Stock Price Forecasting with Empirical Mode Decomposition Based Ensemble ν-Support Vector Regression Model

Xueheng Qiu[1], Huilin Zhu[1], P.N. Suganthan[1(✉)], and Gehan A.J. Amaratunga[2]

[1] Nanyang Technological University, Singapore, Singapore
{qiux0004,hzhu007,epnsugan}@ntu.edu.sg
[2] University of Cambridge, Cambridge, UK
gaja1@cam.ac.uk

Abstract. Stock price forecasting is one of the most challenging tasks of time series forecasting due to the inherent non-linearity and non-stationary characteristics of the stock market and financial time series. In this paper, an ensemble method composed of Empirical Mode Decomposition (EMD) algorithm and ν-Support Vector Regression (ν-SVR) is presented for short-term stock price forecasting. First of all, the historical stock price time series were decomposed into several intrinsic mode functions (IMFs). Then each IMF was modeled by a ν-SVR model to generate the corresponding forecasting IMF value. Finally, the prediction results of all IMFs were combined to formulate an aggregated output for stock price. The stock market price datasets of three power related companies are used to test the effectiveness of the proposed EMD-ν-SVR method. Simulation results demonstrated attractiveness of the proposed method compared with six forecasting methods.

Keywords: Empirical mode decomposition · Ensemble method · Time series forecasting · Stock price forecasting · ν-Support vector regression · Artificial neural network

In modern financial markets and industrial fields, big data mining, analysis and forecasting play a important role for companies to optimize their plans and strategies in order to keep themselves competitive. Stock price forecasting belongs to time series (TS) forecasting paradigm, which aims to predict the future stock market price ranging from hours to several days ahead by analyzing TS data itself and extracting meaningful characteristics [29]. Among all the challenging tasks in the field of financial time series forecasting, stock price forecasting is regarded as one of the most difficult one due to the highly non-linear and non-stationary patterns of stock price TS caused by numerous influence factors, such as economy, affecting government, enterprise and investors [19]. Table 1 shows the list of nomenclatures and acronyms used in this paper.

In the literature, various methods and ideas have been published for stock price forecasting with varying degrees of success, which can be categorised into linear

J.K. Mandal et al. (Eds.): CICBA 2017, Part I, CCIS 775, pp. 22–34, 2017.
DOI: 10.1007/978-981-10-6427-2_2

Table 1. Nomenclature

TS	Time Series
ANN	Artificial Neural Network
SVM	Support Vector Machine
SVR	Support Vector Regression
EMD	Empirical Mode Decomposition
MAPE	Mean Absolute Percentage Error
RMSE	Root Mean Square Error
ARIMA	Auto-Regressive Integrated Moving Average
GARCH	Generalized Autoregressive Conditional Heteroscedasticity

statistical methods and nonlinear machine learning models [32]. For linear models, normally statistical theories and mathematical equations are used for extrapolating the future values of TS. The most successful linear models include linear regression [28], Holt-Winters exponential smoothing [20], Autoregressive Integrated Moving Average (ARIMA) [5], and so on. These conventional models are developed based on the assumption that the TS being forecasted are linear and stationary. However, in the literature, it has been verified that these conventional modeling techniques are not adequate for stock market price forecasting [23].

Nonlinear machine learning methods can learn features from and also make predictions on TS data, which build a model from example inputs in order to make data-driven predictions, instead of following strictly static program instructions [25]. With the rapid development of computational intelligence, machine learning methods have been widely applied for various research fields including stock price forecasting. The most widely used machine learning algorithms include generalized autoregressive conditional heteroscedasticity (GARCH) [4,15], artificial neural network (ANN) [11], support vector regression (SVR) [10] and fuzzy comprehensive evaluation [40]. For example, in [3], an evolutionary Levenberg-Marquardt neural networks based hybrid model was proposed for stock price forecasting, along with some data pre-processing techniques.

ANNs have been widely used for stock price forecasting in the literature [24,41], but still suffer from local minimum traps and over-fitting problems [36]. On the contrary, the kernel machines, such as SVR, have global optimums due to their structural risk minimization principle. SVR considers both the training error and the capacity of the regression model to avoid under-fitting and over-fitting problems in the training process [39]. For example, in [23], a forecasting model based on chaotic mapping, firefly algorithm and SVR has been proposed to predict stock market price with higher accuracy than ANNs. However, in the literature, most publications are focused on the original form of SVR, namely ϵ-SVR, while the updated version ν-SVR still needs further investigation [37]. Therefore, in this work, ν-SVR is mainly focused, at the same time, a comparison between ν-SVR and ϵ-SVR is also conducted.

Ensemble learning methods, or hybrid methods, aim to obtain better forecasting performance by strategically combining multiple algorithms. Dietterich has concluded the success of ensemble methods due to three fundamental reasons: statistical, computational and representational [13]. Ensemble learning can be divided into two categories according to the way of combination sequential and parallel [33]. In a sequentially combined ensemble method, the outputs from several forecasting models are treated as the inputs to another forecasting method [6,31]. For a parallel combined ensemble method, the training TS is decomposed into a collection of sub-datasets [9]. Then we train a forecasting model for each TS, and aggregate the outputs from all the models to calculate final prediction results. There are many examples of parallel ensemble methods in the literature, such as wavelet decomposition [17,21], empirical mode decomposition (EMD) [22,30] and negative correlation learning [2].

In this paper, an ensemble method composed of EMD and ν-SVR is proposed for short-term stock price forecasting. The attractiveness of the proposed method is demonstrated on real world stock price datasets of power related companies, and compared with six benchmark learning algorithms: Persistence, ANN, ϵ-SVR, ν-SVR, EMD based ANN and EMD based ϵ-SVR models.

The remaining of this paper is organized as follows: Sect. 1 explains the theoretical background on forecasting methods. Section 2 presents the algorithm of proposed EMD-ν-SVR approach. Section 3 shows the procedures for experiment setup, followed by the discussion about experiment results in Sect. 4. Finally in Sect. 5, the conclusions and future works are stated.

1 Review of Forecasting Models

1.1 Support Vector Regression

The Support Vector Machine (SVM) is a machine learning algorithm proposed by Cortes and Vapnik [10] based on statistical learning theory. Structural risk minimization is the basic concept of this method. A version of SVM for regression was proposed in [14]. Support vector regression has been widely applied in time series forecasting problems [35]. There are two versions of SVR: ϵ-SVR and ν-SVR.

ϵ-Support Vector Regression. Suppose a time series data set is given as follows

$$D = \{(\boldsymbol{x}_i, y_i)\}, 1 \leq i \leq N \tag{1}$$

where $\mathbf{x}_i$ is the input vector at time i with m elements and y_i is the corresponding output data. The regression function can be defined as

$$f(\boldsymbol{x}) = \boldsymbol{w}^T \phi(\boldsymbol{x}) + b \tag{2}$$

where $\boldsymbol{w}$ is the weight vector, b is the bias, and $\phi(\boldsymbol{x})$ maps the input vector $\boldsymbol{x}$ to a higher dimensional feature space. $\boldsymbol{w}$ and b can be obtained by solving the following optimization problem:

$$\min_{\boldsymbol{w},b,\boldsymbol{\xi},\boldsymbol{\xi}^*} \frac{1}{2}\boldsymbol{w}^T\boldsymbol{w} + C\sum_{i=1}^{N}(\xi_i + \xi_i^*) \tag{3}$$

Subject to:

$$\begin{aligned} & y_i - \boldsymbol{w}^T\phi(\boldsymbol{x}_i) - b \leq \epsilon + \xi_i \\ & \boldsymbol{w}^T\phi(\boldsymbol{x}_i) + b - y_i \leq \epsilon + \xi_i^* \\ & \xi_i, \xi_i^* \geq 0 \end{aligned} \tag{4}$$

where C is a predefined positive trade-off parameter between model simplicity and generalization ability, ξ_i and ξ_i^* are the slack variables measuring the cost of the errors.

For nonlinear input data set, kernel functions can be used to map from original space onto a higher dimensional feature space in which a linear regression model can be built. The dual problem is

$$\min_{\boldsymbol{\alpha},\boldsymbol{\alpha}^*} \frac{1}{2}(\boldsymbol{\alpha} - \boldsymbol{\alpha}^*)^T Q(\boldsymbol{\alpha} - \boldsymbol{\alpha}^*) + \epsilon\sum_{i=1}^{N}(\alpha_i + \alpha_i^*) + \sum_{i=1}^{N} y_i(\alpha_i - \alpha_i^*) \tag{5}$$

Subject to:

$$\begin{aligned} & \boldsymbol{e}^T(\boldsymbol{\alpha} - \boldsymbol{\alpha}^*) = 0, \\ & 0 \leq \alpha_i, \alpha_i^* \leq C \end{aligned} \tag{6}$$

where $\boldsymbol{e} = [1, \ldots, 1]^T$ is the vector of all ones, Q is an l by l positive semi-definite matrix, K is the kernel function, $Q_{ij} = K(\boldsymbol{x}_i, \boldsymbol{x}_j) \equiv \phi(\boldsymbol{x}_i)^T\phi(\boldsymbol{x}_j)$.

Thus, the final SVR function is obtained as

$$y_i = f(\boldsymbol{x}_i) = \sum_{i=1}^{N}(\alpha_i^* - \alpha_i)K(\boldsymbol{x}_i, \boldsymbol{x}) + b \tag{7}$$

where α_i and α_i^* are the Lagrange multipliers. The most frequently used kernel function is the Gaussian radial function (RBF) with a width of σ

$$K(\boldsymbol{x}_i, \boldsymbol{x}_j) = exp(-\left\|\boldsymbol{x}_i - \boldsymbol{x}_j\right\|^2 / (2\sigma^2)) \tag{8}$$

$\boldsymbol{\nu}$-Support Vector Regression. ν-SVR uses a parameter $\nu \in (0, 1]$ to control the number of support vectors [34]. It is proved that ν is an upper bound on the fraction of training errors and a lower bound of the fraction of support vectors [7].

Similar with ϵ-SVR, with (C, ν) as parameters, ν-SVR solves

$$\min_{\boldsymbol{w},b,\boldsymbol{\xi},\boldsymbol{\xi}^*,\epsilon} \frac{1}{2}\boldsymbol{w}^T\boldsymbol{w} + C(\nu\epsilon + \frac{1}{N}\sum_{i=1}^{N}(\xi_i + \xi_i^*)) \tag{9}$$

Subject to:

$$\begin{aligned} & y_i - \boldsymbol{w}^T\phi(\boldsymbol{x}_i) - b \leq \epsilon + \xi_i \\ & \boldsymbol{w}^T\phi(\boldsymbol{x}_i) + b - y_i \leq \epsilon + \xi_i^* \\ & \xi_i, \xi_i^* \geq 0, \epsilon \geq 0. \end{aligned} \tag{10}$$

The dual problem is

$$\min_{\boldsymbol{\alpha},\boldsymbol{\alpha}^*} \frac{1}{2}(\boldsymbol{\alpha}-\boldsymbol{\alpha}^*)^T Q(\boldsymbol{\alpha}-\boldsymbol{\alpha}^*) + \boldsymbol{y}^T(\boldsymbol{\alpha}-\boldsymbol{\alpha}^*) \tag{11}$$

Subject to:

$$\begin{aligned} &\boldsymbol{e}^T(\boldsymbol{\alpha}-\boldsymbol{\alpha}^*) = 0, \boldsymbol{e}^T(\boldsymbol{\alpha}-\boldsymbol{\alpha}^*) \le C\nu, \\ &0 \le \alpha_i, \alpha_i^* \le C/N \end{aligned} \tag{12}$$

The final approximate function is

$$y_i = f(\boldsymbol{x}_i) = \sum_{i=1}^{N}(\alpha_i^* - \alpha_i)K(\boldsymbol{x}_i, \boldsymbol{x}) + b \tag{13}$$

1.2 Empirical Mode Decomposition

EMD [22], also known as Hilbert-Huang transform (HHT), is a method to decompose a signal into several intrinsic mode functions (IMF) along with a residue which stands for the trend. EMD is an empirical approach to obtain instantaneous frequency data from non-stationary and nonlinear data sets.

The system load is a random non-stationary process composed of thousands of individual components. The system load behavior is influenced by a number of factors, which can be classified as: economic factors, time, day, season, weather and random effects. Thus, EMD algorithm can be very effective for load demand forecasting.

An IMF is a function that has only one extreme between zero crossings, along with a mean value of zero. The shifting process which EMD uses to decompose the signal into IMFs is described as follows:

1. For a time series signal $x(t)$, let m_1 be the mean of its upper and lower envelopes as determined by a cubic-spline interpolation of local maxima and minima.
2. The first component h_1 is computed by subtracting the mean from the original time series: $h_1 = x(t) - m_1$.
3. In the second shifting process, h_1 is treated as the data, and m_{11} is the mean of h_1's upper and lower envelopes: $h_{11} = h_1 - m_{11}$.
4. This shifting procedure is repeated k times until one of the following stop criterion is satisfied: (i) m_{1k} approaches zero, (ii) the numbers of zero-crossings and extrema of h_{1k} differs at most by one, or (iii) the predefined maximum iteration is reached. h_{1k} can be treated as an IMF in this case and computed by: $h_{1k} = h_{1(k-1)} - m_{1k}$.
5. Then it is designated as $c_1 = h_{1k}$, the first IMF component from the data, which contains the shortest period component of the signal. We separate it from the rest of the data: $x(t) - c_1 = r_1$. The procedure is repeated on r_j: $r_1 - c_2 = r_2, \ldots, r_{(n-1)} - c_n = r_n$.

As a result, the original time series signal is decomposed as a set of functions: $x(t) = \sum_{i=1}^{n}(c_i) + r_n$, where the number of functions n in the set depends on the original signal.

2 Proposed Ensemble Method

"Divide and conquer" is an ensemble method which works by decomposing the original TS into a series of sub-datasets until they are simple enough to be analyzed. For proposed EMD-ν-SVR approach, as mentioned above, the stock price data is decomposed into several IMFs and one residue by EMD method. Then a ν-SVR model is trained for each IMF including the residue. The final prediction results are given by combining the outputs from all sub-series using another ν-SVR model. Figure 1 is the schematic diagram of this proposed ensemble method, and the procedures can be concluded as:

1. Use EMD to decompose the original TS into several IMFs and one residue.
2. Construct the training matrix as the input of each ν-SVR for each IMF and residue.
3. Train ν-SVR models to obtain the prediction results for each of the extracted IMF and residue.
4. Combine all the prediction results by another ν-SVR model to formulate an ensemble output for TS forecasting.

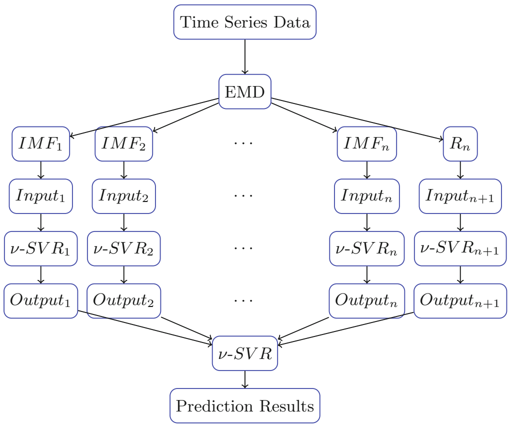

Fig. 1. Schematic diagram of the proposed EMD-KRR-SVR approach

3 Experiment Setup

3.1 Datasets

In this paper, the stock price datasets of power related companies were used for evaluating the performance of benchmark learning models. Daily closing stock market prices for Surgutneftegas (from 31/12/2007 to 2/9/2016), Lukoil (from

18/11/1996 to 13/1/2017) and Exxonmobil (from 3/1/1993 to 3/31/2017) were extracted from Yahoo Finance [1]. For each dataset, to compare the performance of learning models with different forecasting horizons, three kinds of simulations were conducted: one day ahead, two days ahead and one week ahead forecasting. For each dataset, 80% of the data points were used for training, while the remaining 20% was used for testing.

3.2 Methodology

For the time series stock price datasets, all the training and testing values are linearly scaled to $[0, 1]$. The scaling formula is:

$$\bar{y}_i = \frac{y_{max} - y_i}{y_{max} - y_{min}} \tag{14}$$

To implement the simulation, LIBSVM toolbox was used for SVR based models, including ϵ-SVR, ν-SVR, EMD-ϵ-SVR, and the proposed EMD-ν-SVR [8]. Neural network toolbox in Matlab was used for constructing neural networks, including ANN and EMD based ANN (EMD-ANN).

For SVR based models, we use the RBF kernel function with parameters chosen by a grid search. The range of C is $[2^{-4}, 2^4]$, and the range of σ is $[10^{-3}, 10^{-1}]$. For ANN and EMD-ANN, the size of neural networks is determined by the size of input vector. The number of iterations for back propagation is set as 1000.

3.3 Error Measurement

To examine the accuracy of the prediction model, two evaluation measures are used in this study: Root Mean Square Error (RMSE) and Mean Absolute Percentage Error (MAPE). They are defined as:

$$\begin{aligned} RMSE &= \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i' - y_i)^2} \\ MAPE &= \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_i' - y_i}{y_i}\right| \end{aligned} \tag{15}$$

where y_i' is the predicted value of corresponding y_i, and n is the number of data points in the testing time series.

4 Results and Discussion

In this work, six benchmark methods were implemented for stock market price forecasting to perform a comparison with the proposed EMD-ν-SVR model. First of all, the persistence method was employed as the baseline for comparing the performance of learning models. This method assumes the conditions at the future time the same as the current values and has relatively reasonable accuracy for stock price TS. The prediction results for short-term stock price forecasting are shown in Table 2, where the forecasting horizons are one day, two days and one week. The numbers in bold mean that the corresponding method has the best performance for this dataset under this performance measure. According to the prediction results, we can conclude that all the machine learning models outperform the persistence method for short-term stock price forecasting.

To reveal the advantages of EMD based ensemble methods, we implemented the single structure models ANN, ϵ-SVR and ν-SVR for stock price forecasting, and conducted an comparison with their EMD hybrid models. From the results shown in Table 2, it is obviously that EMD can bring significant improvements for short-term stock price forecasting. Besides that, ν-SVR generally performs better than ϵ-SVR in all cases. However, in some cases, the differences between these two models are not so significant. Moreover, the proposed EMD-ν-SVR achieves the best performance for every dataset, which means that the proposed method has more advantages compared with the benchmark models.

Table 2. Prediction results for stock market price forecasting

Dataset	Horizon	Metrics	Prediction model						
			Persistence	ANN [18]	ϵ-SVR [10]	ν-SVR [34]	EMD-ANN [26]	EMD-ϵ-SVR [38]	Proposed
Lukoil	1 day	RMSE	1.305	0.957	0.950	0.950	0.540	0.463	**0.461**
		MAPE	2.106%	1.543%	1.531%	1.532%	0.881%	0.738%	**0.737%**
	2 days	RMSE	1.588	1.325	1.309	1.306	0.706	0.709	**0.704**
		MAPE	2.608%	2.158%	2.118%	2.115%	1.136%	1.119%	**1.114%**
	1 week	RMSE	2.179	2.126	2.063	2.040	1.167	0.912	**0.884**
		MAPE	3.623%	3.602%	3.463%	3.404%	1.956%	1.498%	**1.461%**
Surgutneftegas	1 day	RMSE	18.762	18.488	13.288	13.181	12.134	10.516	**7.486**
		MAPE	2.846%	2.699%	1.943%	1.913%	1.843%	1.795%	**1.141%**
	2 days	RMSE	22.628	22.388	18.417	18.407	14.567	8.973	**8.762**
		MAPE	3.389%	3.258%	2.696%	2.694%	2.246%	1.281%	**1.256%**
	1 week	RMSE	36.038	32.350	30.272	29.888	17.038	18.382	**15.575**
		MAPE	5.764%	4.579%	4.469%	4.321%	2.707%	2.651%	**2.280%**
Exxonmobil	1 day	RMSE	1.337	0.967	0.963	0.962	0.707	0.486	**0.483**
		MAPE	1.130%	0.818%	0.811%	0.811%	0.617%	0.416%	**0.414%**
	2 days	RMSE	1.635	1.414	1.339	1.335	0.823	0.665	**0.662**
		MAPE	1.389%	1.226%	1.142%	1.138%	0.727%	0.570%	**0.566%**
	1 week	RMSE	2.161	1.949	1.990	1.986	1.177	0.906	**0.904**
		MAPE	1.882%	1.760%	1.730%	1.724%	1.023%	0.773%	**0.763%**

In order to give a detailed analysis of these results, we employ Friedman test [16] and Nemenyi post-hoc test [27] to test the significance of the differences among these learning models. The Friedman test ranks the algorithms for each dataset separately, and then assign average ranks in case of ties. The null-hypothesis states that all the algorithms have the same performance. If the null-hypothesis is rejected, in order to tell whether the performances of two among totally k learning models are significantly different, the Nemenyi post-hoc test is applied to compare all the learning models with each other. The comparison results of statistical test based on RMSE and MAPE are shown in Figs. 2 and 3, respectively. The methods with better ranks are at the top whereas the methods with worse ranks are at the bottom. It is worth noting that the models within a vertical line whose length is less than or equal to a critical distance have statistically the same performance. The critical distance for Nemenyi test is defined as:

$$CD = q_{\alpha}\sqrt{\frac{k(k+1)}{6N}} \tag{16}$$

where k is the number of algorithms, N is the number of data sets, and q_{α} is the critical value based on the Studentized range statistic divided by $\sqrt{2}$ [12]. From the statistical test results, the proposed EMD-ν-SVR achieves the best rank and significantly outperforms the non-EMD based methods with a 95% confidence.

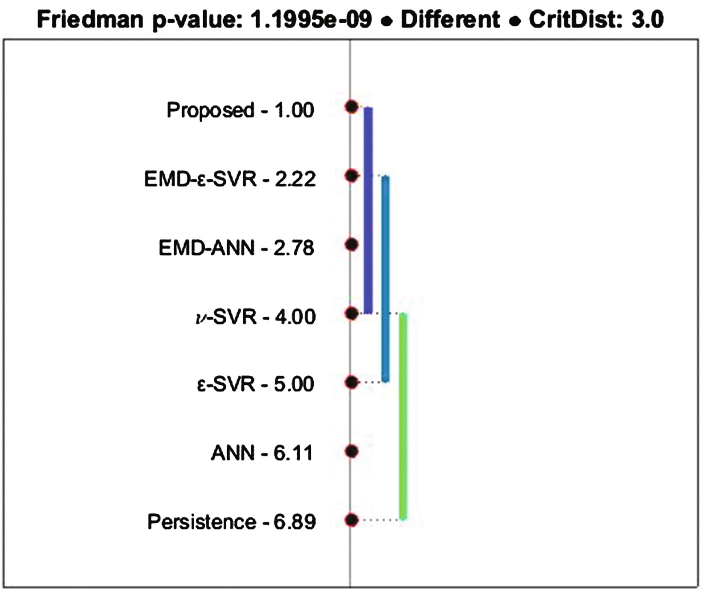

Fig. 2. Nemenyi test for stock price forecasting based on RMSE. The critical distance is 3.0.

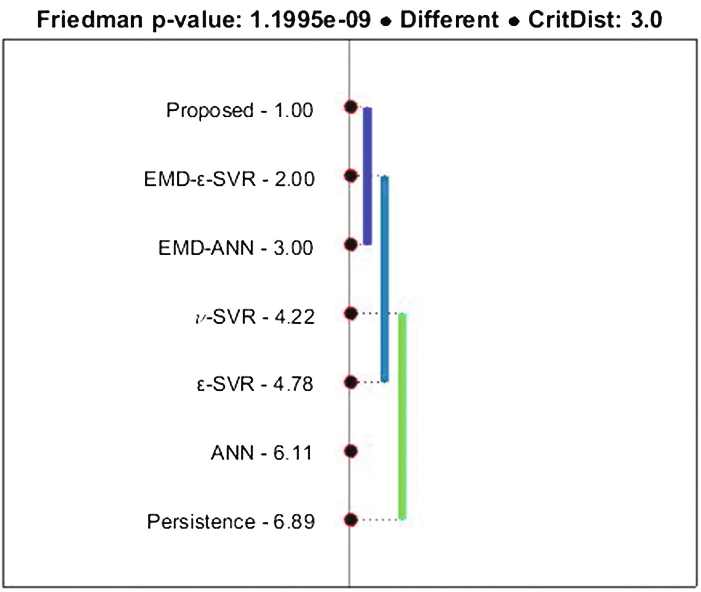

Fig. 3. Nemenyi test for stock price forecasting based on MAPE. The critical distance is 3.0.

5 Conclusion

In this paper, we proposed an ensemble learning approach for short-term stock market price forecasting composed of EMD and ν-SVR. Three stock price datasets from power related companies were used for evaluating the performance of the proposed method. Moreover, six benchmarks methods were implemented to perform a comparison with the proposed method. From the forecasting results, the following conclusions are made:

1. ν-SVR generally performs better than ϵ-SVR for short-term stock price forecasting. However, in some cases, the differences between these two models are not so significant.
2. EMD based hybrid methods, including EMD-ANN, EMD-ϵ-SVR, and the proposed EMD-ν-SVR, significantly outperform the corresponding single structure models for short-term stock price time series forecasting.
3. The proposed EMD-ν-SVR approach achieves the best performance for stock price forecasting according to the statistical testing.

For future research directions, more influence factors, such as economy, affecting government, enterprise and investors need to be considered in order to construct much more complex multivariate models for stock price forecasting. Moreover, ν-SVR and its ensemble models can also be tested using other types of TS, such as renewable energy data, to evaluate the performance in the generic situation.

Acknowledgment. This project is funded by the National Research Foundation Singapore under its Campus for Research Excellence and Technological Enterprise (CREATE) programme.

References

1. Yahoo finance, April 2017. http://www.finance.yahoo.com/
2. Alhamdoosh, M., Wang, D.: Fast decorrelated neural network ensembles with random weights. Inf. Sci. **264**, 104–117 (2014)
3. Asadi, S., Hadavandi, E., Mehmanpazir, F., Nakhostin, M.M.: Hybridization of evolutionary LevenbergMarquardt neural networks and data pre-processing for stock market prediction. Knowl. Based Syst. **35**, 245–258 (2012)
4. Bollerslev, T.: Generalized autoregressive conditional heteroskedasticity. J. Econometrics **31**, 307–327 (1986)
5. Box, G.E.P., Jenkins, G.M.: Time Series Analysis: Forecasting and Control. Holden-Day Series in Time Series Analysis and Digital Processing. Holden-Day, San Francisco (1976)
6. Breiman, L.: Stacked regressions. Mach. Learn. **24**, 49–64 (1996)
7. Chang, C.C., Lin, C.J.: Training ν-support vector regression: theory and algorithms. Neural Comput. **14**, 1959–1977 (2002)
8. Chang, C.C., Lin, C.J.: LIBSVM: a library for support vector machines. ACM Trans. Intell. Syst. Technol. (TIST) **2**(3), 27 (2011)
9. Cormen, T.H., Leiserson, C.E., Rivest, R.L., Stein, C.: Introduction to Algorithms. MIT Press, Cambridge (2000)
10. Cortes, C., Vapnik, V.: Support-vector networks. Mach. Learn. **20**(3), 273–297 (1995)
11. Darbellay, G.A., Slama, M.: Forecasting the short-term demand for electricity: do neural networks stand a better chance? Int. J. Forecast. **16**, 71–83 (2000)
12. Demšar, J.: Statistical comparisons of classifiers over multiple data sets. J. Mach. Learn. Res. **7**, 1–30 (2006)
13. Dietterich, T.G.: Ensemble methods in machine learning. In: Kittler, J., Roli, F. (eds.) MCS 2000. LNCS, vol. 1857, pp. 1–15. Springer, Heidelberg (2000). doi:10.1007/3-540-45014-9_1
14. Drucker, H., Burges, C.J., Kaufman, L., Smola, A., Vapnik, V.: Support vector regression machines. Adv. Neural Inf. Process. Syst. **9**, 155–161 (1997)
15. Engle, R.F.: Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. Econometrica **50**, 987–1008 (1982)
16. Friedman, M.: The use of ranks to avoid the assumption of normality implicit in the analysis of variance. J. Am. Stat. Assoc. **32**(200), 675–701 (1937)
17. Guan, C., Luh, P.B., Michel, L.D., Wang, Y., Friedland, P.B.: Very short-term load forecasting: wavelet neural networks with data pre-filtering. IEEE Trans. Power Syst. **28**(1), 30–41 (2013)
18. Haykin, S.: Neural Networks: A Comprehensive Foundation. Prentice Hall, London (1999). International edition
19. He, K., Zha, R., Wu, J., Lai, K.K.: Multivariate EMD-based modeling and forecasting of crude oil price. Sustainability **8**(4), 387 (2016)
20. Holt, C.C.: Forecasting seasonals and trends by exponentially weighted moving averages. Int. J. Forecast. **20**, 5–10 (2004)

21. Hooshmand, R.A., Amooshahi, H., Parastegari, M.: A hybrid intelligent algorithms based short-term load forecasting approach. Int. J. Electr. Power Energ. Syst. **45**, 313–324 (2013)
22. Huang, N.E., Shen, Z., Long, S.R., Wu, M.C., Shih, H.H., Zheng, Q., Yen, N.C., Tung, C.C., Liu, H.H.: The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. Roy. Soc. London A. **454**, 903–995 (1998)
23. Kazem, A., Sharifi, E., Hussain, F.K., Saberi, M., Hussain, O.K.: Support vector regression with chaos-based firefly algorithm for stock market price forecasting. Appl. Soft Comput. **13**, 947–958 (2013)
24. Jae Kim, K., Han, I.: Genetic algorithms approach to feature discretization in artificial neural networks for the prediction of stock price index. Exp. Syst. Appl. **19**, 125–132 (2008)
25. Kohavi, R., Provost, F.: Glossary of terms. Mach. Learn. **30**, 271–274 (1998)
26. Liu, H., Chen, C., Tian, H., Li, Y.: A hybrid model for wind speed prediction using empirical mode decomposition and artificial neural networks. Renew. Energ. **48**, 545–556 (2012)
27. Nemenyi, P.: Distribution-free Multiple Comparisons. Princeton University, Princeton (1963)
28. Papalexopoulos, A.D., Hesterberg, T.C.: A regression-based approach to short-term system load forecasting. IEEE Trans. Power Syst. **5**, 1535–1547 (1990)
29. Qiu, X., Ren, Y., Suganthan, P.N., Amaratumga, G.A.J.: Empirical mode decomposition based ensemble deep learning for load demand time series forecasting. Appl. Soft Comput. **54**, 246–255 (2017)
30. Qiu, X., Suganthan, P.N., Amaratunga, G.A.J.: Electricity load demand time series forecasting with empirical mode decomposition based random vector functional link network. In: Proceedings of IEEE Conference on Systems, Man and Cybernetics (SMC 2016), Budapest, Hungary, December 2016
31. Qiu, X., Zhang, L., Ren, Y., Suganthan, P.N., Amaratunga, G.: Ensemble deep learning for regression and time series forecasting. In: Proceedings of IEEE Symposium on Computational Intelligence in Ensemble Learning (CIEL), Orlando, US, December 2014
32. Ren, Y., Suganthan, P.N., Srikanth, N., Amaratunga, G.: Random vector functional link network for short-term electricity load demand forecasting. Inf. Sci. **367–368**, 1078–1093 (2016)
33. Ren, Y., Zhang, L., Suganthan, P.: Ensemble classification and regression-recent developments, applications and future directions. IEEE Comput. Intell. Mag. **11**(1), 41–53 (2016). [review article]
34. Schölkopf, B., Smola, A.J., Williamson, R.C., Bartlett, P.L.: New support vector algorithms. Neural Comput. **12**, 1207–1245 (2000)
35. Sousa, J.C., Jorge, H.M., Neves, L.P.: Short-term load forecasting based on support vector regression and load profiling. Int. J. Energ. Res. **38**(3), 350–362 (2014)
36. Vapnik, V.: The Nature of Statistical Learning Theory. Springer, Heidelberg (1995)
37. Wang, J., Hou, R., Wang, C., Shen, L.: Improved ν-Support vector regression model based on variable selection and brain storm optimization for stock price forecasting. Appl. Soft Comput. **49**, 164–178 (2016)
38. Ye, L., Liu, P.: Combined model based on emd-svm for short-term wind power prediction. Proc. Chin. Soc. Electr. Eng. (CSEE) **31**, 102–108 (2011)
39. Yeh, C.Y., Huang, C.W., Lee, S.J.: A multiple-kernel support vector regression approach for stock market price forecasting. Expert Syst. Appl. **38**, 2177–2186 (2010)

40. Ying, L.C., Pan, M.C.: Using adaptive network based fuzzy inference system to forecast regional electricity loads. Energ. Convers. Manag. **49**, 205–211 (2008)
41. Yu, L., Wang, S., Lai, K.K.: A neural-network-based nonlinear metamodeling approach to financial time series forecasting. Appl. Soft Comput. **9**, 563–574 (2009)

http://www.springer.com/978-981-10-6426-5

Computational Intelligence, Communications, and
Business Analytics
First International Conference, CICBA 2017, Kolkata,
India, March 24 – 25, 2017, Revised Selected Papers,
Part I
Mandal, J.K.; Dutta, P.; Mukhopadhyay, S. (Eds.)
2017, XXXI, 595 p. 295 illus., Softcover
ISBN: 978-981-10-6426-5