Communications for Statistical Applications and Methods
2015, Vol. 22, No. 6, 675–683

DOI: http://dx.doi.org/10.5351/CSAM.2015.22.6.675
Print ISSN 2287-7843 / Online ISSN 2383-4757

A Study on the Comparison of Electricity Forecasting Models: Korea and China

Xueyan Zheng[a], Sahm Kim[1,a]

[a]Department of Applied Statistics, Chung-Ang University, Korea

Abstract

In the 21st century, we now face the serious problems of the enormous consumption of the energy resources. Depending on the power consumption increases, both China and South Korea face a reduction in available resources. This paper considers the regression models and time-series models to compare the performance of the forecasting accuracy based on Mean Absolute Percentage Error (MAPE) in order to forecast the electricity demand accurately on the short-term period (68 months) data in Northeast China and find the relationship with Korea. Among the models the support vector regression (SVR) model shows superior performance than time-series models for the short-term period data and the time-series models show similar results with the SVR model when we use long-term period data.

Keywords: ϵ-support vector regression (ϵ-SVR), regression, time series, electricity demand forecasting, mean absolute percentage error (MAPE)

1. Introduction

Industrial development and improved comfortable standards have led to the explosion of resource consumption and resource depletion. We have to match the consumption and demand of electricity because it is impossible to store. A mismatch between supply and demand will make will lead to "Black Out" situation that has a serious impact on the utility industry due to the lack of electricity.

In case of China, the 80% of the electricity power generation source is the thermal power. We can assume that the Chinese government may utilize nuclear power and solar power to expand the development of alternative measures as a new energy source if we consider the dramatically increasing demand from China over the past decades.

Economic development, social changes, industrial policy, and seasonal components are factors that influence electrical demand. Establishing a single prediction model that can be a factor in any influence on the account is not realistic (Wang *et al.*, 2009). Until now, forecasting models can be classified into several models: Holt-winters model (Taylor, 2003), Triple seasonal Holt-winters model (Taylor, 2010). Multiple linear regression model (Karin, 2011), ϵ-SVR model (Claveria *et al.*, 2015). These methods do not need a large amount of historical data and composed of three components: the trend, seasonal are components and stochastic errors. Yoon *et al.* (2009) studied the electricity patterns in Korea based monthly maximum load data.

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (NRF-2013R1A1A2008277).
[1] Corresponding author: Department of Applied Statistics, Chung-Ang University, 221, Heukseok-dong, Dongjak-gu, Seoul 06974, Korea. E-mail: sahm@cau.ac.kr

Published 30 November 2015 / journal homepage: http://csam.or.kr

In this paper, we focus on comparing the performance of proposed models. We find best performance models among the models at different data patterns and different periods of training data. This paper is organized as following: Section 2 introduces the proposed models; Section 3 shows the results of the performance evaluation different data; Section 4 has the conclusion in the paper.

2. Forecasting Models

2.1. Linear regression model

Linear regression model with trend and seasonality components can be written as Equation (2.1).

$$Z_t = \beta_0 + \beta_1 t + \beta_2 \sin\left(\frac{2\pi t}{12}\right) + \beta_3 \cos\left(\frac{2\pi t}{12}\right) + \epsilon_t. \tag{2.1}$$

Where $Z_t$is electricity demand here, t is sampling time, the unknown parameters $\beta = (\beta_0, \beta_1, \beta_2, \beta_3)'$ is regression coefficients and ϵ_t is model error at time t, and we assume $\epsilon_t \sim N(0, \sigma_\epsilon^2)$.

2.2. Holt-Winters model

2.2.1. Additive seasonal model

Winters' additive seasonal model at time $n+l$ was proposed in Winters (1960) and defined as Equation (2.2).

$$Z_{n+l} = T_{n+l} + S_{n+l} + I_{n+l}. \tag{2.2}$$

Note that I_{n+l} is error term, T_{n+l} is trend component and S_{n+l} is seasonality component with the period of s. If we assume linear trend of T_{n+l} as Equation (2.3),

$$T_{n+l} = \beta_{0,n} + \beta_{1,n}(n + l) = (\beta_{0,n} + \beta_{1,n}n) + \beta_{1,n}l = T_n + \beta_{1,n}l \tag{2.3}$$

additive seasonality component of s as $S_i = S_{i+s} = S_{i+2s} = \cdots$ $(i = 1, 2, \ldots, s)$ and $\sum_{i-0}^{s} S_i = 0$. Then Z_{n+l} the forecasting value of future time on n is

$$\begin{aligned}
\hat{Z}_n(l) &= \hat{T}_n + \hat{\beta}_{1,n}l + \hat{S}_{n+l-s}, \quad l = 1, 2, \ldots, s,\\
\hat{Z}_n(l) &= \hat{T}_n + \hat{\beta}_{1,n}l + \hat{S}_{n+l-2s}, \quad l = s + 1, s + 2, \ldots, 2s,\\
\hat{Z}_n(l) &= \hat{T}_n + \hat{\beta}_{1,n}l + \hat{S}_{n+l-3s}, \quad l = 2s + 1, 2s + 2, \ldots, 3s.
\end{aligned}$$

2.2.2. Multiple seasonal model

Winters' multiplicative seasonal model is defined as Equation (2.4)

$$\begin{aligned}
Z_{n+l} &= T_{n+l} + S_{n+l} + I_{n+l}\\
&= (T_n + \beta_{1,n}l) S_{n+l} + I_{n+l}.
\end{aligned} \tag{2.4}$$

Where, the notation for T_{n+l},S_{n+l} and I_{n+l} are as above and multiplicative seasonal components s as $S_i = S_{i+s} = S_{i+2s} = \cdots$ $(i = 1, 2, \ldots, s)$ and $\sum_{i=0}^{s} S_i = s$.

Then Z_{n+l} the forecasting value on n is

$$\begin{aligned}
\hat{Z}_n(l) &= \left(\hat{T}_n + \hat{\beta}_{1,n}l\right) \hat{S}_{n+l-s}, \quad l = 1, 2, \ldots, s,\\
\hat{Z}_n(l) &= \left(\hat{T}_n + \hat{\beta}_{1,n}l\right) \hat{S}_{n+l-2s}, \quad l = s + 1, s + 2, \ldots, 2s,\\
\hat{Z}_n(l) &= \left(\hat{T}_n + \hat{\beta}_{1,n}l\right) \hat{S}_{n+l-3s}, \quad l = 2s + 1, 2s + 2, \ldots, 3s.
\end{aligned}$$

2.3. Seasonal ARIMA model

Seasonal ARIMA (SARIMA) model of Box and Jenkins (1970) is given by Equation (2.5).

$$\phi_p(B)\Phi_P(B^s)(1-B)^d(1-B^s)^D(Y_t-\mu)=\theta_q(B)\Theta_Q(B^s)\epsilon_t. \tag{2.5}$$

Where ϵ_t is the white noise process. The ARIMA model is denoted as ARIMA$(p,d,q)\times(P,D,Q)_s$. p, d, q, P, D, Q is integer, B present Back-Shift operator, s present seasonality, d, D are the orders of differencing and seasonal differencing. The equations of the $\phi_p(B)$, $\Phi_P(B^s)$, $\theta_q(B)$, $\Theta_Q(B^s)$ as:

$$\begin{aligned}\phi_P(B)&=1-\phi_1B-\cdots-\phi_PB^P,\\ \Phi_P(B)&=1-\Phi_1B^s-\cdots-\Phi_PB^Ps,\\ \theta_q(B)&=1-\theta_1B-\cdots-\theta_qB^q,\\ \Theta_Q(B^s)&=1-\Theta_1B^s-\cdots-\Phi_QB^Qs.\end{aligned}$$

2.4. ϵ-support vector regression (ϵ-SVR)

The ϵ-SVR seeks to estimate linear functions,

$$f(x)=\langle\omega,x\rangle+b, \tag{2.6}$$

where $\omega,x\in\eta,\ b\in R$

$$(x_1,y_1),\ldots,(x_n,y_n)\in\eta\times R.$$

Vapnik (1995) solved the problem by applying the minimization method as in Equation (2.7).

$$\frac{1}{2}\|\omega\|^2+C\sum_{i=1}^{m}|y_i-f(x_i)|_\epsilon. \tag{2.7}$$

Then we can transform that into a constrained optimization problem by using slack variables, $f(x_i)-y_i>\epsilon$ and $y_i-f(x_i)>\epsilon$. We denote them by ξ and ξ^*.

Formally, the previous optimization problem as Equations (2.8), (2.9)

$$\text{minimize } \tau(\omega,\xi,\xi^*)=\frac{1}{2}\|\omega\|^2+C\sum_{i=1}^{m}(\xi+\xi^*) \tag{2.8}$$

$$\text{subject to}\begin{cases}f(x_i)-y_i\le\epsilon+\xi_i,\\ y_i-f(x_i)\le\epsilon+\xi_i^*,\\ \xi\ \xi^*\ge 0,\end{cases}\qquad\text{for all } i=1,\ldots,m. \tag{2.9}$$

Note that everything above ϵ is restricted in slack variables. The slack variables are penalized in the objective function via a regularization C. C is a constant to determine the trade off between empirical risk and the flatness of model.

In order to find the solution for (2.8), (2.9), the key idea is to construct a Lagrangian from the objective function (optimal objective function) and the corresponding constraints, by introducing a dual set of multipliers which have to satisfy the constraints, $\alpha_i,\alpha_i^*,\eta_i,\eta_i^*\ \ge\ 0$ as described in e.g.

(Fletcher, 1987). Then we define a Lagrangian as:

$$\begin{aligned} L = \frac{1}{2}\|\omega\|^2 + C\sum_{i=1}^{m}\alpha_i\,(\epsilon + \xi_i - y_i + \langle\omega, x\rangle + b) \\ - \sum_{i=1}^{m}\alpha_i^*\,(\epsilon + \xi_i^* - y_i + \langle\omega, x\rangle + b) \\ - \sum_{i=1}^{m}(\eta\xi_i + \eta^*\xi_i^*)\,. \end{aligned} \tag{2.10}$$

It follows from the saddle point condition that the partial derivatives of with respect to the primal variables have to vanish for optimality;

$$\partial_b L = \sum_{i=1}^{m}(\alpha_i - \alpha_i^*) = 0, \tag{2.11}$$

$$\partial_w L = w - \sum_{i=1}^{m}(\alpha_i^* - \alpha_i) = 0, \tag{2.12}$$

$$\partial_{\xi_i\xi_j} * L = C - \sum_{i=1}^{m}(\alpha_i + \alpha_i^*) - \sum_{i=1}^{m}(\eta + \eta^*) = 0. \tag{2.13}$$

Substituting (2.11), (2.12) and (2.13) into (2.10), then

$$\text{maximize}\begin{cases} -\dfrac{1}{2}\displaystyle\sum_{i,j=1}^{m}(\alpha_i^* - \alpha_i)(\alpha_j^* - \alpha_j)\langle x_i, x_j\rangle, \\ -\epsilon\displaystyle\sum_{i=1}^{m}(\alpha_i^* - \alpha_i) + \sum_{i=1}^{m}y_i(\alpha_i^* - \alpha_i), \end{cases} \tag{2.14}$$

$$\text{subject to } \sum_{i=1}^{m}(\alpha_i^* - \alpha_i) = 0 \quad \text{and} \quad \alpha_i, \alpha_i^* \in [0, C].$$

In deriving (2.14), we eliminate the dual variables η_i, η_i^* through condition (2.13). Equation (2.12) can be written as

$$w = \sum_{i=1}^{m}(\alpha_i^* - \alpha_i)\,x_i, \quad \text{thus } f(x) = \sum_{i=1}^{m}(\alpha_i^* - \alpha_i)\,\langle x_i, x\rangle + b.$$

In order to generalize the SV regression to nonlinear (Aizerman *et al.*, 1964), to make the SV algorithm in nonlinear could be achieved by simply preprocessing the training patterns x, by a map $\Phi : \chi \rightarrow F$ into some feature space F and then applying the standard SV regression algorithm. Therefore, if we use kernel trick which substitute $\Phi(x)$ for x, then only one part we need to replace is $\langle x_i, x_j\rangle$ and we could have

$$f(x) = \sum_{i=1}^{m}(\alpha_i^* - \alpha_i)K(x, x_i) + b. \tag{2.15}$$

Here, kernel function K takes the form of Equation (2.14)

$$\text{maximize}\begin{cases} -\frac{1}{2}\sum_{i,j=1}^{m}(\alpha_i^* - \alpha_i)(\alpha_j^* - \alpha_j)K(x_i, x_j), \\ -\epsilon\sum_{i=1}^{m}(\alpha_i^* - \alpha_i) + \sum_{i=1}^{m} y_i(\alpha_i^* - \alpha_i), \end{cases} \tag{2.16}$$

$$\text{subject to } \sum_{i=1}^{m}(\alpha_i - \alpha_i^*) = 0 \quad \text{and} \quad \alpha_i, \alpha_i^* \in [0, C].$$

We use the Gaussian Radial basis function (RBF) kernel as the evaluation of forecasting shows SVR with Gaussian RBF kernel outperforms in most cases (Claveria *et al.*, 2015). The Gaussian RBF kernel function is defined as

$$K(x, x_i) = \exp\left(\frac{1}{\sigma^2}(x - x_i)^2\right), \tag{2.17}$$

where σ^2 is the bandwidth of the Gaussian RBF kernel.

3. Data Analysis and Results

In order to compare the proposed models, the data used in this paper was applied in Wang *et al.* (2009) which is the electricity demand of northeast China from 'January, 2004' to 'April, 2009'. We use the electricity sales volume from 'January, 1965' to 'April, 2009' to compare the performance of short period data and long period data. The reason to choose Northeast China data and Korea data is that we may consider climate effect occurs by different locations.

Regarding the ϵ-SVR model, data are divided into two data sets: the training data set and the testing data set (7 months).

The selection of three parameters in ϵ-SVR, σ (controls the Gaussian function width), ϵ (controls the width of the ϵ-insensitive loss function), C. We fix $\sigma = 0.5$, $\epsilon = 0.5$ and $C = 100$ to compare the performance with benchmarking model.

The original data plots for electricity demand for Northeast China from 'January, 2004' to 'August, 2008', electricity sales in Korea from 'January, 2004' to 'August, 2008', and from 'January, 1965' to 'August, 2008' are as Figure 1. February or March are festival months that have seasonality and other months have similar trends. So we considered seasonality.

We take trend and seasonality as independent variables in linear regression model. Table 1 shows the fitting model for each of 3 data set for electricity demand (DATA1: Northeast China 'January, 2004' to 'September, 2008', DATA2: Korea 'January, 2004' to 'September, 2008', DATA3: Korea 'January, 1965' to 'September, 2008').

In this paper, we select multiplicative seasonality for Holt-Winters model. The parameters of Smoothing constant (α), seasonal constant (β) and trend constant (γ) in Holt-Winters model for the three dataset models are estimated by 'Holt-Winters' function and smoothing constant α in simple exponential smoothing is estimated by 'Ets' function in *R* program (Table 2).

Regarding ARIMA model, this paper select the model with minimum AIC (Akaike's Information Criterion) which automatically computed by auto.arima function in *R* program. According to the AIC, the best fitted seasonal ARIMA models for Data 1, 2, 3 are ARIMA$(0, 1, 1) \times (0, 0, 1)_1 2$,

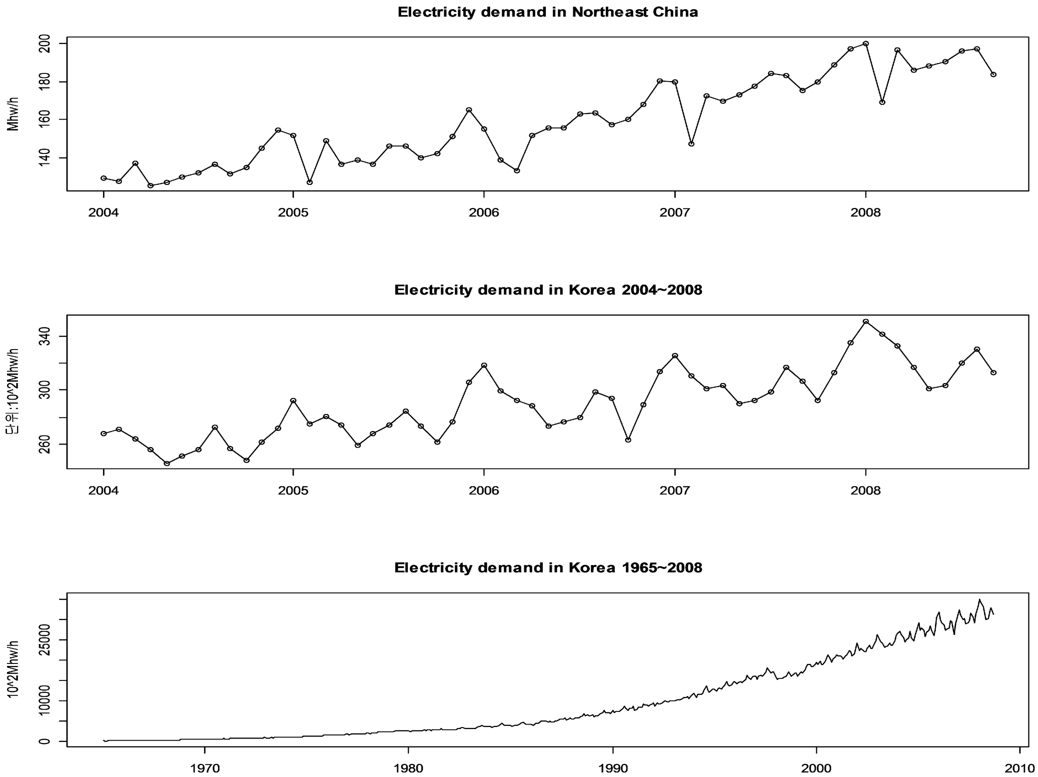

Figure 1: *Original time series data.*

Table 1: Parameter estimation of linear regression model

Parameter	Data 1	Data 2	Data 3
Trend	1.2600	131611	59355
Season 2	−22.6900	−1263038	−477559
Season 3	−8.1094	−1944665	−423042
Season 4	−13.1151	−2713336	−517987
Season 5	−11.6678	−4222956	−790242
Season 6	−11.5524	−3933078	−643828
Season 7	−6.4061	−3325752	−415332
Season 8	−6.6658	−1971268	−12643
Season 9	−15.6395	−3273378	−399264
Season 10	−12.9036	−4877693	−936114
Season 11	−5.1508	−3123890	−623574
Season 12	5.0246	−1088342	−333128

$\text{ARIMA}(0,0,1) \times (0,1,1)_{1}2$, $\text{ARIMA}(2,1,3) \times (0,1,1)_{1}2$ respectively. Non-Seasonal ARIMA models are $\text{ARIMA}(1,0,1)$, $\text{ARIMA}(1,1,1)$ and $\text{ARIMA}(2,2,2)$.

$$\text{MAPE} = \frac{1}{n} \sum_{t=1}^{n} \left| \frac{Z_t - F_t}{Z_t} \right| \times 100(\%). \tag{3.1}$$

Table 2: Parameter estimation of Holt-Winters

Parameter	Data 1	Data 2	Data 3
α	0.5214	0.506	0.2726
β	0.0001	0.1037	0.0062
γ	0.0001	0.0001	0.5784

Table 3: Parameter estimation of ARIMA model

Non-seasonal ARIMA for Data 1			
Parameter	Estimate	Standard error	p-value
Intercept	159.6526	23.8932	< .0001
ϕ_1	0.9894	0.0148	< .0001
θ_1	−0.6061	0.1028	< .0001
Non-seasonal ARIMA for Data 2			
Parameter	Estimate	Standard error	p-value
ϕ_1	−0.6067	0.1358	< .0001
θ_1	0.9633	0.0700	< .0001
drift	83000.72	209296.48	0.6917
Non-seasonal ARIMA for Data 3			
Parameter	Estimate	Standard error	p-value
ϕ_1	0.7452	0.0449	< .0001
ϕ_2	−0.3394	0.0431	< .0001
θ_1	−1.8855	0.0241	< .0001
θ_2	0.8893	0.0239	< .0001
SARIMA for Data 1			
Parameter	Estimate	Standard error	p-value
θ_1	−0.8689	0.0977	< .0001
Θ_1	0.6416	0.1605	< .0001
drift	1.2021	0.2173	< .0001
SARIMA for Data 2			
Parameter	Estimate	Standard error	p-value
θ_1	0.3581	0.1900	0.0594
Θ_1	−0.6787	0.3842	0.0773
drift	132303.7	7205.195	< .0001
SARIMA for Data 3			
Parameter	Estimate	Standard error	p-value
ϕ_1	−0.0457	0.0540	0.3973
ϕ_1	0.7825	0.0430	< .0001
θ_1	−0.3076	0.0756	< .0001
θ_2	−0.9607	0.0151	< .0001
θ_3	0.3195	0.0663	< .0001
Θ_1	−0.4238	0.0442	< .0001

Where n is the number of the data, F_t is the predicted value at time t. In this paper, we use the MAPE (mean absolute percentage error) to evaluate forecast performance among models traditionally applied to measure forecasting accuracy electricity demand. It is easy to be interpreted as the measure to compare the forecasting accuracy since the MAPE captures the differences between the forecast error and the real values. We defined the equation of MAPE as follows. Tables 4, 5, 6 shows the forecasting value for 3 dataset by each models. Table 7 are the results of MAPE for each of dataset. We know that ϵ-SVR outperforms than other models when the training set period is short-term. The performance results performance in Korea and Northeast China are almost same when we use the same period of the training set; in addition, the performance of seasonal ARIMA model is improved when the training period is long-term.

Table 4: Compare 4 model for Data 1

Time	Original	ϵ-SVR	Linear regression	Holt-Winters	ARIMA	SARIMA
2008/10	181.07	173.93	191.90	190.32	189.31	194.78
2008/11	180.56	181.01	200.91	200.53	189.00	199.23
2008/12	189.03	185.83	212.35	214.36	188.69	202.82
2009/01	182.07	192.12	208.58	208.95	188.38	203.41
2009/02	167.35	169.06	187.15	180.40	188.07	196.65
2009/03	189.30	189.93	202.99	200.80	187.77	201.17
2009/04	175.84	177.33	199.24	195.25	187.47	201.41

Table 5: Compare 4 models for Data 2

Time	Original	ϵ-SVR	Linear regression	Holt-Winters	ARIMA	SARIMA
2008/10	30403363	29919215	30562941	29615225	30688266	30257564
2008/11	30863876	30684207	32448355	31551030	30439605	32507955
2008/12	32670358	32769385	34615515	33865652	30097313	34671953
2009/01	34349683	34044781	35835468	35204958	29901663	36068707
2009/02	33306296	33519572	34704041	33739110	29701588	34887317
2009/03	32615254	32854583	34154025	32959241	29565050	34225113
2009/04	32478163	30740669	33516965	32105364	29442530	33493561

Table 6: Compare 4 models for Data 3

Time	Original	ϵ-SVR	Linear regression	Holt-Winters	ARIMA	SARIMA
2008/10	30403363	27825347	24621948	30004024	31455923	30001874
2008/11	30863876	30050942	24993843	32185875	32203795	32087537
2008/12	32670358	32471323	25343644	34628975	32793750	34459612
2009/01	34349683	34190264	25736127	36169171	33055888	35909645
2009/02	33306296	33249388	25317923	34834388	33124373	34794484
2009/03	32615254	32206817	25431795	33806665	33158066	33985596
2009/04	32478163	30422399	25396204	32755678	33231250	33090748

Table 7: MAPE of the models (%)

Dataset	ϵ-SVR	Linear regression	Holt-Winters	ARIMA	SARIMA
1	1.9437	10.9315	9.8911	4.6682	10.7503
2	1.4413	4.0073	2.0670	7.5226	4.2495
3	2.8485	21.9070	3.7120	2.3538	3.6942

4. Conclusion

We compared the performance of the models for forecasting electricity demand. Of the models, the seasonal ARIMA and ϵ-SVR approaches have been widely applied. The results show that the SVR with a Gaussian RBF kernel outperforms the rest of the models in all data sets. In addition, the Korea data shows more seasonal patterns and trends than the China data. Therefore, time series models with trend and seasonal components are outperformed in the case of Korea versus China.

In this study, only the historical electricity demand are used into consideration to forecast electricity demand. Exogenous variables such as average of temperature, illuminance, humidity, and CPI are important factors to improve forecasting accuracy. Researchers should include more extensive comparison with different type of kernel functions employing to advance the performance of the ϵ-SVR model and compare models at various types of data pattern. Other topics that employ more extensive comparison with different type of kernel functions should advance the performance of the ϵ-SVR model, comparing the methods at various types of data pattern. More work is required in the research field of smart grid and management for energy demand.

References

Aizerman, A., Braverman, E. M. and Rozoner, L. I. (1964). Theoretical foundations of the potential function method in pattern recognition learning, *Automation and Remote Control*, **25**, 821–837.

Box, G. E. P. and Jenkins, G. M. (1970). Time series analysis, *Forecasting and Control*, Holden-Day, Oakland, CA.

Claveria, O., Monte, E. and Torra, S. (2015). *Regional Forecasting with Support Vector Regressions: The Case of Spain*, University of Barcelona, Regional Quantitative Analysis Group.

Fletcher, R. (1987). *Practical Methods of Optimization*, Wiley, New York, **1**, 784–794.

Karin, K. (2011). Forecasting electricity demand in Thailand with and artificial neural network approach, *Journal of Energies*, **4**, 1246–1257.

Taylor, J. W. (2003). Short-term electricity demand forecasting using double seasonal exponential smoothing, *Journal of the Operational Research Society*, **54**, 799–805.

Taylor, J. W. (2010). Triple seasonal methods for short-term electricity demand forecasting, *European Journal of Operational Research*, **204**, 139–152.

Vapnik, V. (1995). *The Nature of Statistical Learning Theory,* Springer, New York.

Wang, J., Zhu, W., Zhang, W. and Sun, D. (2009). A trend fixed on firstly and seasonal adjustment model combined with the ϵ-SVR for short-term forecasting of electricity demand, *Energy Policy*, **37**, 4901–4909.

Winters, P. R. (1960). Forecasting sales by exponentially weighted moving averages, *Management Science*, **6**, 324–342.

Yoon, S. H., Lee, Y. S. and Park, J. S. (2009). Statistical modeling for forecasting maximum electricity demand in Korea, *Communications for Statistical Applications and Method*, **16**, 127–135.

Received October 05, 2015; Revised November 11, 2015; Accepted November 16, 2015