Research Article
A Simple Hybrid Model for Short-Term Load Forecasting

Suseelatha Annamareddi,[1] Sudheer Gopinathan,[1] and Bharathi Dora[2]

[1] *Mathematics Department, GVP College of Engineering for Women, Visakhapatnam 530048, India*
[2] *Engineering Mathematics Department, A.U College of Engineering, Visakhapatnam 530003, India*

Correspondence should be addressed to Sudheer Gopinathan; g_sudheer@hotmail.com

The paper proposes a simple hybrid model to forecast the electrical load data based on the wavelet transform technique and double exponential smoothing. The historical noisy load series data is decomposed into deterministic and fluctuation components using suitable wavelet coefficient thresholds and wavelet reconstruction method. The variation characteristics of the resulting series are analyzed to arrive at reasonable thresholds that yield good denoising results. The constitutive series are then forecasted using appropriate exponential adaptive smoothing models. A case study performed on California energy market data demonstrates that the proposed method can offer high forecasting precision for very short-term forecasts, considering a time horizon of two weeks.

1. Introduction

Short-term load forecasting (STLF) is an integral part of power system operations as it is essential for ensuring supply of electrical energy with minimum expenses. This type of forecasting is used to predict load demands from an hour to a week ahead so that the day-to-day operations of a power system can be efficiently planned to minimize the operational costs [1]. The ability to forecast accurately the load even a few hours ahead is beneficial from different points of view, ranging from technical to commercial [2]. This importance has led to the development of several mathematical models/techniques. These techniques can be broadly classified as (i) classical time series model and (ii) machine intelligence models. The classical time series models have been criticized by researchers for their weakness of nonlinear fitting capability [3]. Moreover, they require huge amounts of data to arrive at optimal models and perform poorly for data with inherent special events [4–6]. The artificial neural network (ANN) is one of the most popular machine intelligence models [7] that have been used for short-term load forecasting with encouraging results [8–10]. An excellent review of neural networks for short-term load forecasting has been presented in [11]. However, no single model has performed well in short term load forecasting [12]. This has led to the development of hybrid models that try

to deduct the best features of different models and integrate them to achieve good forecasting results [13–15]. Hybrid models include a time series and ANN-based model as in [16], a combination of ANN and fuzzy expert systems as in [17], and integration of wavelet-based approach with ANN as in [18]. It is to be noted that the aim of any forecasting model for STLF is to identify the different characteristics of load over different time horizons and incorporate them into the model [19]. It is here that wavelet multiresolution analysis is found to be useful. The wavelet transforms are found to capture all the information in a time series and associate it with specific time horizons and locations in time and thus help to unfold the inner load characteristics that are useful for more precise forecasting [20, 21]. This paper describes a hybrid model based on the discrete wavelet transform for short-term load forecasting. The idea is to combine the wavelet transform technique with the exponential smoothing method that is intuitively appealing, easy to update, and has minimal computer storage requirements [22]. Noisy characteristics in the power load data will affect the forecasting precision, so it is first preprocessed using the discrete wavelet transform according to the methodology outlined in Section 3. The proposed model is verified by forecasting the electrical load of California energy market. The hybrid model is found to provide good forecasting precision for four-hour-ahead forecasts.

2. Theoretical Background

In modeling the time series data, noise and nonstationarity are two problems that a forecasting methodology should take care of. The presence of noisy characteristics in the data prevents the full capture of the dependency between the past and the future behavior of the time series. The non-stationarity of data infers that the time series switch their dynamics among different regions [23]. Thus, in the time series data, one observes that the intrinsic property of the data (deterministic feature) is superimposed by rapid variations on much shorter time scales. Therefore, it seems appropriate to separate the fluctuating component from the deterministic component and wavelet decomposition is an excellent technique that can separate the two.

2.1. Discrete Wavelet Transform [24, 25]. A dyadic discretized wavelet is given by

$$\psi_{m,n}(t) = 2^{-m/2}\psi\left(2^{-m}t - n\right), \tag{1}$$

where the control parameters $m, n \in Z$. The discrete wavelet transform (DWT) of a signal $x(t)$ can be written as

$$T_{m,n} = \int_{-\infty}^{\infty} x(t)\,\psi_{m,n}(t)\,dt, \tag{2}$$

where $T_{m,n}$ is known as the wavelet (or detail) coefficient at scale and location indices m, n. The dyadic grid wavelet leads to the construction of an orthonormal wavelet basis. They are associated with scaling function and their dilation equations. The scaling function is associated with the smoothing of the signal and is given by

$$\phi_{m,n}(t) = 2^{-m/2}\phi\left(2^{-m}t - n\right). \tag{3}$$

The scaling function can be convolved with the signal to produce approximation coefficients as

$$S_{m,n} = \int_{-\infty}^{\infty} x(t)\,\phi_{m,n}(t)\,dt. \tag{4}$$

The multiresolution decomposition theorem [26] gives a method of generating approximation coefficients and detail wavelet coefficients at different scales. The approximation and wavelet coefficients at scale index $m+1$ can be generated using the coefficients at the previous scale as

$$S_{m+1,n} = \frac{1}{\sqrt{2}}\sum_{k} c_k S_{m,2n+k} = \frac{1}{\sqrt{2}}\sum_{k} c_{k-2n} S_{m,k},$$

$$T_{m+1,n} = \frac{1}{\sqrt{2}}\sum_{k} b_k S_{m,2n+k} = \frac{1}{\sqrt{2}}\sum_{k} b_{k-2n} S_{m,k}. \tag{5}$$

In general, the discrete input signal is taken to be signal approximation coefficients at scale index $m = 0$. For a signal of length $N = 2^M$ the range of scales that can be investigated is $0 < m < M$, and the input signal can be expressed as

$$x_0(t) = x_M(t) + \sum_{m=1}^{M} d_m(t), \tag{6}$$

where the mean signal approximation at scale M is

$$x_M(t) = S_{M,n}\phi_{M,n}(t), \tag{7}$$

and detail signal approximation corresponding to scales index m is given by

$$d_m(t) = \sum_{m=0}^{2^{n-m}-1} T_{m,n}\psi_{m,n}(t), \tag{8}$$

that is, adding the approximation of the signal at scale index M to the sum of all detail signal components across scales $0 < m < M$ gives the approximation of the original signal at scale index 0.

2.2. Wavelet Denoising. It is generally assumed that the noisy part of a signal shows random characteristics and mainly reflects the inherent uncertainties in the data, while the original signal is composed of deterministic components and mainly reflects the deterministic characteristics of the data studied. When a discrete wavelet transform (DWT) is applied to a time series data, small wavelet coefficients are presumed to be dominated by noises and carry little useful information, but the original series carry useful information and are concentrated in a limited number of wavelet coefficients [27]. Wavelet shrinkage denoising is not to be confused with smoothing. Smoothing removes high frequencies and retains low ones whereas denoising attempts to remove whatever noise is present regardless of the signal frequency content [28]. Wavelet shrinkage denoising involves shrinking in the wavelet transform domain and consists of three essential steps: a linear forward wavelet transform, a nonlinear shrinkage denoising, and a linear inverse wavelet transform. If the observed data $X(t)$ is assumed to be of the form

$$X(t) = S(t) + N(t), \tag{9}$$

where $S(t)$ is true signal and $N(t)$ is noise, then the denoising procedure can be outlined as follows.

Let $W(\cdot)$ and $W^{-1}(\cdot)$ denote the forward and inverse wavelet operators, let $D(\cdot, \lambda)$ denote the denoising operator with soft thresholding λ, and, $\widehat{S}(t)$ be an estimate of $S(t)$. The three steps

$$Y = W(X),$$
$$Z = D(Y, \lambda), \tag{10}$$
$$\widehat{S} = W^{-1}(Z)$$

summarize the procedure.

Given threshold λ for data U, the rule

$$D(U, \lambda) = \text{sgn}(U)\max(0, |U| - \lambda) \tag{11}$$

defines nonlinear soft thresholding. Setting the threshold λ is an essential part of denoising. Different thresholding techniques have been provided in [29, 30]. A hybrid method called heursure [31] that determines at each multiresolution level a threshold that is either universal or Stein's unbiased risk estimator (SURE) is used in the present work.

2.3. Exponential Smoothing. Exponential smoothing is a pragmatic approach for forecasting wherein the prediction is constructed from an exponentially weighted average of past observations [32]. Its robustness and accuracy have led to its widespread use in a variety of applications. In the present work, the wavelet preprocessed data is used for prediction using double exponential smoothing. The specific formula for double exponential smoothing with additive trend and no seasonality [33, 34] is

$$S_t = \alpha X_t + (1 - \alpha)(S_{t-1} + T_{t-1}), \tag{12}$$

$$T_t = \gamma(S_t - S_{t-1}) + (1 - \gamma)T_{t-1}; \quad \alpha, \gamma \in [0\ 1], \tag{13}$$

$$\widehat{X}_t(m) = S_t + mT_t. \tag{14}$$

Here α is the smoothing parameter for the level of the series, γ is the smoothing parameter for the trend, S_t is the smoothed level of the series computed after X_t is observed, T_t is the smoothed additive trend at the end point of the period t, X_t is the observed value of the time series at period t, $\widehat{X}_t(m)$ is the forecast for m periods ahead from origin t. The double exponential smoothing methodology is used in the present work to forecast the constitutive series obtained via Haar wavelet decomposition.

3. Methodology

The DWT provides a handy tool for decomposing the time series into deterministic and fluctuation components and wavelet denoising method is a means of achieving the decomposition. The wavelet threshold denoising method is influenced by several key issues such as choice of wavelet, choice of decomposition level, threshold estimation, and thresholding rule [35]. In the present work, the Haar wavelet transform is used for the decomposition due to the following characteristics of the wavelet function.

(1) As the real world data is collected periodically, it is easy to assume that they are piecewise constant functions [20], and Haar wavelet transform is the most appropriate for these functions.

(2) The Haar wavelet transform is not affected by border distortions when one performs filtering/convolution on finite length signals and thus eliminates the boundary effect that is a problem with other wavelet transforms.

With the choice of Haar wavelet transform, the best resolution level is tested and it is observed that resolution level three is optimum for this case study. There are several threshold estimation methods such as universal threshold algorithm, minimax algorithm, Stein's unbiased risk estimation (SURE) [29, 30] algorithm, and heuristic SURE algorithm [31] that can be used to select a threshold. The denoising results are analyzed to see if the denoising results are good [36] and hence determine the optimum threshold. It is observed here that heursure threshold selection criteria gave the best denoising results with multilevel thresholding. With the key issues of wavelet denoising method being fixed, the proposed

methodology is given as a flowchart in Figure 1 and is outlined as follows.

(1) Decompose the available load series (X) via the Haar wavelet transform into one approximation series denoted as a_3 and three detail series denoted as d_1, d_2, d_3.

(2) The wavelet coefficients in the detail series are thresholded using Heursure threshold selection criteria to obtain the new detail series $\widehat{d}_1, \widehat{d}_2, \widehat{d}_3$.

(3) The approximation series a_3 and detail series $\widehat{d}_1, \widehat{d}_2, \widehat{d}_3$ are used by the inverse Haar wavelet transform to construct the deterministic series (X_t). The fluctuation series (X_f) is obtained by $X_f = X - X_t$.

(4) The deterministic series $X_t = (x_1, x_2, \ldots, x_n)$ is now forecasted using double exponential smoothing method through (12)–(14). The strategy is summarized in the following steps:

    (i) the initial values for S_t and T_t are chosen as $S_1 = x_1$; $T_1 = (x_n - x_1)/(n - 1)$

    (ii) the parameter α is varied from 0 to 1 in steps of 0.05 and γ is varied from 1 to 0 in decreasing steps of 0.05. For every possible combination of α and γ, the four-step-ahead forecast is computed. The best parameters are those resulting in the smallest four-step-ahead forecasting errors. The minimum value of the mean squared forecast error determines the optimum values of the parameters. With these parameters the four hour ahead deterministic component $(\widehat{X}_t)$ is forecast.

(5) The mean corrected fluctuation series (subtracting the sample mean) is modeled using the double exponential smoothing method in a procedure similar to the one outlined in step (4) to forecast the four-hour-ahead fluctuation component $(\widehat{X}_f)$ of the load series.

(6) The four-hour ahead load forecast is now obtained as $\widehat{X} = \widehat{X}_t + \widehat{X}_f$.

4. Results and Discussions

4.1. Case Study: California Energy Market. The hourly load data of California energy market [37] is considered in this real world case study. The system-wise load for every hour of the year 2000 is considered in this work. To build the forecasting model for each of the considered days, the information available to the method is the hourly load data of the 14 days (2 weeks) previous to the day of the week whose loads are to be forecasted. The forecasting has been done in two time scales, that is, four hours ahead and one day ahead. The behavior of the proposed technique is illustrated by providing results corresponding to different seasons (summer, spring, and winter) of the year. The four hour-ahead forecast, of the 15th, 22nd, 29th, 36th, 43rd, 50th, 57th, 64th, 71st, 78th, 85th, 92nd, 99th, and 106th days in each of the seasons are
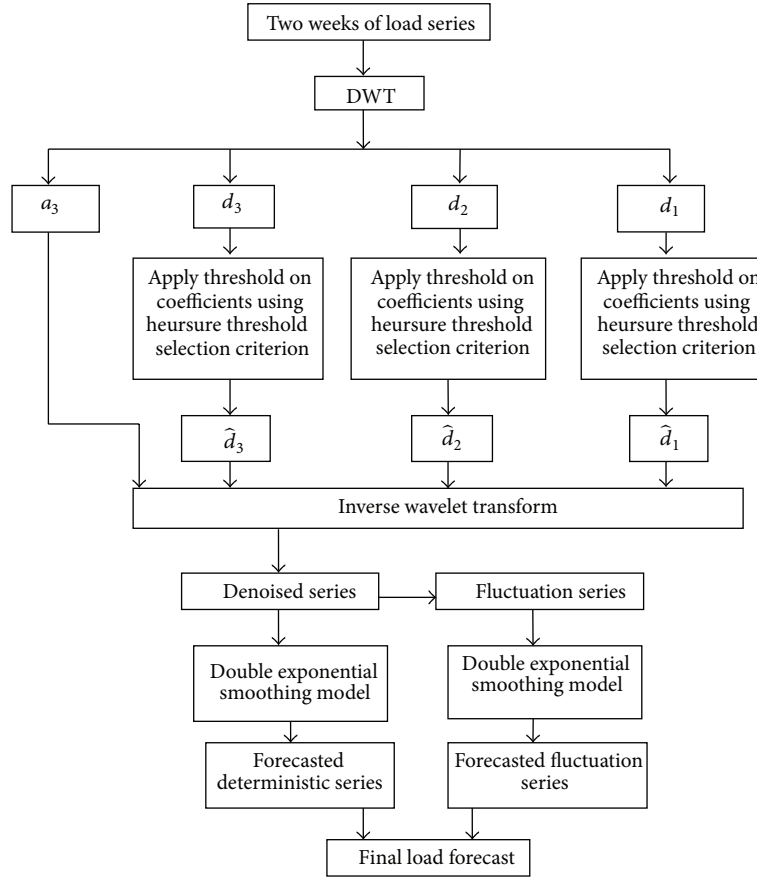
FIGURE 1: Flow chart of the algorithm.

presented. The one-day forecasts are carried out using an adaptive scheme. The 24-hour forecast (one day) is obtained by considering a moving window of 336 hours (2 weeks) prior to the four hours whose load is to be forecast and shifting the window by four hours until the entire day is covered. The forecasted on different test days (one in each season) is considered. The considered days are Jan 15th (Saturday a Holiday), April 11th (Tuesday a working day), and July 14th (Friday a day close to weekend) for a fair comparison.

4.2. Results. The forecast methodology outlined in the previous section is performed on the data considered. The forecast accuracy is examined using the two different evaluation metrics: the root mean square error (RMSE) and the mean absolute percentage error (MAPE). They are defined as follows:

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(\text{Actual load}_i - \text{Forecasted load}_i\right)^2},$$

$$\text{MAPE} = \frac{100}{n}\sum_{i=1}^{n}\frac{\left|\text{Actual load}_i - \text{Forecasted load}_i\right|}{\text{Actual load}_i}. \tag{15}$$

The hourly load data for the year 2000 is shown in **Figure 2**.

Observe the unstable mean and variance present in this series. This behavior makes forecasting a difficult task. The

TABLE 1: Forecasting accuracy metrics for 24-hour-ahead forecast of a typical day in each seasons.

Evaluation metric	Model	24-hr-ahead adaptive forecast		
		July 14th	April 11th	Jan 15th
RMSE	DEM	1.2591	1.0554	0.8334
	Proposed model	0.2520	0.4756	0.5390
MAPE	DEM	2.9645	2.7862	2.7722
	Proposed model	0.7438	1.3075	1.4501

RMSE: root mean square error; MAPE: mean absolute percentage error; DEM: double exponential smoothing model.

original series, denoised series, and fluctuation series of two weeks in summer, spring, and winter are shown in Figures 3(a), 3(b), and 3(c), respectively.

The values for the RMSE and MAPE metrics to evaluate the accuracy of the proposed methodology in forecasting the day-ahead forecasts for the days in different time periods and different seasons are presented in Table 1. The results are compared with the double exponential smoothing method (DEM). The parameters of the adaptively fitted DEM model to forecast the 24 hrs load data of the days January 15, April 11, and July 14 using the wavelet decomposed series of the load data are given in Table 2. The days considered include a holiday, a working day, and a day close to the weekend in different seasons. This has been done to show
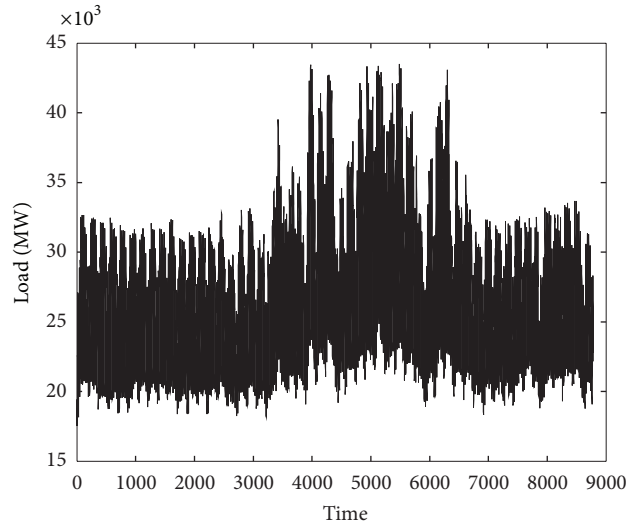
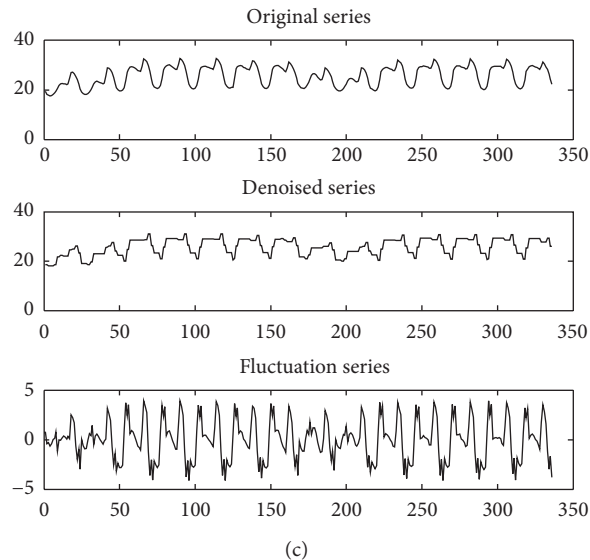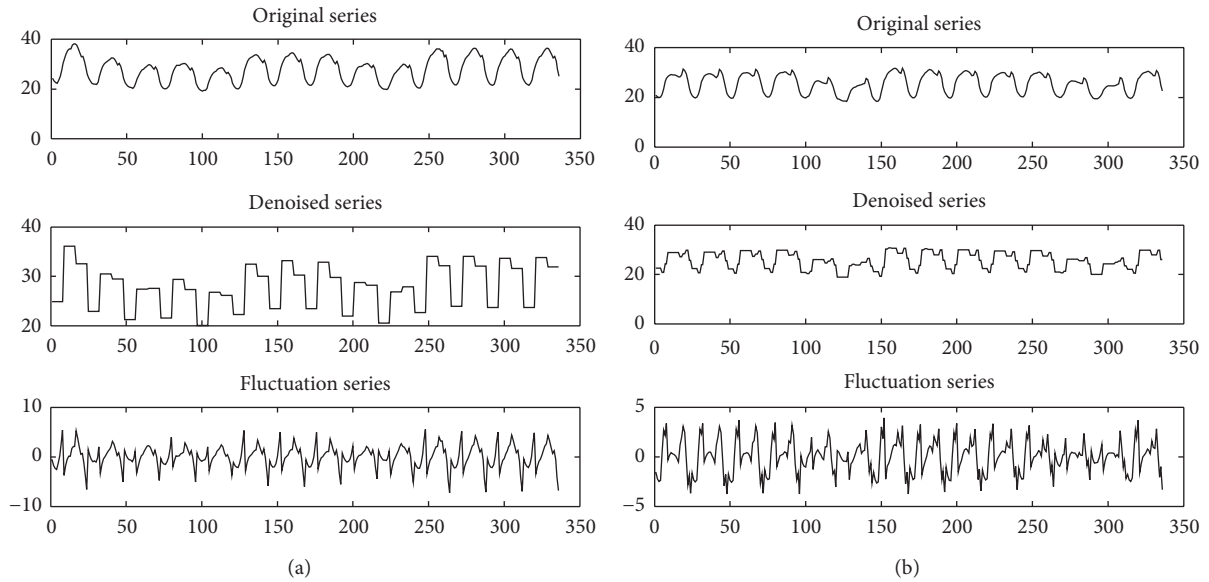FIGURE 2: Hourly load data of California energy market in the year 2000.

FIGURE 3: (a) Original series, denoised series, and fluctuation series of two weeks in summer. (b) Original series, denoised series, and fluctuation series of two weeks in spring. (c) Original series, denoised series, and fluctuation series of two weeks in winter.

TABLE 2: The parameters of the adaptively fitted double exponential smoothing models.

| Day | Period of hourly forecast | Parameters | | | |
| | | Deterministic part | | Fluctuation part | |
		α	γ	α	γ
January 15th	01.00 AM–04.00 AM	0.15	0.92	0.95	0.04
	05.00 AM–08.00 AM	0.3	1	0.91	0.58
	09.00 AM–12.00 PM	0.1	0.05	0.98	0.43
	12.00 PM–04.00 PM	0.3	0.65	0.34	0
	05.00 PM–08.00 PM	0.25	0.5	0.83	0.09
	09.00 PM–12.00 AM	0.3	1	0.14	1
April 11th	01.00 AM–04.00 AM	0.15	0.95	1	0.03
	05.00 AM–08.00 AM	0.3	1	0.33	0
	09.00 AM–12.00 PM	0.1	0.2	0.43	0.72
	12.00 PM–04.00 PM	0.2	1	0.29	0.29
	05.00 PM–08.00 PM	0.75	0.8	0.04	0.08
	09.00 PM–12.00 AM	0.1	0.95	0.13	0.95
July 14th	01.00 AM–04.00 AM	0.2	1	0.84	0.29
	05.00 AM–08.00 AM	0.25	1	0.61	1
	09.00 AM–12.00 PM	0.1	0.25	0.44	0.44
	12.00 PM–04.00 PM	0.8	0.15	0.62	0.98
	05.00 PM–08.00 PM	0.3	0.8	0.68	0.05
	09.00 PM–12.00 AM	0.6	0.5	0.13	0.99

the effectiveness of the model in handling special events/days. The evaluation metrics of the four-hour-ahead forecasts on different days (mentioned in the case study) for the three seasons of the proposed model are presented in Table 3. The results presented confirm the assumption that wavelet shrinkage denoising produces constitutive (deterministic and fluctuating) series that can be predicted more accurately by the exponential smoothing methods.

5. Conclusion

This paper proposes a technique of applying the simple Haar wavelet transform in a hybrid model that utilizes the double exponential smoothing model for short-term load forecasting. The heursure wavelet thresholding method is used to separate the load series into deterministic and fluctuation (noisy) components. The problems associated with wavelet denoising such as selection of appropriate wavelet function, suitable level of decomposition, and threshold selection are handled to a certain extent by analyzing the load data. It is seen that the trend in both the components can be taken care of effectively using double exponential smoothing models. The results demonstrate that the proposed wavelet model is advantageous in reducing the complex structure of load series and thereby enhancing the accuracy of short-term forecasts.

TABLE 3: Forecasting accuracy metrics for 4-hr-ahead forecast.

| Day | Proposed method | | DEM | |
	RMSE	MAPE	RMSE	MAPE
15th				
Summer	0.2657	1.2575	0.4448	2.2143
Spring	0.1934	0.8417	0.4359	0.8325
Winter	0.1623	0.7955	0.4017	1.8907
22nd				
Summer	0.1885	0.9445	0.5145	2.4405
Spring	0.1521	0.6962	0.4598	1.9832
Winter	0.1794	0.8262	0.4392	2.0424
29th				
Summer	0.1467	0.7307	0.4710	2.2014
Spring	0.1802	0.7902	0.4809	1.9876
Winter	0.2787	1.305	0.5234	2.4684
36th				
Summer	0.1712	0.8445	0.5418	2.4261
Spring	0.2558	1.1315	0.5949	2.5174
Winter	0.128	0.6315	0.4594	2.0987
43rd				
Summer	0.1511	0.7337	0.5046	2.2080
Spring	0.2406	1.0435	0.6057	2.5245
Winter	0.2288	1.0105	0.4926	2.2277
50th				
Summer	0.1290	0.6277	0.4629	0.9494
Spring	0.2646	1.1047	0.6031	2.3827
Winter	0.1725	0.7425	0.4728	2.0265
57th				
Summer	0.1761	0.799	0.4430	2.0739
Spring	0.2568	1.0695	0.5661	2.2892
Winter	0.1511	0.7272	0.2321	0.9605
64th				
Summer	0.2265	1.0865	0.4518	2.0590
Spring	0.1937	0.836	0.5109	2.1099
Winter	0.1790	0.9077	0.4198	1.9776
71st				
Summer	0.1029	0.5025	0.5640	2.4710
Spring	0.2164	0.925	0.5596	2.2997
Winter	0.1698	0.8467	0.5577	2.6100
78th				
Summer	0.2512	1.2687	0.5116	2.3770
Spring	0.1573	0.7472	0.3502	1.5144
Winter	0.1781	0.8737	0.3948	1.7977
85th				
Summer	0.2071	0.969	0.4885	2.1521
Spring	0.2105	0.983	0.6110	2.6769
Winter	0.1631	0.797	0.3001	1.3905
92nd				
Summer	0.1225	0.5535	0.4293	1.8439
Spring	0.2168	0.9395	0.4838	2.0605
Winter	0.1736	0.7222	0.3972	1.6758
99th				
Summer	0.1945	0.9087	0.6095	2.5899
Spring	0.1505	0.7075	0.4343	1.9034
Winter	0.2741	1.2167	0.4307	1.9024
106th				
Summer	0.1785	0.8467	0.4720	2.1271
Spring	0.1677	0.783	0.5330	2.1665
Winter	0.1809	0.914	0.2348	1.0906

References

[1] S. Adbel-hady Soliman and A. M. Al-Kandari, *Electrical Load Forecasting: Modeling and Model Construction*, Butterworth-Heinemann, Boston, Mass, USA, 2010.

[2] L. A. Aguirre, D. D. Rodrigues, S. T. Lima, and C. B. Martinez, "Dynamical prediction and pattern mapping in short-term load forecasting," *International Journal of Electrical Power and Energy Systems*, vol. 30, no. 1, pp. 73–82, 2008.

[3] J. Wang, L. Li, D. Niu, and Z. Tan, "An annual load forecasting model based on support vector regression with differential evolution algorithm," *Applied Energy*, vol. 94, pp. 65–70, 2012.

[4] I. Moghram and S. Rahman, "Analysis and evaluation of five short-term load forecasting techniques," *IEEE Transactions on Power Systems*, vol. 4, no. 4, pp. 1484–1491, 1989.

[5] S. R. Huang, "Short-term load forecasting using threshold autoregressive models," *IEE Proceedings on Generation, Transmission and Distribution*, vol. 144, no. 5, pp. 477–481, 1997.

[6] C.-I. Kim, I.-K. Yu, and Y. H. Song, "Kohonen neural network and wavelet transform based approach to short-term load forecasting," *Electric Power Systems Research*, vol. 63, no. 3, pp. 169–176, 2002.

[7] N. S. Sisworahardjo, A. A. El-Keib, J. Choi, J. Valenzuela, R. Brooks, and I. El-Agtal, "A stochastic load model for an electricity market," *Electric Power Systems Research*, vol. 76, no. 6-7, pp. 500–508, 2006.

[8] W. Charytoniuk, E. Don Box, W.-J. Lee, M.-S. Chen, P. Kotas, and P. Van Olinda, "Neural-network-based demand forecasting in a deregulated environment," *IEEE Transactions on Industry Applications*, vol. 36, no. 3, pp. 893–898, 2000.

[9] T. Senjyu, H. Takara, K. Uezato, and T. Funabashi, "One-hour-ahead load forecasting using neural network," *IEEE Transactions on Power Systems*, vol. 17, no. 1, pp. 113–118, 2002.

[10] K. Y. Lee, Y. T. Cha, and J. H. Park, "Short-term load forecasting using an artificial neural network," *IEEE Transactions on Power Systems*, vol. 7, no. 1, pp. 124–132, 1992.

[11] H. S. Hippert, C. E. Pedreira, and R. C. Souza, "Neural networks for short-term load forecasting: a review and evaluation," *IEEE Transactions on Power Systems*, vol. 16, no. 1, pp. 44–55, 2001.

[12] C.-M. Lee and C.-N. Ko, "Short-term load forecasting using lifting scheme and ARIMA models," *Expert Systems with Applications*, vol. 38, no. 5, pp. 5902–5911, 2011.

[13] A. A. El-Keib, X. Ma, and H. Ma, "Advancement of statistical based modeling techniques for short-term load forecasting," *Electric Power Systems Research*, vol. 35, no. 1, pp. 51–58, 1995.

[14] T. Zheng, A. A. Girgis, and E. B. Makram, "Hybrid wavelet-Kalman filter method for load forecasting," *Electric Power Systems Research*, vol. 54, no. 1, pp. 11–17, 2000.

[15] G. J. Chen, K. K. Li, T. S. Chung, H. B. Sun, and G. Q. Tang, "Application of an innovative combined forecasting method in power system load forecasting," *Electric Power Systems Research*, vol. 59, no. 2, pp. 131–137, 2001.

[16] A. A. El Desouky and M. M. Elkateb, "Hybrid adaptive techniques for electric-load forecast using ANN and ARIMA," *IEE Proceedings on Generation, Transmission and Distribution*, vol. 147, no. 4, pp. 213–217, 2000.

[17] K.-H. Kim, J.-K. Park, K.-J. Hwang, and S.-H. Kim, "Implementation of hybrid short-term load forecasting system using artificial neural networks and fuzzy expert systems," *IEEE Transactions on Power Systems*, vol. 10, no. 3, pp. 1534–1539, 1995.

[18] D. Benaouda, F. Murtagh, J.-L. Starck, and O. Renaud, "Wavelet-based nonlinear multiscale decomposition model for electricity load forecasting," *Neurocomputing*, vol. 70, no. 1–3, pp. 139–154, 2006.

[19] T. Nengling, J. Stenzel, and W. Hongxiao, "Techniques of applying wavelet transform into combined model for short-term load forecasting," *Electric Power Systems Research*, vol. 76, no. 6-7, pp. 525–533, 2006.

[20] R. Gencay, F. Selçuk, and B. Whitcher, *An Introduction to Wavelets and Other Filtering Methods in Finance and Economics*, Academic press, 2002.

[21] A. J. da Rocha Reis and A. P. Alves da Silva, "Feature extraction via multiresolution analysis for short-term load forecasting," *IEEE Transactions on Power Systems*, vol. 20, no. 1, pp. 189–198, 2005.

[22] R. D. Snyder, A. B. Koehler, R. J. Hyndman, and J. K. Ord, "Exponential smoothing models: means and variances for lead-time demand," *European Journal of Operational Research*, vol. 158, no. 2, pp. 444–455, 2004.

[23] L. Cao, "Support vector machines experts for time series forecasting," *Neurocomputing*, vol. 51, pp. 321–339, 2003.

[24] P. Addison, J. Walker, and R. Guido, "Time-frequency analysis of biosignals," *IEEE Engineering in Medicine and Biology Magazine*, vol. 28, no. 5, pp. 14–29, 2009.

[25] S. Mallat, *A Wavelet Tour of Signal Processing*, Academic Press, 3rd edition, 2008.

[26] S. G. Mallat, "Theory for multiresolution signal decomposition: the wavelet representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 11, no. 7, pp. 674–693, 1989.

[27] M. Jansen and A. Bultheel, "Asymptotic behavior of the minimum mean squared error threshold for noisy wavelet coefficients of piecewise smooth signals," *IEEE Transactions on Signal Processing*, vol. 49, no. 6, pp. 1113–1119, 2001.

[28] C. Taswell, "What, how, and why of wavelet shrinkage denoising," *Computing in Science and Engineering*, vol. 2, no. 3, pp. 12–19, 2000.

[29] R. Weron, *Modeling and Forecasting Electricity Load and Prices: A Statistical Approach*, The Wiley Finance series, 2006.

[30] D. L. Donoho and J. M. Johnstone, "Ideal spatial adaptation by wavelet shrinkage," *Biometrika*, vol. 81, no. 3, pp. 425–455, 1994.

[31] D. L. Donoho and I. M. Johnstone, "Adapting to unknown smoothness via wavelet shrinkage," *Journal of the American Statistical Association*, vol. 90, no. 432, pp. 1200–1224, 1995.

[32] M. Misiti, Y. Misiti, G. Oppenheim, and J.-M. Poggi, Wavelet Toolbox User's Guide.

[33] E. S. Gardner Jr., "Exponential smoothing: the state of the art-Part II," *International Journal of Forecasting*, vol. 22, no. 4, pp. 637–666, 2006.

[34] C. C. Holt, *Forecasting Trends and Seasonals by Exponentially Weighted Averages*, vol. 52, Carnegie Institute of Technology, Pittsburgh, Pa, USA, 1957.

[35] Y.-F. Sang, D. Wang, J.-C. Wu, Q.-P. Zhu, and L. Wang, "Entropy-based wavelet de-noising method for time series analysis," *Entropy*, vol. 11, no. 4, pp. 1123–1147, 2009.

[36] Y.-F. Sang, D. Wang, J.-C. Wu, Q.-P. Zhu, and L. Wang, "The relation between periods' identification and noises in hydrologic series data," *Journal of Hydrology*, vol. 368, no. 1–4, pp. 165–177, 2009.

[37] http://www.ucei.berkeley.edu/datamine/uceidata/.