure 3. The Original Short-Term Load Data from March 1, 2013 to June 10, 2013 of Yangquan City

Through the two-level decomposition by DWT, the power load curve after noise reduction is shown in Figure 4. By contrast with Figure 3, the power load signal is changed into smooth one which has smaller burr and has better connectedness after the DWT process.

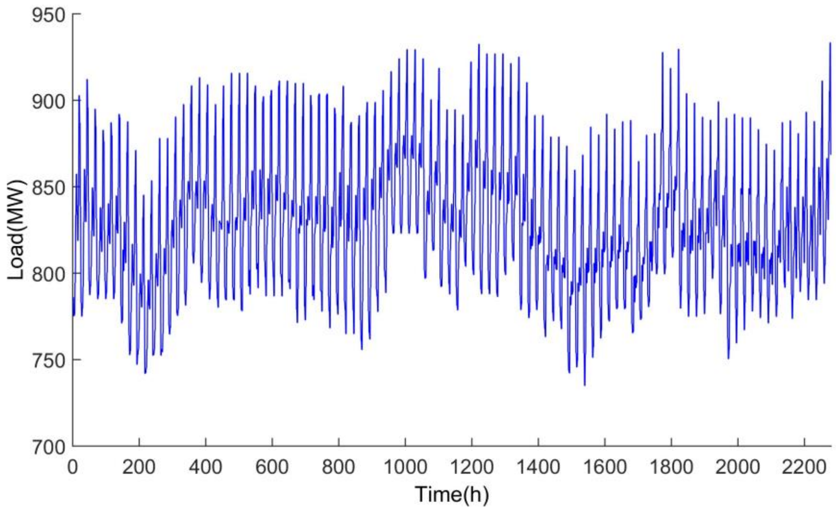

Figure 4. The Short-Term Load Data After DWT Decomposition

4.2. Selection of Input

Currently, the input variable selection is one of the most important parts of every forecasting method. Therefore, this paper presents the refinement of the argument by detailing the classification and selection of the input to improve the accuracy of the forecasting.

Nowadays, the common used classification of day types, which is set as $\{0, 0.8, 1\}$, is based on weekdays, weekends and holidays. As is shown in Figure 5， the time series of load displays a clear daily and weekly seasonality. The weekly seasonality effects of the days of the week and holidays. Most papers claim that Monday, Tuesdays, Wednesdays, Thursdays and Fridays can be modeled as a single type of day. After the study of the short-term load in Yangquan, researchers find that, in this region, each daily load has its own unique characteristics. Figure 5, shows Yangquan weekly load curve, from it we can know that the evening load in Friday and Saturday is higher than other days in the same period. Since the large amount of data we have, researchers prefer to model each day as a dummy variable. Dummies for holidays and days around holidays are also considered. Table 1, gives a summary of the variables used， and the $\{0, 0.8, 1\}$ day type classification is compared with the proposed classification later.

Table 1. Types of Days Used in the DWT-QDE-LSSVM Model

Code	Description
1	Sunday
2	Monday
3	Tuesday
4	Wednesday
5	Thursday
6	Friday
7	Saturday
8	Holiday (official or religious)
9	Working day before a holiday
10	Working day after a holiday

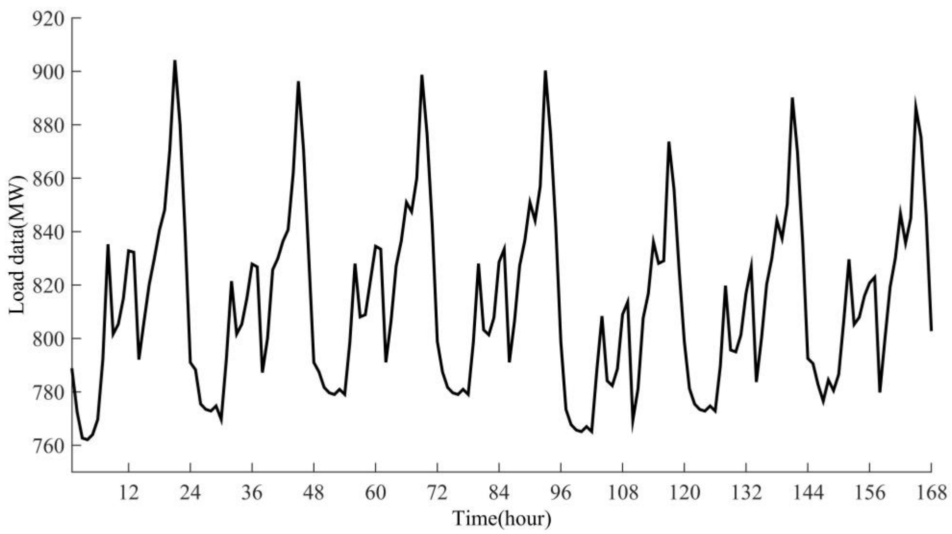

Figure 5. Weekly Load Values in Yangquan

In this region, grid load is greatly influenced by weather and seasonal temperature differences. Thus, the related factors in this paper include: days type; daily temperature (including daily highest temperature and daily lowest temperature); daily precipitation and daily wind velocity. 2160 data are selected as training samples from May 1, 2013 to June 5, 2013, and the data from June 6, 2013 to June 10, 2013 are selected as test samples. Meanwhile, in order to reflect the different load conditions, 24 hours test results are obtained.

4.3. Statistic Measure to Determine the Accuracy of the Forecast

This paper assesses the prediction model accurately by using appropriate indicators. The paper applies mean absolute percent error MAPE, mean absolute error MAE and daily peak error DPE to measure the feasibility of the model. These measures are defined as follows:

$$MAPE = \frac{1}{N}\sum_{i=1}^{N} \left| y_i^{'} - y_i \right| \Big/ y_i \ *100\% \tag{20}$$

$$MAE = \frac{1}{N}\sum_{i=1}^{N} \left| y_i^{'} - y_i \right| \tag{21}$$

$$DPE = \max_{1<i<24}\left(\left| y_i^{'} - y_i \right| \Big/ y_i \right) \tag{22}$$

in which, $y_i^{'}$ and y_i denote output value and the actual value of the test sample, respectively.

4.4. DWT-QDE-LSSVM Results Analysis

As discussed previously, the performance of LSSVM modeling lies on its parameters. QDE is developed to tune regularization parameter γ and kernel parameter σ^2 of LSSVM by minimizing errors generated in the training set and test set. The regularization parameter γ and kernel parameter σ^2 of LSSVM optimized by QDE are 152.173 and 0.046, respectively. The main parameters of QDE and BPNN are listed in Table 2.

Table 2. Parameters of QDE and BPNN

QDE Parameters	Value	BPNN Parameters	Value
Population	10	Node in hidden layer	2
Mutation factor	0.05	Maximum number of convergence	1000
Crossover factor	0.5	Learning rate	0.5
		Error	0.004

The auto correlation and partial correlation of the forecasting residual series by DWT-QDE-LSSVM are considered to exam the power load forecasting. The results are shown in Figure 6. It is obvious to find that there are no significant auto correlation and partial correlation of the forecasting residual series. Accordingly, we can draw the following conclusions that the information contained in power load is mined well by LSSVM whose parameters are tuned by QDE.

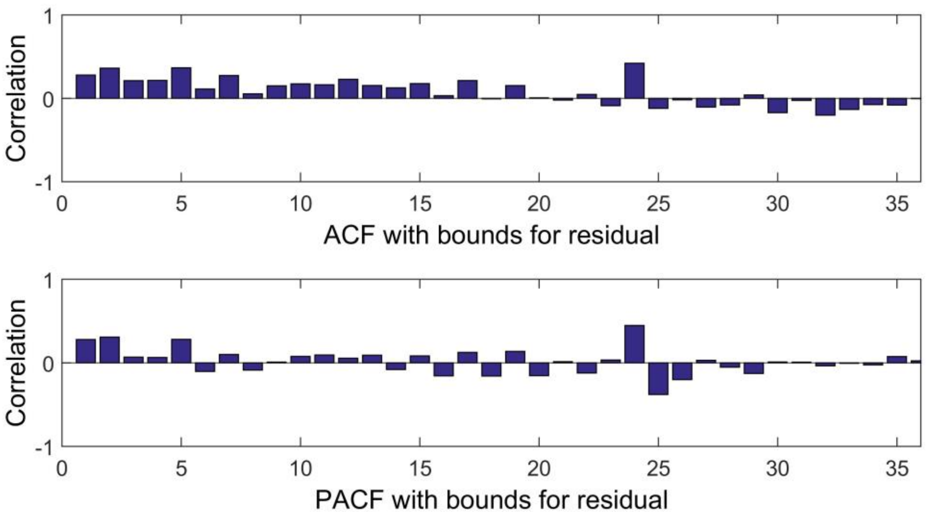

Figure 6. Auto Correlation and Partial Correlation of the Residual Series by DWT-QDE-LSSVM

4.5. Comparative Analysis

To demonstrate the advanced nature of the proposed model, this paper uses DWT-QDE-LSSVM, QDE-LSSVM, LSSVM, BPNN and ARMA to forecast the 24h short-term load from June 6, 2013 to June 10, 2013. The DWT-QDE-LSSVM with and without day type refinement and other forecasting results with day type refinement from June 6, 2013 to June 10, 2013 are shown in Figure 7.

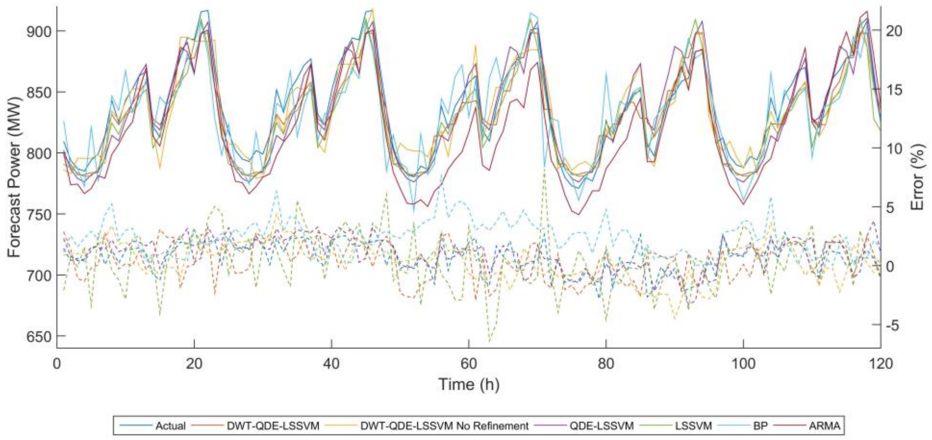

Figure 7. Forecasting Results of Different Models from June 6, 2013 to June 10, 2013

Figure 7, shows that: (a) the accuracy of DWT-QDE-LSSVM is improved after day type refinement, the highest error reduces from 3.083 to 2.891 which proves that the day type refinement can enhance the prediction performance of the model; (b) DWT-QDE-LSSVM shows the best fitting degree to the actual data while ARMA gives the worst one, [-4.48% 4.38%], [-3.32% 3.80%] and [-6.38% 8.17%] are the error range of QDE-LSSVM, LSSVM and BPNN, respectively which means that the fitting degree of QDE-LSSVM is better than LSSVM and BPNN; (c) because the optimal selection of the parameters by QDE, the prediction accuracy of LSSVM has been significantly improved. QDE algorithm uses quantum coding scheme to break the local mechanism trap. Thus, it can increase the ability to traverse the solution space during the short-term load forecasting. In the individual updating, the general quantum rotation gate strategy is applied to accelerate convergence; (d) the method based on LSSVM is better than BPNN in short-term load forecasting. As an optimization method for local search, BPNN is very sensitive to the initial network weights which may converge to different local minima. Aiming at the problems of over-fitting and generalization of BPNN, LSSVM has optimal combination of approximation precision and generalization ability; (e) compared with QDE-LSSVM, the results of DWT-QDE-LSSVM show that estimates are in good agreement with true values. It is mainly because DWT processing filters out the non-significant information of the original time series which eliminates the impact of non-significant information on forecasting results.

This paper uses MAPE, MAE and DPE to measure the performance of the three forecasting models, which are shown in Table 3.

Table 3. MAPE, MAE and DPE of DWT-QDE-LSSVM, DWT-QDE-LSSVM without Day Type Refinement, QDE-LSSVM, LSSVM, BPNN and ARMA Models

Forecasting error	Different models for short-term load forecasting					
	DWT-QDE-LSSVM	DWT-QDE-LSSVM without day type refinement	QDE-LSSVM	LSSVM	BPNN	ARMA
MAPE (%)	1.320	1.350	1.429	1.496	1.937	2.369
MAE (MW)	11.286	11.977	11.327	12.655	16.281	19.745
DPE (%)	2.890	3.083	3.809	4.089	8.179	7.631

Table 4, shows the MAPE, MAP and the DPE of the prediction methods. From Table 4, it can be concluded that: (a) the day type refinement decreases the MAPE, MAE and DPE of DWT-QDE-LSSVM by 0.03%, 0.691MW and 0.193%, respectively; (b) DWT-QDE-LSSVM has the lowest MAPE and MAE in most days and ARMA has the highest ones. The MAPE and MAE of DWT-QDE-LSSVM are up to 1.049% and 8.459MW lower than the ARMA. This suggests that, compared with ARMA, DWT-QDE-LSSVM has a significant advantage on small sample set regression for LSSVM meets the structural risk minimization principle; (b) the DPE of LSSVM is up to 1.199% higher than DWT-QDE-LSSVM while 0.280% higher than QDE-LSSVM, which indicates that the prediction efficiency of DWT-QDE-LSSVM and QDE-LSSVM are superior to LSSVM; (c) the DPE of DWT-QDE-LSSVM is 2.890% while the BPNN is 8.179%, which means the prediction results of BPNN cannot meet the stability principle of the of short-term load forecasting requirements.

5. Conclusions

Considering the historical power load data, date types, and meteorological factors, this paper uses DWT-QDE-LSSVM model to forecast the short-term load. Based on different forecasting models in this paper, it can be concluded that: (a) the day type refinement can highly improve the forecast precision; (b) using DWT-QDE-LSSVM model to forecast the short-term load can achieve higher prediction accuracy compared with QDE-LSSVM model, LSSVM model, BPNN model and ARMA model; (c) DWT increases the stability of short-term load data, thus it can improve the accuracy of the proposed model; (d) through parameters selection by QDE, the prediction accuracy of LSSVM has been significantly improved; (e) as a hybrid heuristic algorithm, the algorithm in this paper is able to provide a basis for smart grid when it establishes scientific and rational power generation plan.

Although DWT-QDE-LSSVM has obvious advantages, the application in short-term load fore-casting is worth improving, for example, the computation time is too long. The future research emphasis in this field will include not only the optimization of the proposed method but also the reduction of the computation time in order to further improve the model practicality.

References

[1] G. F. Fan, L. L. Peng, W. C. Hong and F. Sun, "Electric load forecasting by the SVR model with differential empirical mode decomposition and auto regression", Neurocomputing, vol. 173, **(2015)**, pp. 958-970.

[2] N. Tang and D. J. Zhang. "Application of a Load Forecasting Model Based on Improved Grey Neural Network in the Smart Grid", Energy Procedia, vol. 12, **(2011)**, pp. 180-184.

[3] R. Ye, P. N. Suganthan, N. Srikanth and G. Amaratunga, "Random vector functional link network for short-term electricity load demand forecasting", Information Sciences, **(2016)**.

[4] L. Zjavka and V. Snášel, "Short-term power load forecasting with ordinary differential equation substitutions of polynomial networks", Electric Power Systems Research, vol. 137, **(2016)**, pp.113-123.

[5] A. E. Clements, A. S. Hurn and Z. Li, "Forecasting day-ahead electricity load using a multiple equation time series approach", European Journal of Operational Research, vol. 251, no. 2, **(2015)**, pp. 522-530.

[6] T. Niknam, M. Zare and J. Aghaei, "Scenario-based multiobjective volt/var control in distribution networks including renewable energy sources", Power Delivery, IEEE Transactions on, vol. 27, no. 4, **(2012)**, pp. 2004-2019.

[7] S. S. Pappas, L. Ekonomou, D. C. Karamousantas, G. Chatzarakis, S. Katsikas and P. Liatsis, "Electricity demand loads modeling using AutoRegressive Moving Average (ARMA) models", Energy, vol. 33, no. 9, **(2008)**, pp. 1353-1360.

[8] L. Suganthi and A. A. Samuel, "Energy models for demand forecasting—A review", Renewable and sustainable energy reviews, vol. 16, no. 2, **(2012)**, pp. 1223-1240.

[9] M. Cellura, V. L. Brano and A. Marvuglia, "Forecasting daily urban electric load profiles using artificial neural networks", Energy Conversion and Management, vol. 45, no. 18, **(2004)**, pp. 2879-2900.

[10] D. X. Niu, Y. Wang and D. D. Wu, "Power load forecasting using support vector machine and ant colony optimization", Expert Systems with Applications, vol. 37, no. 3, **(2010)**, pp. 2531-2539.

[11] B. J. Chen, M. W. Chang and C. J. Lin, "Load forecasting using support vector machines: A study on EUNITE competition 2001", Power Systems, IEEE Transactions on, vol. 19, no. 4, **(2004)**, pp. 1821-1830.

[12] M. Mohandes, "Support vector machines for short-term electrical load forecasting", International Journal of Energy Research, vol. 26, no. 26, **(2002)**, pp. 335-345.

[13] J. C. Sousa, H. M. Jorge and L. P. Neves, "Short-term load forecasting based on support vector regression and load profiling", International Journal of Energy Research, vol. 38, no. 38, **(2014)**, pp. 350-362.

[14] K. Pelckmans, J. A. K. Suykens, T. V. Gestel, J. D. Brabanter, L. Lukas, B. Hamers, B. De Moor and J. Vandewalle, "LS-SVMlab: a matlab/c toolbox for least squares support vector machines", Tutorial Kuleuven, **(2002)**.

[15] H. Shayeghi, A. Ghasemi, M. Moradzadeh and M. Nooshyar, "Simultaneous day-ahead forecasting of electricity price and load in smart grids", Energy Conversion and Management, vol. 95, **(2015)**, pp. 371-384.

[16] W. Sun and Y. Liang, "Research of least squares support vector regression based on differential evolution algorithm in short-term load forecasting model", Journal of Renewable and Sustainable Energy, vol. 6, no. 5, **(2014)**, pp. 1-10.

[17] Z. A. Bashir and M. E. EI-Hawary, "Applying wavelets to short-term load forecasting using PSO-based neural networks", Power Systems, IEEE Transactions on, vol. 24, no. 1, **(2009)**, pp. 20-27.

[18] S. Bahrami, R. A. Hooshmand and M. Parastegari, "Short term electric load forecasting by wavelet transform and grey model improved by PSO (particle swarm optimization) algorithm", Energy, vol. 72, no. 7, **(2014)**, pp. 434-442.

[19] H. J. Su, Y. P. Yang and L. Zhao, "Classification rule discovery with DE/QDE algorithm", Expert Systems with Applications, vol. 37, no. 2, **(2010)**, pp. 1216-1222.

Authors

Sun Wei, she is a professor from North China Electric Power University major in electricity economy and power market.