Songklanakarin J. Sci. Technol.
41 (6), 1294-1300, Nov. - Dec. 2019

http://www.sjst.psu.ac.th

Original Article

Forecasting petroleum consumption using hybrid SVR-DE model emphasizing on optimal parameter selection technique

Thoranin Sujjaviriyasup*

Department of Logistics Engineering, School of Engineering,
University of the Thai Chamber of Commerce, Dindeang, Bangkok, 10400 Thailand

Received: 17 January 2018; Revised: 23 March 2018; Accepted: 23 July 2018

Abstract

At present, liquid fuels remain the dominant source of transportation energy consumption all over the world. Accordingly, the future demand prediction of petroleum consumption is a very challenging task with regard to efficient supply management. In this paper, a hybrid SVR-DE model is developed and proposed to address the problem. The developed model takes ability of SVR model to formulate complex predictive function while DE algorithm is used to search the optimal parameters of SVR model. Moreover, the hybrid model is compared withboth ARIMA and SVR models as traditional single models. The experimental results indicated that the developed model outperforms traditional forecasting models based on MAE, MAPE, and sMAPE criteria. Furthermore, the forecast performance of hybrid model is significantly different from both traditional single models at 0.05 significance level. Consequently, the proposed model can be a promising tool for annual petroleum consumption.

Keywords: combined model, petroleum consumption, ARIMA, support vector regression, differential evolution

1. Introduction

From a logistics point of view, one of several key activities is transportation that is used to transfer goods and services in order to fulfill customer satisfactions. Subsequently, numerous demands on energy are generated to drive the transportation activity. Concerning transportation sector energy consumption (U.S. Energy Information Administration, 20 16), liquid fuels still remain the main source of transportation energy consumption all over the world. Although renewable energy has been developed and proposed to fulfill transporttation energy consumption, the demand of liquid fuels tends to increase annually.

Pertaining to efficient supply management, a future demand of petroleum consumption is useful information (Chiroma *et al*., 2016; de Oliveira & Oliveira, 2018; Li, Wang, Wang, & Li, 2018) and is needed to realize before making a critical decision. Thus, the accuracy of future demand is a very challenging problem. However, the pattern of future demand is fluctuation and hard to be estimated by traditional methods. Accordingly, several forecasting methods are proposed to estimate the future demand in recent years. One of forecasting methods is time series analysis, which is formulated from previous observations and is widely used by decision makers all over the world to extrapolate the future demand of energy consumption as a guideline for decision making.

For statistical methods, autoregressive integrated moving average (ARIMA) is a time series forecasting model and has dominated in linear problems. The ARIMA models are widely used in literatures of energy consumption and are extended in many fields of science in present. Nonetheless, its major disadvantage is an assumption based on linear form of time series so approximation by ARIMA models may not be sufficient for complex nonlinear real-world problems. In addition, it is often difficult to determine whether a time series is generated from a linear or nonlinear underlying process. Therefore, the ARIMA models are still used to forecast energy consumption in recent years (Al-Musaylh, Deo, Adamowski, & Li, 2018; Ediger & Akar, 2007; Hussain, Rahman, & Memon, 2016; Sen, Roy, & Pal, 2016).

*Corresponding author
Email address: thoranin_suj@utcc.ac.th

For that reason, machine learning methods are widely developed and proposed in energy forecasting. In this regard, its advantages over statistical models those make them attractive in forecasting problems. First, these models impose few prior assumptions on the model formulation due to data-driven models. Second, the models have flexible nonlinear function mapping ability. The last, these models are adaptive in nature.

One of several machine learning methods is support vector regression (SVR), which is implemented by using structural risk minimization principle. Based on the principle, the SVR models achieve an optimum network structure and always provide a unique as well as globally optimal solution. Therefore, the SVR models are successfully in energy forecasting tasks (Hu, Bao, Chiong, & Xiong, 2015; Jiang & Dong, 2017; Wang, Hou, Wang, & Shen, 2016; Yang, Che, Li, Zhao, & Zhu, 2016). However, the performance of SVR models depends heavily on an appropriate selection of hyperplane parameters. Consequently, the optimal selection of hyper-plane parameter is a major concern for developing predictive models regarding SVR models. In order to reduce a risk of improper parameter selection of SVR models, metaheuristics are employed to provide a sufficiently good solution andthe optimal parameter selection of SVR models.

Differential evolution (DE) is a method in field of evolutionary computation and has proved to be a well-known algorithm of the most successful evolutionary algorithms (Wu *et al*., 2018; Jadon, Tiwari, Sharma, & Bansal, 2017). Thus, the DE can be a useful method to provide the sufficiently good parameters of SVR model (Wang, Li, Niu, & Tan, 2012; Zhang, Deb, Lee, Yang, & Shah, 2016).

In this paper, a hybridization of SVR and DE is developed and proposed to forecast seven datasets of petroleum consumption. Relating to the proposed model, the SVR models are utilized to formulate a prediction function. Meanwhile, the DE is exploited to search the most appropriate parameters of SVR models. Moreover, the proposed model is compared to both ARIMA and SVR models based on three accuracy measures. All forecasting models are evaluated its performances based on MAE, MAPE, and sMAPE. For comparison across different datasets, the MAPE is used to identify significantly difference between forecast performances of those forecasting models based on randomized block design as analysis of variance and Tukey's Honestly Significant Difference as post hoc test.

2. Materials and Methods

2.1 Datasets of petroleum consumption

The annual datasets of petroleum consumption from 1980 to 2015 are obtained from U.S. Energy Information Administration (U.S. Energy Information Administration, 2015). Relating to the annual data sets, trend component is the dominant component of time series pattern. The summary of all datasets is presented in Figure 1.

2.2 Methodologies

2.2.1 The autoregressive integrated moving average model

The ARIMA model is generalization of autoregressive moving average in a case of non-stationary time series data and is generally referred to as an ARIMA (*p, d, q*) model. The mathematical expression of the ARIMA (*p, d, q*) model with the mean is described as Equation (1).

$$\left(1-\sum_{i=1}^{p}\varphi_i B^i\right)(1-B)^d\left(y_t-\mu\right)=\left(1-\sum_{j=1}^{q}\theta_j B^j\right)\varepsilon_t \quad (1)$$

where y_t and ε_t are the actual value and random error at time period t, respectively. B is the backward shift operator; p and q are referred to orders of autoregressive integrated moving average model, which are integers as well. Recently, several software vendors have implemented automated time series forecasting procedure for univariate autoregressive integrated moving average by using several criteria i.e., Akaike's information criterion (AIC), Bayesian information criterion (BIC), and Akaike's information criterion with a correction for small sample sizes (AICc). Since the various ARIMA models can mimic the behavior of type of series. The

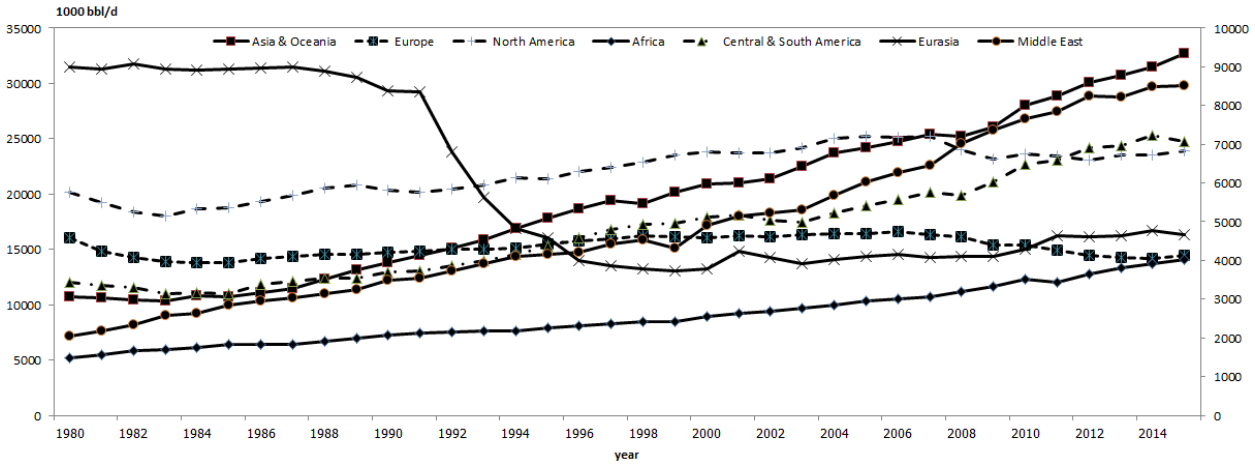

Figure 1. Datasets of petroleum consumption.

automated function (Hyndman *et al.*, 2015; Melard & Pasteels, 2000; Müller & Bogenberger, 2015), namely auto.arima, is employed in this research. Referring the automated function, the function will conduct a search over possible model within the order constraints provided and will return the best ARIMA model based on the lowest AICc criterion. According to the best ARIMA model depends on the change of the previous observation update. Therefore, the ARIMA model is used to stand for ARIMA (*p*, *d*, *q*) in this research. The criterion used to select the proper model is AICc as shown in Equations (2) and (3).

$$\text{AICc} = \text{AIC} + \frac{2k(k+1)}{n-k-1} \tag{2}$$

$$\text{AIC} = 2k - 2\ln(L) \tag{3}$$

where *n* and *k* are sample size and parameters, respectively. *L* is the maximum value of the likelihood function for the model.

2.2.2 The support vector regression model

The support vector machine model (Meyer, Dimitriadou, Hornik, Weingessel, & Leisch, 2014) is one of computational intelligence models, which is supervised learning model and is usually exploited in classification. Regarding regression analysis, The SVR model is an extension model of support vector machine concerning regression problems and generates a regression function by applying a set of high dimensional linear function and nonlinear function by using kernel functions. In addition, the SVR models are able to be a useful tool for solving complex problems and unknown functional relationship as well. The model formulation of $\varepsilon-$SVR for regression problems is shown as Equation (4)

$$f(x_i) = \sum_{i=1}^{T}(\alpha_i - \alpha_i^*)K(x, x_i) + b \tag{4}$$

where α_i and α_i^* are the so-called Lagrange multipliers, *b* is a scalar threshold, $K(x, x_i)$ is kernel function. The kernel functions are the most used for SVR models, which are defined as follows:

Linear: $K(x, x_i) = x^T x_i$,

Polynomial: $K(x, x_i) = (\gamma x^T x_i + r)^p$

Radial basis: $K(x, x_i) = \exp(-\gamma \|x - x_i\|^2)$,

Sigmoid: $K(x, x_i) = \tanh(\gamma x^T x_i + r)$

where γ, *r*, and *p* are kernel parameters.

In this research, the SVR models are used to forecast time series data by rearranging previous observations into *m* columns of previous observations, namely SVR (*m*). The first *m* – 1 columns and the last column are used as input data and target data, respectively. The data preparation of previous dataset before employing the SVR models is presented as Figure 2.

$$\begin{bmatrix} y_1 & y_2 & y_3 & \cdots & y_m \\ y_2 & y_3 & y_4 & \cdots & y_{m+1} \\ y_3 & y_4 & y_5 & \cdots & y_{m+2} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ y_{t-m+1} & \cdots & y_{t-2} & y_{t-1} & y_t \end{bmatrix}$$

Figure 2. Data preparation of previous dataset.

In order to explore the most fitted SVR model, the best SVR model is selected based on the lowest mean square error. Moreover, the appropriate parameter estimation is searched by using grid search.

2.2.3 The differential evolution optimization

The differential evolution (DE) (Ardia, Mullen, Peterson, & Ulrich, 2013) is a population based stochastic search heuristic introduced by Storn and Price (1997) as a global optimization algorithm on continuous numerical minimization problems.The algorithm of differential evolution is presented in Figure 3.

2.2.4 Hybridization of SVR and DE

The main objective of the hybrid model is to developed complex models emphasizing on parameter selection and optimization technique. The proposed model uses capability of SVR models to model the predictive function whilst the DE algorithm is used to find the best parameters of SVR model in order to guarantee the best SVR model based on a given search space. The algorithm of the developed model is presented as follows:

For *m* equal to 2 to a termination criterion do

- The kernel function is selected to formulate the model formulation of $\varepsilon-$SVR to predict the future data as Equation (4).
- The DE provides initial parameters with regard to dimensional features of the kernel function as presented in Equation (5).

$$\theta_{i,G} = [\theta_{1,i,G}, \theta_{2,i,G}, \theta_{3,i,G}, ..., \theta_{d,i,G}] \quad i = 1,2,3,...,N \tag{5}$$

Data: Let $f : \Re^d \to \Re$ be the cost function which must be minimized. Let $\mathbf{x} \in \Re^d$ designate a candidate solution in population with *d* real parameter. Where *N* and *G* are the size of the population, and the number of generations, respectively. The DE algorithm can be described as follows:

1. Initialize all agents **x** with random positions in the search space.
2. Until a termination criterion is met repeat the following:

   For each agent **x** in the population does:

   - An initial mutant parameter vector $v_{i,G}$ is created by selecting three members of the population, $x_{r_0,G}$, $x_{r_1,G}$ and $x_{r_2,G}$ randomly. Where *G* is the number of generations, r_0, r_1, and r_2 are randomly chosen numbers within the population size; and $i = 1,2,3,..., N$.
   - The crossover operator generates a trial vector $u_{i,G}$ Where *G* is the number of generations; $i = 1,2,3,..., N$.
   - The differential evolution uses a greedy selection operator.

3. Choose the agent from the population that has the highest fitness or lowest cost and return it as the best found candidate solution.

Figure 3. DEalgorithm.

where θ is a vector of kernel function parameters, d is dimensional parameters of the kernel function, G is the number of generation, and N is the size of population.

- Until a termination criterion is met repeat as follows:

For each agent θ in the population does:

For t equal to 70% of the previous observations to the observation before the last observation do

- Rearrange the previous observations into m columns of the previous observations.

$$\begin{bmatrix} y_1 & y_2 & y_3 & \cdots & y_m \\ y_2 & y_3 & y_4 & \cdots & y_{m+1} \\ y_3 & y_4 & y_5 & \cdots & y_{m+2} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ y_{t-m+1} & \cdots & y_{t-2} & y_{t-1} & y_t \end{bmatrix}$$

- For the first $m-1$ columns of the matrix of the previous observations are used as input data whilst the last column of the matrix of the previous observations is adopted as target data.
- Use the parameters of the kernel function as Equation (6).

$$\hat{y}_t = f\left(y_{t-1}, y_{t-2}, y_{t-3}, \ldots, y_{t-m+1} : \theta\right) \tag{6}$$

where f is the prediction function determined by the support vector regression; y_t is the actual value at time period t; $\hat{y}_t$ is the predicted value at time period t; θ is a set of the kernel function parameters; m is integer that represents the number of columns.

- The SVR model uses the parameters to formulate the fitted prediction function.
- The fitted SVR model is exploited to forecast the future data.
- Calculate MAPE

End

- An initial mutant parameter vector $v_{i,G}$ is created by selecting three members of the population, $\theta_{r_0,G}$, $\theta_{r_1,G}$ and $\theta_{r_2,G}$ at random as Equation (7).

$$v_{i,G} = \theta_{i,G} + \left(\theta_{best,G} - \theta_{i,G}\right) + \theta_{r_0,G} + 0.8 \times \left(\theta_{r_1,G} - \theta_{r_2,G}\right) \tag{7}$$

where $\theta_{i,G}$ and $\theta_{best,G}$ are the i-th vector of the population at the current generation and the best individual vector with the best fitness, respectively. G is the number of generations; $r_0, r_1,$ and r_2 are randomly chosen numbers within the population size; and $i = 1,2,3,\ldots,N$

- The crossover operator generates a trial vector $u_{i,G}$ as Equation (8)

$$u_{i,G} = \begin{cases} v_{j,i,G} & \text{if } rand_{j,i} \le 0.5 \text{ or } j = I_{rand} \\ \theta_{j,i,G} & \text{otherwise} \end{cases} \tag{8}$$

where $i = 1,2,3,\ldots,N$; $j = 1,2,3,\ldots,d$, $rand_{j,i} \sim U[0,1]$, I_{rand} is a random integer from $[1,2,\ldots,d]$. G is the number of generations.

- The differential evolution uses a greedy selection operator as Equation (9).

$$\theta_{i,G+1} = \begin{cases} u_{i,G} & \text{if } f\left(u_{i,G}\right) \le f\left(\theta_{i,G}\right) \\ \theta_{i,G} & \text{otherwise} \end{cases} \tag{9}$$

where $f\left(u_{i,G}\right)$ is the MAPE of the trial vector and $f\left(\theta_{i,G}\right)$ equal to MAPE of the target vector. G is the number of generations; and $i = 1,2,3,\ldots,N$

End

- Choose the agent from the population that has the lowest MAPE and return it as the best found parameters of kernel function.

End

2.3 Cross-validation

All forecasting models are evaluated their forecast accuracy by using seven time series datasets of petroleum consumption in each continent, which are obtained from U.S. Energy Information Administration. Each dataset of the petroleum consumptions is separated into two subsets as training dataset and test dataset. For training dataset, 70% of each time series data of petroleum consumption is used to formulate the fitted model and to forecast one step ahead. The rest of each time series data of petroleum consumption is used to evaluate those forecast models as the hold–out set, which is 11–fold cross–validation. After the actual data is realized, then it is included into training dataset to formulate and predict one step ahead until the last data of hold–out set.

For existing measures of accuracy (Hyndman & Koehler, 2006), the most commonly used measures are MAE and MAPE. The MAE is recommended to evaluate forecast accuracy on same scale of data sets due to scale-dependent measure. Meanwhile, the MAPE is also recommended to evaluate forecast accuracy and is the primary measure in the M-competition. An advantage of MAPE measure is percentage error that has advantage of being scale-independent. Thus, it is frequently used to compare forecast performance across different data sets rather than MAE as scale-dependent measure. Even though, there are arguments in against of using MAPE in some situations (i.e., meaningful zero, heavier penalty on positive errors than on negative errors), it may still be preferred for reasons of simplicity to explain. With regard to reduce the disadvantages of MAPE, sMAPE is developed and proposed to address the problems. In order to evaluate the forecast performances, three accuracy measures used in this research are MAE, MAPE, and sMAPE in cross-validation. All mathematical expressions of the accuracy measures are presented in Equation (10) to (12).

$$\mathrm{MAE} = \frac{\sum_{i=1}^{n} |y_i - \hat{y}_i|}{n} \tag{10}$$

$$\mathrm{MAPE} = \frac{\sum_{i=1}^{n} |y_i - \hat{y}_i| / y_i}{n} \times 100 \tag{11}$$

$$\mathrm{sMAPE} = \frac{\sum_{i=1}^{n} 2 \times |y_i - \hat{y}_i| / (y_i + \hat{y}_i)}{n} \times 100 \tag{12}$$

3. Results and Discussion

For a superior ability of the most proper model with regard to petroleum consumption, all forecasting models are evaluated by using several criteria under many situations. First, all forecasting models are compared based on the three measures of forecast accuracy. Second, descriptive statistical analysis is described by using box plots based on MAPE to display measures of central tendency. The last, both analysis of variance test and post hoc test are analyzed to identify significantly difference between the forecast models. The summary of all forecasting models based on three measures of forecast accuracy is demonstrated in Table 1.

In Table 1, the proposed model provides higher accuracy than both traditional single models. This finding revealed that the developed model may be a meaningful model to deal with these problems. Moreover, the experimental results indicated that the utilization of optimization technique can overcome the traditional SVR model in all cases. Meanwhile the SVR model outperforms ARIMA model approximately 82% of all cases. Consequently, this evidence supports that the technique can enhance forecast accuracy and reduces the risk of using improper parameters of SVR model. The most appropriate models of SVR regarding petroleum consumption are described in Table 2.

Furthermore, the box plot is used to display descriptive statistics of each forecast model based on MAPE as presented in Figure 4.

Table 2. Most appropriate models of SVR regarding petroleum consumption.

Data	SVR model	Kernel function	parameter	
			Cost	Gamma
Africa	SVR(4)	Linear	9.600441	NA
Asia Oceania	SVR(2)	Radial basis	144.051613	0.01009
Central South America	SVR(3)	Linear	0.365700	NA
Eurasia	SVR(2)	Radial basis	0.355321	158.013591
Europe	SVR(4)	Linear	0.511490	NA
Middle East	SVR(5)	Radial basis	360.582893	0.037667
North America	SVR(2)	Radial basis	165.570822	3.549593

Table 1. Summary of all forecasting models based on three measures of forecast accuracy.

Data	ARIMA			SVR			SVR-DE		
	MAE	MAPE	sMAPE	MAE	MAPE	sMAPE	MAE	MAPE	sMAPE
Africa	77.849	2.26	2.27	59.431	1.77	1.76	59.204	1.76	1.75
Asia Oceania	479.484	1.74	1.75	525.575	1.85	1.88	353.720	1.28	1.30
Central South America	179.418	2.80	2.80	154.173	2.43	2.43	139.733	2.20	2.20
Eurasia	116.613	2.60	2.60	133.186	2.94	2.99	57.499	1.32	1.32
Europe	305.950	2.02	2.02	234.922	1.55	1.53	226.326	1.49	1.48
Middle East	177.996	2.46	2.48	126.105	1.64	1.63	40.764	0.58	0.58
North America	473.003	1.98	1.97	406.132	1.70	1.69	327.307	1.39	1.37

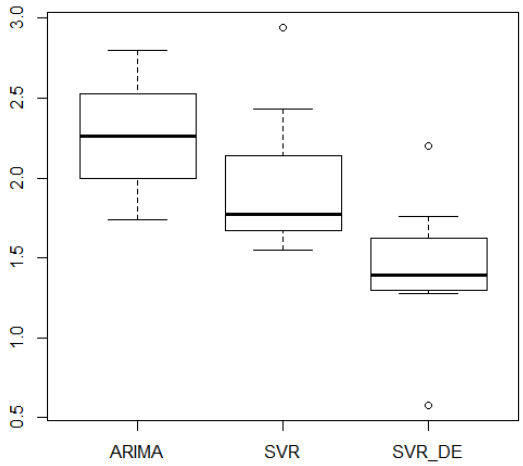

Figure 4. Visual data display of box plot based on MAPE.

Referencing Figure 4, the proposed model demonstrated that it provides both the lowest mean and median of MAPE compared to other candidate forecasting models. Moreover, both mean and median of SVR model are also lower than ARIMA model. In order to identify significantly difference of forecast performances, the randomized block design is ex-ploited to investigate the finding. However, the normality test of MAPE has to be performed by using Shapiro-Wilk test before applying the analysis of variance test. Given results of normality test, p-values of Shapiro-Wilk test are more than 0.05, which indicated that MAPE of each forecasting model comes from a normally distributed population at 0.05 signi-ficance levels. Subsequently, the summary of analysis of va-riance test based on MAPE is presented in Table 3.

As given results in Table 3, this evidence indicated that at least one prediction performance of forecasting model is significantly different from one other at the 0.05 significance level. However, the datasets are not significantly different from one other at 0.05 significance level. For pairwise comparisons of means, the Tukey's Honestly Significant Difference is used to identify significantly difference. The summary of multiple comparisons is demonstrated in Figure5.

As results of multiple comparisons, the forecasting performance of proposed model is significantly different from both ARIMA and SVR models. On the other hand, the SVR model is not significantly different from ARIMA model.

4. Conclusions

According to all evidences, they revealed that the proposed model has superior ability and provides higher accuracy than other compared models at 0.05 significance levels. Consequently, it can support to conclude that the optimization technique concerning parameter selection of SVR model is able to improve forecast accuracy compared with candidate models. In other words, the technique of optimal parameter selection can reduce the risk of using improper parameters of SVR models. Thus, the developed model can overcome the limitation of each other regarding datasets of petroleum consumption and can be a promising tool to predict annual petroleum consumption.

Table 3. Summary of analysis of variance test based on MAPE.

	Df	Sum square	Mean square	P-value
Forecasting model (treatment)	2	2.520	1.2601	0.00334
Dataset (block)	6	2.293	0.3822	0.05574
Residuals	12	1.588	0.1323	

```
                  diff       lwr        upr     p adj
SVR-ARIMA    -0.2828571 -0.8016471  0.23593284 0.3458282
SVR_DE-ARIMA -0.8342857 -1.3530757 -0.31549573 0.0027860
SVR_DE-SVR   -0.5514286 -1.0702186 -0.03263859 0.0371610
```

Figure 5. Summary of multiple comparisons based on MAPE.

References

Al-Musaylh, M. S., Deo, R. C., Adamowski, J. F., & Li, Y. (2018). Short-term electricity demand forecasting with MARS, SVR and ARIMA models using aggregated demand data in Queensland, Australia. *Advanced Engineering Informatics, 35*, 1-16.

Ardia, D., Mullen, K. M., Peterson, B. G., & Ulrich, J. (2013). DE-optim: Differential Evolution in R.

Chiroma, H., Khan, A., Abubakar, A. I., Saadi, Y., Hamza, M. F., Shuib, L., . . . Herawan, T. (2016). A new approach for forecasting OPEC petroleum consumption based on neural network train by using flower pollination algorithm. *Applied Soft Computing, 48*, 50-58.

de Oliveira, E. M., & Oliveira, F. L. C. (2018). Forecasting mid-long term electric energy consumption through bagging ARIMA and exponential smoothing methods. *Energy, 144*, 776-788.

Ediger, V. Ş., &Akar, S. (2007). ARIMA forecasting of primary energy demand by fuel in Turkey. *Energy Policy, 35*(3), 1701-1708.

Hu, Z., Bao, Y., Chiong, R., & Xiong, T. (2015). Mid-term interval load forecasting using multi-output support vector regression with a memetic algorithm for feature selection. *Energy, 84*, 419-431.

Hussain, A., Rahman, M., & Memon, J. A. (2016). Forecasting electricity consumption in Pakistan: the way forward. *Energy Policy, 90*, 73-80.

Hyndman, R. J., Athanasopoulos, G., Razbash, S., Schmidt, D., Zhou, Z., Khan, Y., . . . Wang, E. (2015). Forecast: Forecasting functions for time series and linear models. *R package version, 6*(6), 7.

Hyndman, R. J., & Koehler, A. B. (2006). Another look at measures of forecast accuracy. *International Journal of Forecasting, 22*(4), 679-688.

Jadon, S. S., Tiwari, R., Sharma, H., & Bansal, J. C. (2017). Hybrid artificial bee colony algorithm with differential evolution. *Applied Soft Computing, 58*, 11-24.

Jiang, H., & Dong, Y. (2017). Global horizontal radiation forecast using forward regression on a quadratic kernel support vector machine: Case study of the Tibet Autonomous Region in China. *Energy, 133*. doi:10.1016/j.energy.2017.05.124

Li, J., Wang, R., Wang, J., & Li, Y. (2018). Analysis and forecasting of the oil consumption in China based on combination models optimized by artificial intelligence algorithms. *Energy, 144*, 243-264.

Melard, G., & Pasteels, J. M. (2000). Automatic ARIMA modeling including interventions, using time series expert software. *International Journal of Forecasting, 16*(4), 497-508.

Meyer, D., Dimitriadou, E., Hornik, K., Weingessel, A., & Leisch, F. (2014). e1071: Misc Functions of the Department of Statistics (e1071), TU Wien. R package version 1. 6-3. Retrieved from http://CRAN.R-project. org/package=e1071

Müller, J., & Bogenberger, K. (2015). Time series analysis of booking data of a free-floating Carsharing system in Berlin. *Transportation Research Procedia, 10*, 345-354.

Sen, P., Roy, M., & Pal, P. (2016). Application of ARIMA for forecasting energy consumption and GHG emission: A case study of an Indian pig iron manufacturing organization. *Energy, 116*, 1031-1038.

U.S. Energy Information Administration. (2017, December 22). International Energy Outlook 2016. Retrieved from https://www.eia.gov/outlooks/ieo/pdf/0484(2016).pdf

U.S. Energy Information Administration. (2017, December 22). Total petroleum and other liquids production 2015. Retrieved from https://www.eia.gov/beta/international/data/browser/#/?pa=0000001&c=4100000002000060000000000000g0002&tl_id=5-A&vs=INTL.5-2ASOC-TBPD.A&vo=0&v=H&end=2015.

Wang, J., Hou, R., Wang, C., & Shen, L. (2016). Improved v-Support vector regression model based on variable selection and brain storm optimization for stock price forecasting. *Applied Soft Computing, 49*, 164-178.

Wang, J., Li, L., Niu, D., & Tan, Z. (2012). An annual load forecasting model based on support vector regression with differential evolution algorithm. *Applied Energy, 94*, 65-70.

Wu, G., Shen, X., Li, H., Chen, H., Lin, A., & Suganthan, P. N. (2018). Ensemble of differential evolution variants. *Information Sciences, 423*, 172-186.

Yang, Y., Che, J., Li, Y., Zhao, Y., & Zhu, S. (2016). An incremental electric load forecasting model based on support vector regression. *Energy, 113*, 796-808.

Zhang, F., Deb, C., Lee, S. E., Yang, J., & Shah, K. W. (2016). Time series forecasting for building energy consumption using weighted Support Vector Regression with differential evolution optimization technique. *Energy and Buildings, 126*, 94-103.