Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 108C (2017) 1308–1317

Procedia Computer Science

International Conference on Computational Science, ICCS 2017, 12-14 June 2017, Zurich, Switzerland

Short-term Electricity Price Forecasting with Empirical Mode Decomposition based Ensemble Kernel Machines

Xueheng Qiu[1], P. N. Suganthan[1], and Gehan A. J. Amaratunga[2]

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

Abstract

Short-term electricity price forecasting is a critical issue for the operation of both electricity markets and power systems. An ensemble method composed of Empirical Mode Decomposition (EMD), Kernel Ridge Regression (KRR) and Support Vector Regression (SVR) is presented in this paper. For this purpose, the electricity price signal was first decomposed into several intrinsic mode functions (IMFs) by EMD, followed by a KRR which was used to model each extracted IMF and predict the tendencies. Finally, the prediction results of all IMFs were combined by an SVR to obtain an aggregated output for electricity price. The electricity price datasets from Australian Energy Market Operator (AEMO) are used to test the effectiveness of the proposed EMD-KRR-SVR approach. Simulation results demonstrated attractiveness of the proposed method based on both accuracy and efficiency.

Peer-review under responsibility of the scientific committee of the International Conference on Computational Science

Keywords: Electricity Price Forecasting, Kernel Ridge Regression, Support Vector Regression, Empirical Mode Decomposition, Ensemble Learning

1 Introduction

Electricity price forecasting plays an important role in the power market operation nowadays. Under the help of accurate short term electricity price forecasting methods, not only the power suppliers are able to adjust their biding strategies to achieve the maximum benefit, but also consumers can decide whether to buy electricity from the pool or use self-production capability to avoid unacceptable high prices [24]. Short term electricity price forecasting belongs to time series (TS) forecasting paradigm, which aims to predict the future electricity price ranging from hours to on day ahead by analyzing TS data itself and extracting meaningful characteristics. However, electricity is economically non-storable, while a constant balance between production and consumption is needed for stable power supply. In practice, electricity load demand TS often performs highly nonlinear patterns due to various exogenous factors such as climate change, economic fluctuation, special occasions, and so on [13, 27]. These unique and specific reasons

1877-0509 © 2017 The Authors. Published by Elsevier B.V.
Peer-review under responsibility of the scientific committee of the International Conference on Computational Science
10.1016/j.procs.2017.05.055

lead to price dynamics not observed in any other market and thus make accurate electricity price forecasting a challenging task [32].

Over the past seventeen years since the year 2000, a wide variety of methods and ideas have been published for electricity price forecasting (EPF) with varying degrees of success, which can be categorised into linear statistical methods and nonlinear machine learning models [28]. 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 [25], Holt-Winters exponential smoothing [18], Autoregressive Integrated Moving Average (ARIMA) [3], and so on. 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 [21]. With the rapid development of computational intelligence, machine learning methods have been widely applied for various research fields including short-term electricity price forecasting. The most widely used machine learning algorithms include artificial neural network (ANN) [9], support vector regression (SVR) [8], fuzzy comprehensive evaluation [34], etc.

Kernel machines has become very popular since Support Vector Machine (SVM) being introduced in 1995 [8]. To define complex functions of the input space, SVM performs a non-linear mapping of the data into a high dimensional space, which is known as "kernel tricks". SVM has the advantage of giving a single solution that is characterized by the global minimum of the optimized functional, compared to ANN which is frequently trapped in a local minimum. Many SVM based electricity price forecasting algorithms exist in the literature. For example, in [6], a hybrid model called SVR-ARIMA that combines both SVR and ARIMA models was proposed for short term EPF problems. Besides for SVM, possibly the most elementary algorithm that can be kernelized is ridge regression. In other words, Kernel Ridge Regression (KRR) combines Ridge Regression (linear least squares with $l2-$norm regularization) with the kernel trick. In contrast to SVR, fitting KRR can be done in closed-form and is typically faster for medium-sized datasets [17, 35].

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 [11]. Ensemble learning can be divided into two categories according to the way of combination sequential and parallel [29]. In a sequentially combined ensemble method, the outputs from several forecasting models are treated as the inputs to another forecasting method [4, 27]. For a parallel combined ensemble method, the training TS is decomposed into a collection of sub-datasets [7]. 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 [15, 19], empirical mode decomposition (EMD) [20] and negative correlation learning [2].

In this paper, an ensemble method composed of EMD, KRR and SVR is proposed for short-term electricity load demand forecasting. The attractiveness of the proposed method are demonstrated on real world datasets compared with six benchmark learning algorithms: Persistence, SVR, SLFN, KRR, EMD based SVR and EMD based SLFN models.

The remaining of this paper is organized as follows: Section 2 explains the theoretical background on forecasting methods. Section 3 presents the algorithm of proposed EMD-KRR-SVR approach. Section 4 shows the procedures for experiment setup, followed by the discussion about experiment results in Section 5. Finally in Section 6, the conclusions and future works are stated.

2 Review of Forecasting Models

2.1 Support Vector Regression

The Support Vector Machine (SVM) is a machine learning algorithm proposed by Cortes and Vapnik [8] based on statistical learning theory. Structural risk minimization is the basic concept of this method. A version of SVM for regression was proposed in [12]. Support vector regression (SVR) has been widely applied in time series forecasting problems [30].

Suppose a time series data set is given as follows

$$D = \{(X_i, y_i)\}, 1 \leq i \leq N \tag{1}$$

where 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

$$f(X_i) = W^T \phi(X_i) + b \tag{2}$$

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

$$\text{Min}\ \frac{1}{2}\|W\|^2 + C\sum_{i=1}^{N}(\varepsilon_i + \varepsilon_i^*) \tag{3}$$

Subject to:

$$\begin{aligned} & y_i - W^T(\varphi(x)) - b \leq \xi + \varepsilon_i \\ & W^T(\varphi(x)) + b - y_i \leq \xi + \varepsilon_i^* \\ & \varepsilon_i, \varepsilon_i^* \geq 0 \end{aligned} \tag{4}$$

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. Thus, the final SVR function is obtained as

$$y_i = f(X_i) = \sum_{i=1}^{N}((\alpha_i - \alpha_i^*)K(X_i, X_j)) + b \tag{5}$$

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

$$K(X_i, X_j) = exp(-\|X_i - X_j\|^2/(2\sigma^2)) \tag{6}$$

2.2 Kernel Ridge Regression

Ridge Regression is a linear model which addresses ordinary least squares by imposing a penalty on the size of coefficients (l2-norm regularization) [26]. The ridge coefficients minimize a penalized residual sum of squares which is shown as follows:

$$\min_w ||Xw - y||_2^2 + \alpha||w||_2^2 \tag{7}$$

where α is a complexity parameter that controls the amount of shrinkage. The coefficients are more robust to collinearity as α becomes lager.

Kernel ridge regression (KRR) combines Ridge Regression with the kernel trick [17, 35]. Thus it constructs a linear model in the space induced by the kernel we used for the data. The form of the model learned by KRR is similar with SVR, except for the different loss functions. KRR uses squared error loss instead of ε-insensitive loss which is applied in SVR. Moreover, KRR can be trained in closed-form and is typically faster for medium-sized datasets.

2.3 Empirical Mode Decomposition

EMD [20], 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.

3 Proposed Ensemble Method

In this work, an ensemble method called "divide and conquer" is employed, which works by decomposing the original TS into a series of sub-datasets until they are simple enough to be analyzed. For proposed EMD-KRR-SVR approach, as mentioned above, the electricity price data is decomposed into several IMFs and one residue by EMD method. Then a KRR network is trained for each IMF including the residue, which is much more efficient than using SVR or

SLFN. The final prediction results are given by combining the outputs from all sub-series using an SVR model, which ensures the overall accuracy. Figure 1 is the schematic diagram of this proposed ensemble method, and the procedures can be concluded as:

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

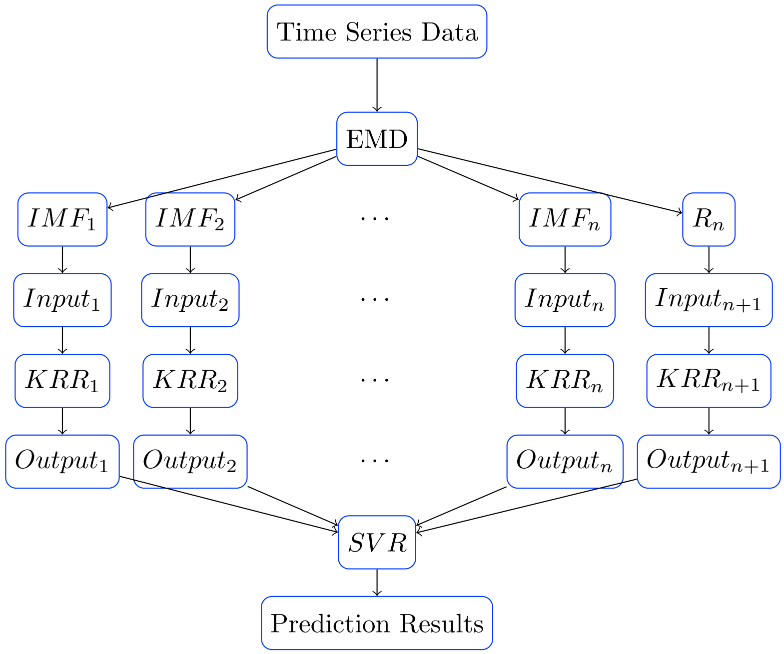

Figure 1: Schematic Diagram of the Proposed EMD-KRR-SVR approach

4 Experiment setup

4.1 Datasets

In this paper, the electricity price datasets from Australian Energy Market Operator (AEMO) [1] were used for evaluating the performance of benchmark learning models. There are totally three electricity price datasets of year 2016 from three states of Australia: New South Wales (NSW), Tasmania (TAS) and Queensland (QLD). For each dataset, to reduce the influence of climate change due to different season, four months were selected to perform comparison: January, April, July and October. For each month, the first three weeks were used for training, while the remaining one week was used for testing.

4.2 Methodology

For the time series electricity price datasets, all the training and testing values are linearly scaled to $[0, 1]$. To implement the simulation, LIBSVM toolbox was used for the SVR model [5], while neural network toolbox in Matlab was used for constructing neural networks, including SLFN and EMD based SLFN (EMD-SLFN). Moreover, the Kernel Methods Toolbox for Matlab was used for KRR and the proposed EMD-KRR-SVR approach [31].

For SVR and EMD based SVR, 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 SLFN and EMD-SLFN, the size of neural networks is determined by the size of input vector. The number of iterations for back propagation is set as 1000. We choose Gaussian kernel as the kernel in KRR. The regularization constant is searched within the range $[10^{-8}, 10^8]$ with the stepsize of $10^{0.2}$; while the range of Gaussian kernel width is $[10^{-4}, 10^4]$ with the same stepsize.

4.3 Error Measurement

In this paper, Root Mean Square Error (RMSE) is used to evaluate the performance of learning models. It is defined as

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i' - y_i)^2} \tag{8}$$

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

5 Results and Discussion

In this section, six benchmark methods were implemented for electricity price forecasting to perform a comparison with the proposed EMD-KRR-SVR model.

5.1 Performance comparison for short-term electricity price forecasting

In this work, 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, which has good accuracy due to the highly periodic characteristic of electricity price TS. The prediction results for short-term electricity price forecasting are shown in Table 1, where the forecasting horizon is half an hour. 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 electricity price forecasting.

To reveal the advantages of EMD based ensemble methods, we implemented the single structure models SVR, SLFN and KRR for EPF, and conducted an comparison with their EMD hybrid models. Moreover, all of the EMD based ensemble methods have the best performance cases, which shows that they have comparable performance with each other. However, the proposed EMD-KRR-SVR achieves the best performance in most cases, which means that the proposed method has more advantages compared with the benchmark models.

Table 1: Prediction results for half-an-hour ahead electricity price forecasting ($/MWh)

Dataset	Month	Prediction model						
		Persistence	SVR [8]	SLFN [16]	KRR [17]	EMD-SVR [33]	EMD-SLFN [22]	Proposed
NSW	Jan	20.585	18.410	20.203	19.991	12.681	12.409	**12.356**
	Apr	34.512	30.004	32.566	32.994	25.120	22.131	**20.383**
	Jul	27.387	24.342	24.792	25.741	**19.101**	19.723	20.472
	Oct	19.336	16.729	18.058	18.963	12.345	12.767	**11.729**
TAS	Jan	22.403	21.184	21.757	21.831	18.303	**16.497**	18.045
	Apr	23.395	20.856	21.163	22.544	20.394	**19.196**	19.734
	Jul	25.636	23.278	24.797	23.104	**14.839**	15.891	15.615
	Oct	16.185	15.718	15.967	15.794	12.289	12.980	**12.131**
QLD	Jan	335.873	240.409	268.652	241.917	229.549	232.249	**227.077**
	Apr	31.803	30.582	30.426	31.469	20.748	23.367	**20.719**
	Jul	28.837	26.235	15.258	26.788	**19.160**	19.636	20.453
	Oct	21.221	18.693	23.697	20.098	13.246	14.985	**12.813**

In order to give a detailed analysis of these results, we employ Friedman test [14] and Nemenyi post-hoc test [23] 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 is shown in Figure 2. 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:

$$CD = q_\alpha \sqrt{\frac{k(k+1)}{6N}} \tag{9}$$

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 $\sqrt{2}$ [10]. From the statistical test results, the proposed EMD-KRR-SVR achieves the best rank and significantly outperforms the non-EMD based methods with a 95% confidence.

5.2 Computation time comparison

Figure 3 shows the computation time of benchmark methods for load demand forecasting in Tasmania (TAS). Obviously, the computational speed of KRR is superior than SLFN and SVR. SVR requires a grid search on C and σ, and SLFN is iteratively tuned by BP algorithm to convergence to the optimal weights. These repetitive parameter tuning processes cause SLFN and SVR less efficient than KRR, which has closed form solutions.

6 Conclusion

In this paper, we proposed an ensemble kernel machines for short-term electricity price forecasting composed of EMD, KRR and SVR. The electricity price signal was first decomposed into several intrinsic mode functions (IMFs) by EMD, followed by a KRR which was used to

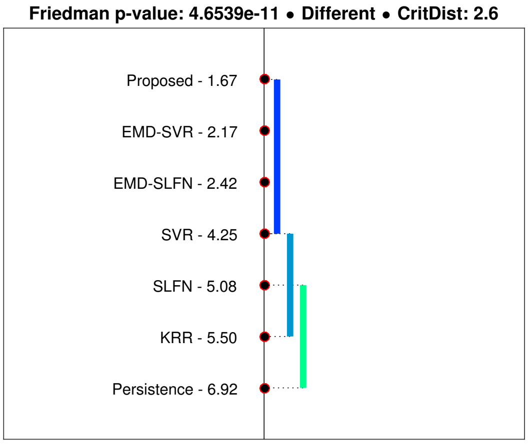

Figure 2: Nemenyi test for electricity price forecasting based on RMSE. The critical distance is 2.6.

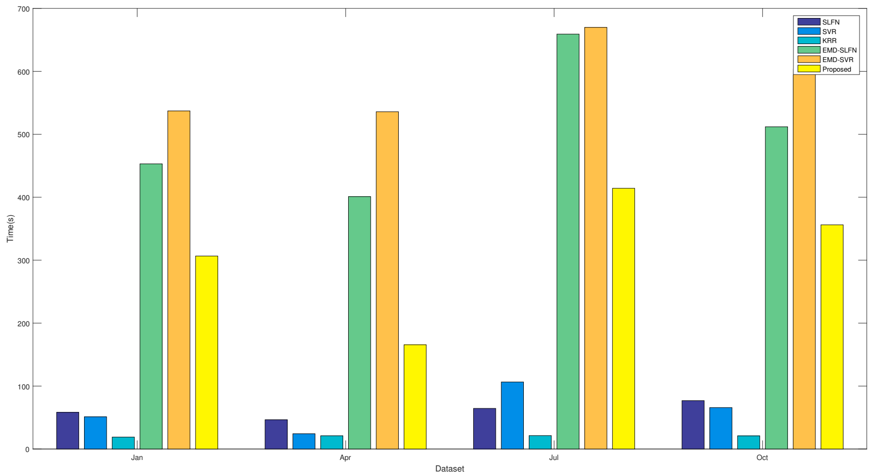

Figure 3: Computation time of learning models for electricity price forecasting in Tasmania (TAS)

model each extracted IMF and predict the tendencies. Finally, the prediction results of all IMFs were combined by an SVR to obtain an aggregated output for electricity price. Three electricity price datasets from AEMO 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. EMD based hybrid methods, including EMD-SVR, EMD-SLFN and the proposed EMD-KRR-SVR, significantly outperform the corresponding single structure models for short-

term electricity price time series forecasting.

2. The computation time of KRR is the shortest among all of the benchmark models.
3. The proposed EMD-KRR-SVR approach achieves the best performance for short-term electricity price forecasting, and also has the advantages of efficiency.

For future research directions, more ensemble methods, such as bagging, can be combined with KRR to make use of its advantages in computation time. Moreover, the concept of deep learning can also be applied with KRR to develop deep kernel machines. Further, Kernel machines and its ensemble models can also be tested using other types of TS, such as financial data and 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] Australian energy market operator, Dec. 2016.

[2] M. Alhamdoosh and D. Wang. Fast decorrelated neural network ensembles with random weights. *Information Sciences*, 264:104–117, 2014.

[3] G. E. P. Box and G. M. Jenkins. *Time series analysis: forecasting and control*. Holden-Day series in time series analysis and digital processing. Holden-Day, 1976.

[4] L. Breiman. Stacked regressions. *Machine Learning*, 24:49–64, 1996.

[5] C.-C. Chang and C.-J. Lin. LIBSVM: a library for support vector machines. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 2(3):27, 2011.

[6] J. Che and J. Wang. Short-term electricity prices forecasting based on support vector regression and auto-regressive integrated moving average modeling. *Energy Conversion and Management*, 51:1911–1917, 2010.

[7] T. H. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein. *Introduction to Algorithms*. MIT Press, 2000.

[8] C. Cortes and V. Vapnik. Support-vector networks. *Machine learning*, 20(3):273–297, 1995.

[9] G. A. Darbellay and M. Slama. Forecasting the short-term demand for electricity: Do neural networks stand a better chance? *International Journal of Forecasting*, 16:71–83, 2000.

[10] J. Demšar. Statistical comparisons of classifiers over multiple data sets. *Journal of Machine Learning Research*, 7:1–30, 2006.

[11] T. G. Dietterich. Ensemble methods in machine learning. In *Multiple classifier systems*, pages 1–15. Springer, 2000.

[12] H. Drucker, C. J. Burges, L. Kaufman, A. Smola, and V. Vapnik. Support vector regression machines. *Advances in neural information processing systems*, 9:155–161, 1997.

[13] M. D. Felice and X. Yao. Short-term load forecasting with neural network ensembles: a comparative study [application notes]. *IEEE Computational Intelligence Magazine*, 6:47–56, 2011.

[14] M. Friedman. The use of ranks to avoid the assumption of normality implicit in the analysis of variance. *Journal of the American Statistical Association*, 32(200):675–701, 1937.

[15] C. Guan, P. B. Luh, L. D. Michel, Y. Wang, and P. B. Friedland. Very short-term load forecasting: wavelet neural networks with data pre-filtering. *IEEE Transactions on Power Systems*, 28(1):30–41, 2013.

[16] S. Haykin. *Neural Networks: A Comprehensive Foundation*. International edition. Prentice Hall, 1999.

[17] T. Hofmann, B. Schlkopf, and A. J. Smola. Kernel methods in machine learning. *Annals of Statistics*, 36(3):1171–1220, 2008.

[18] C. C. Holt. Forecasting seasonals and trends by exponentially weighted moving averages. *International Journal of Forecasting*, 20:5–10, 2004.

[19] R.-A. Hooshmand, H. Amooshahi, and M. Parastegari. A hybrid intelligent algorithms based short-term load forecasting approach. *International Journal of Electrical Power & Energy Systems*, 45:313–324, 2013.

[20] N. E. Huang, Z. Shen, S. R. Long, M. C. Wu, H. H. Shih, Q. Zheng, N.-C. Yen, C. C. Tung, and H. H. Liu. The empirical mode decomposition and the hilbert spectrum for nonlinear and non-stationary time series analysis. In *Roy. Soc. London A*, volume 454, pages 903–995, 1998.

[21] R. Kohavi and F. Provost. Glossary of terms. *Machine Learning*, 30:271–274, 1998.

[22] H. Liu, C. Chen, H. Tian, and Y. Li. A hybrid model for wind speed prediction using empirical mode decomposition and artificial neural networks. *Renewable Energy*, 48:545–556, 2012.

[23] P. Nemenyi. *Distribution-free Multiple Comparisons*. Princeton University, 1963.

[24] J. Olamaee, M. Mohammadi, A. Noruzi, and S. M. H. Hosseini. Day-ahead price forecasting based on hybrid prediction model. *Complexity*, 21(S2):156–164, 2016.

[25] A. D. Papalexopoulos and T. C. Hesterberg. A regression-based approach to short-term system load forecasting. *IEEE Transactions on Power Systems*, 5:1535–1547, 1990.

[26] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830, 2011.

[27] X. Qiu, L. Zhang, Y. Ren, P. N. Suganthan, and G. Amaratunga. Ensemble deep learning for regression and time series forecasting. In *Proc. IEEE Symposium on Computational Intelligence in Ensemble Learning (CIEL)*, pages 1–6, 2014.

[28] Y. Ren, P. N. Suganthan, N. Srikanth, and G. Amaratunga. Random vector functional link network for short-term electricity load demand forecasting. *Information Sciences*, 000:1–16, 2016.

[29] Y. Ren, L. Zhang, and P. Suganthan. Ensemble classification and regression-recent developments, applications and future directions [review article]. *IEEE Computational Intelligence Magazine*, 11(1):41–53, 2016.

[30] 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*, 38(3):350–362, 2014.

[31] S. Van Vaerenbergh. *Kernel methods for nonlinear identification, equalization and separation of signals*. PhD thesis, University of Cantabria, Feb. 2010. Software available at `https://github.com/steven2358/kmbox`.

[32] R. Weron. Electricity price forecasting: a review of the state-of-the-art with a look into the future. *International Journal of Forecasting*, 30:1030–1081, 2014.

[33] L. Ye and P. Liu. Combined model based on emd-svm for short-term wind power prediction. In *Proc. Chinese Society for Electrical Engineering (CSEE)*, volume 31, pages 102–108, 2011.

[34] L. C. Ying and M. C. Pan. Using adaptive network based fuzzy inference system to forecast regional electricity loads. *Energy Conversion and Management*, 49:205–211, 2008.

[35] L. Zhang and P. N. Suganthan. Robust visual tracking via co-trained kernelized correlation filters, 2015.