A Hybrid One-step-ahead Time Series Model Based on GA-SVR and EMD for Forecasting Electricity Loads

Mu-Chiun Shiu, Liang-Ying Wei*, Jing-Wei Liu, Deng-Yang Huang, Chien-Chih Tu and Kuo-Hsiung Liao

1Department of Information Management, National Yunlin University of Science and Technology, Yunlin, Taiwan 640, R.O.C.
2Department of Information Management, Yuanpei University of Medical Technology, Hsinchu, Taiwan 30015, R.O.C.
3Department of Sport Information and Communication, National Taiwan University of Sport, Taichung, Taiwan 404, R.O.C.

Abstract

Economic growth increases the demand for electricity, and forecasting electricity loads is critical for providing cheaper electricity. Conventional time series methods have been applied to forecast electricity loads. However, traditional statistical methods, such as regression models, are unable to address nonlinear relationships, such as those of electricity loads. Moreover, most time-series models which use electricity load data with many factors, such as climate conditions and region environments, involutedly would reduce the forecasting performance. To overcome these problems and improve the forecasting ability of time series models, this paper proposes a hybrid one-step-ahead time series model that is based on support vector regression (SVR), empirical mode decomposition (EMD), and a genetic algorithm (GA) to predict electricity loads. The experimental results were generated from 2 electricity load datasets from various countries, and the proposed model was compared with several models. Our findings indicate that the proposed model outperforms the other approaches in terms of mean absolute percentage error (MAPE).

Key Words: Electricity Load, Support Vector Regression, Empirical Mode Decomposition, Time Series, One-step-ahead Method, Genetic Algorithm

1. Introduction

Economic growth provides opportunities for people to lead high-quality lives, but such growth continually increase the demand for electric energy. To meet customers’ requirements, more power plants are being built, and consequently, forecasting electricity demand has become an important topic in planning and operating energy systems. With regard to operating, planning, and reducing the operational costs of generating power and expanding the security, delivery, and transmission of power, forecasting electricity load has historically been a significant issue.

*Corresponding author. E-mail: lywei@mail.ypu.edu.tw
forecasting error by 1% increases operational costs by ≤ 10 million every year. Energy is balanced between generation and demand, when there is a lack of energy storage. Thus, under- or overestimation of electricity load causes overpayment.

In cases of underestimation, power delivery companies should spend more for the energy in demand, potentially creating a shortage of generated power, because it is always more expensive- and sometimes impossible- to purchase electricity on short notice. Conversely, power delivery companies waste resources due to overestimation. Consequently, highly accurate forecasting of load leads to substantial savings in operational costs and greater reliability of power supply. As a result, power delivery companies strive to forecast their demand for electricity to strike a balance between under- and overestimation. Unfortunately, because large load variations are possible in a period of an hour or a day, forecasting electricity loads becomes difficult.

Due to competition between energy suppliers, scheduling energy transfers and dispatching loads are major issues. Many methods have been developed to deal with problems that are related to forecasting electricity load and are divided into four broad categories: time series models, regression-based models, artificial intelligence techniques, and fuzzy logic methods.

Artificial neural networks can learn load series and perform well in forecasting loads. Support vector machine (SVM) models are based on the principle of structural risk minimization (SRM) rather than empirical risk minimization, which is performed by most traditional neural network models and has been extended to solve nonlinear regression estimation problems.

Pai and Hong [16] used a recurrent SVM with genetic algorithms (RSVMG) model to forecast electricity loads. Ertugrul [17] proposed a recurrent extreme learning machine (RELM) tool to forecast electricity loads more accurately. Ghasemi et al. [18] proposed a hybrid algorithm in smart grids with demand-side management to price electricity and forecast loads. Hassan et al. [19] used an interval type-2 fuzzy logic system with an extreme learning machine to forecast electricity load demand. Takeda et al. [20] used the ensemble Kalman filter to forecast electricity loads. Clements et al. [21] implemented a multiple-equation time series approach to forecast day-ahead electricity loads. Niu et al. [22] used a support vector machine and ant colony optimization to forecast power loads. Fuzzy logic models have been proposed to forecast electricity loads and are useful when decision rules are represented by linguistic terms, simplifying this task for a decision-maker. To this end, Pai [14] developed a hybrid ellipsoidal fuzzy system for time series forecasting (HEFST) and applied it to predict regional electricity loads in Taiwan.

Conventional time series models, such as statistical models, were initially adopted for load forecasting problems, including the autoregressive moving average (ARMA) model [8]. The ARMA model generates forecasts under linear stationary conditions. However, for nonstationary condition, the ARIMA (autoregressive integration moving average) [23] model was proposed to describe such homogeneous nonstationary behavior. ARIMA models consider related factors, such as seasonal temperature and day type, in forecasting. Further, raw electricity loads datasets usually constitute an input variable for conventional time-series models to forecast electricity loads. Raw electricity load datasets include such factors as time, climate conditions, and regional environment, and conventional time-series models that use raw data as the input variable have lower forecasting accuracy.

Regression models construct cause-effect relationships between electricity load and independent variables. However, these independent variables are nonlinear, and existing regression models are unable to consider nonlinear relationships well, because electricity loads are nonlinear. Therefore, many nonlinear artificial intelligence forecasting methods, such as the adaptive network-based fuzzy inference system (ANFIS) [1], artificial neural network (ANN) [25], SVM [16], and support vector regression (SVR) [26], have been developed to address the drawbacks above.

The SVM method, which was introduced by [15], has recently been used in various applications, including the prediction of stock market prices [27]. The SVM technique, in general, is regarded as a state-of-the-art classifier, and many studies have indicated that SVM prediction tools are superior to neural network-based approaches [28]. Developed initially to solve classification problems, SVM techniques can be applied to regression problem...
Unlike pattern recognition problems in which the desired outputs are discrete values (e.g., Boolean), SVR deals with ‘real valued’ functions. SVR is derived from the structural risk minimization principle to estimate a function by minimizing the upper bound of the generalization error [30]. The SVR model has solved prediction problems in many domains [26,31]. Although SVR has many applications, such as in economic and financial predictions, its main problem is the determination of its parameters, which requires practitioner experience [32]. Further, setting unsuitable parameters for SVR reduces its forecasting performance.

Recently, GA was applied to solve parameter optimization problems [33]. GA [34] is a search algorithm that was inspired by evolution and is usually used to determine the optimal parameters of machine learning methods. Therefore, our proposed models use GA to optimize the parameters of SVR and improve its forecasting performance.

One-step-ahead prediction using neural networks has been used for many engineering problems, reporting satisfactory results [35]. Pino et al. [10] used the one-step-ahead forecast method [25] in an experimental procedure that was based on ANN, generating encouraging findings [10]. Thus, our paper combines the one-step-ahead forecast method with SVR to improve forecasting performance. Moreover, empirical mode decomposition (EMD), proposed by Huang et al. [36], is suitable for nonlinear signal analysis, which adaptively represents the local characteristic of a given signal [36]. Based on EMD, any complicated signal can be decomposed into a finite and often small number of intrinsic mode functions (IMFs) [36], which have simpler frequency components and stronger correlations, rendering them easier to forecast more accurately. Thus, EMD has been used widely in many fields, such as in analyzing earthquake signals and structures, monitoring the state of bridges and constructions [37], examining sea wave data [38], and diagnosing the fault of machines [39].

There are several drawbacks of the models above: (1) statistical models can not be applied to datasets that do not follow statistical assumptions [33]; (2) traditional time-series models and earlier electricity load forecasting models that use complicated raw data do not forecast well [1], and (3) it is difficult to determine the parameters with SVR [40].

To overcome these drawbacks, this paper takes into account that EMD can decompose complicated raw data (electricity loads) into simpler frequency components and high-correlation variables, incorporating it into the one-step-ahead model for constructing the primary model. Then, the decomposed datasets are used as input variables to forecast electricity loads by SVR. Subsequently, the GA is used to refine SVR to improve forecasting performance. Thus, this study expects that the proposed model can generate huge profits for power delivery companies by providing more accurate forecasts of electricity loads.

The remainder of this paper is organized as follows. The empirical mode decomposition method is presented in section 2. Section 3 then introduces the proposed hybrid one-step-ahead time-series model. The performance of the proposed model is evaluated in section 4, and the conclusions are made in section 5.

2. Empirical Mode Decomposition

Huang et al. [36] proposed the EMD technique—an adaptive time series decomposition tool that uses the Hilbert-Huang transform (HHT) for nonlinear and non-stationary time series data. The basic principle of EMD is to decompose a time series into a sum of oscillatory functions—namely, intrinsic mode functions (IMFs). In EMD, IMFs must satisfy 2 conditions: (1) the number of extrema (sum of maxima and minima) and the number of 0 crossings must differ only by 1, and (2) the local average is 0. The latter condition implies that the mean of the upper and lower envelopes is equal to 0 [36]. The first condition is similar to the traditional narrow band requirements for a stationary Gaussian process [36]. The second condition modifies the classical global requirement to a local one, so that the instantaneous frequency will not have unwanted fluctuations that are induced by asymmetric wave forms [36]. The detailed algorithm for EMD is as follows [36]:

Step 1: Identify local extrema in the experimental data \( \{x(t)\} \). All local maxima are connected by a cubic spline line, \( U(t) \), which forms the upper envelope of the data. Repeat this procedure for the
local minima to produce the lower envelope, \( L(t) \). Both envelopes will cover all data between them. The mean of the upper and lower envelopes, \( m_1(t) \), is given by:

\[
m_1(t) = (U(t) + L(t))/2
\]  

(1)

Subtracting the running mean \( m_1(t) \) from the original time series \( x(t) \), we obtain the first component \( h_1(t) \):

\[
h_1(t) = x(t) - m_1(t)
\]  

(2)

The resulting component \( h_1(t) \) is an IMF if it is symmetrical and if all of its maxima positive and all of its minima are negative. An additional condition of intermittence can be imposed here to sift out waveforms with a certain range of intermittence values for physical consideration. If \( h_1(t) \) is not an IMF, the sifting process must be repeated as many times as it is required to reduce the extracted signal to an IMF. In subsequent sifts, \( h_1(t) \) is treated as the data that is to be repeated in the steps above:

\[
h_{1k}(t) = h_1(t) - m_{1k}(t)
\]  

(3)

Again, if the function \( h_{1k}(t) \) does not satisfy the criteria for IMF, the sifting process continues up to \( k \) times until an acceptable tolerance is reached:

\[
h_{1k}(t) = h_{1(k-1)}(t) - m_{1k}(t)
\]  

(4)

**Step 2:** If the resulting time series is an IMF, it is designated as

\[
c_1 = h_{1k}(t).
\]

The first IMF is then subtracted from the original data, and the difference \( r_1(t) \) is given by:

\[
r_1(t) = x(t) - c_1(t)
\]  

(5)

The residue \( r_1(t) \) is taken as if it were the original data, and we sift sifting again as in Step 1.

Following the procedures above, we continue the process to identify more intrinsic modes \( c_i \) until the final one. The final residue will be a constant or a monotonic function that represents the general trend of the time series. Finally, we obtain

\[
x(t) = \sum_{i=1}^{n} c_i(t) + r_n
\]

\[
r_n(t) = r_1(t)
\]  

(6)

where \( r_n \) is a residue.

Thus, residue \( r_n(t) \) is the mean trend of \( x(t) \). The IMFs \( c_1(t), c_2(t), \ldots, c_n(t) \) includes various frequency bands, ranging from high to low. The frequency components that are contained in each frequency band differ, and they change with variations in signal \( x(t) \), whereas \( r_n(t) \) represents the central tendency of signal \( x(t) \).

### 3. Proposed Model

To enhance forecasting performance, this paper applies the one-step-ahead forecast method to the proposed model. The one-step-ahead method can use current accuracy values to obtain forecasts for the next period. Further, to overcome the drawbacks in literature review, this paper considers that EMD can decompose raw data (electricity load data) into simpler frequency components and highly correlating variables, adopting the one-step-ahead time series model for building the primary model. Then, the results of primary model are refined by SVR, which can overcome the limitations of statistical methods (the data need to obey some mathematical distribution) and handle noisy data involutedly.

Basic SVR concepts can be found in [15]. Given a training set \((x_i, y_i), i = 1, 2, \ldots, m\), where the input variable \(x_i \in \mathbb{R}^n \) is the \(n\)-dimensional vector and the response variable \(y_i \in \mathbb{R}^c \) is the continuous value. SVR builds the linear regression function in the following form:

\[
f(x, w) = w^T x + b
\]  

(7)

Based on Vapnik’s linear \(\varepsilon\)-insensitivity loss (error) function (Equation (8)), the linear regression \(f(x, w)\) is estimated by simultaneously minimizing \(||w||^2\) and the sum of the linear \(\varepsilon\)-insensitivity losses (Equation (10)). The constant \(C\), which influences a trade-off between an approximation error and the weight vector norm \(||w||\), is a design parameter that is chosen by the user.

\[
|y - f(x, w)| = \begin{cases} 0, & \text{if } |y - f(x, w)| \leq \varepsilon \\ |y - f(x, w)| - \varepsilon, & \text{otherwise} \end{cases}
\]  

(8)

\[
R = \frac{1}{2} ||w||^2 + C \sum_{i=1}^{m} \left[ |y_i - f(x_i, w)|_{\varepsilon} \right]
\]  

(9)

Minimizing the risk \(R\) is equivalent to minimizing
the following risk:
\[
R_{\epsilon,\zeta_i} = \frac{1}{2} \| w \|^2 + C \sum_{i=1}^{m} (\zeta_i + \zeta_i^*)
\]  
(10)
under constraints:
\[
w^T x_i + b - y_i \leq \epsilon + \zeta_i
\]  
(11)
\[
y_i - (w^T x_i + b) \leq \epsilon + \zeta_i^*
\]  
(12)
\[
\zeta_i, \zeta_i^* \geq 0, \ i = 1, 2, \ldots, m
\]  
(13)
where \(\zeta_i\) and \(\zeta_i^*\) are slack variables—one for exceeding the target value by more than \(\epsilon\) and the other for being more than \(\epsilon\) below the target. As with procedures that are applied to SVM classifiers [15], this constrained optimization problem is solved by applying the Lagrangian theory and the Karush Kuhn-Tucker condition to obtain the optimal desired weight vector of the regression function. The most popular kernel function is the radial basis function (RBF), as shown in Equation (14).
\[
K(x_i, x_j) = \exp(-\gamma \| x_i - x_j \|^2)
\]  
(14)
Further, the GA optimizes the parameters of the SVR model. The GA [34] searches for global optima using techniques that are inspired by natural evolution, such as inheritance, mutation, selection, and crossover. Further, the GA has been applied successfully in economic and financial predictions [41,42]. This algorithm encodes a potential solution of a specific problem into a simple chromosome-like data structure and applies recombination operators for these structures to preserve critical information [43]. The steps of the GA in the proposed model are based on Goldberg [44].

To summarize the proposed model, the flowchart of the proposed model is shown in Figure 1. To convey our model easily, the proposed algorithm is introduced using regional electricity load datasets for Taiwan as follows.

**Step 1: Collect datasets**

In this step, the annual regional electricity load data in Taiwan [13] (see Table 1, 1981–2000) are collected as research data. The training data are selected from 1981 to 1996, and the remaining data (from 1997 and 2000) are used as testing data for each region.

**Step 2: Build “one-step-ahead” forecast model**

In this step, the “one-step-ahead” forecasting method is applied to each region as follows.

<table>
<thead>
<tr>
<th>Year</th>
<th>Northern region</th>
<th>Central region</th>
<th>Southern region</th>
<th>Eastern region</th>
</tr>
</thead>
<tbody>
<tr>
<td>1981</td>
<td>3388</td>
<td>1663</td>
<td>2272</td>
<td>122</td>
</tr>
<tr>
<td>1982</td>
<td>3523</td>
<td>1829</td>
<td>2346</td>
<td>127</td>
</tr>
<tr>
<td>1983</td>
<td>3752</td>
<td>2157</td>
<td>2494</td>
<td>148</td>
</tr>
<tr>
<td>1984</td>
<td>4296</td>
<td>2219</td>
<td>2686</td>
<td>142</td>
</tr>
<tr>
<td>1985</td>
<td>4520</td>
<td>2190</td>
<td>2829</td>
<td>143</td>
</tr>
<tr>
<td>1986</td>
<td>5013</td>
<td>2638</td>
<td>3172</td>
<td>176</td>
</tr>
<tr>
<td>1987</td>
<td>5745</td>
<td>2812</td>
<td>3351</td>
<td>206</td>
</tr>
<tr>
<td>1988</td>
<td>6320</td>
<td>3265</td>
<td>3655</td>
<td>227</td>
</tr>
<tr>
<td>1989</td>
<td>6844</td>
<td>3376</td>
<td>3823</td>
<td>236</td>
</tr>
<tr>
<td>1990</td>
<td>7613</td>
<td>3655</td>
<td>4256</td>
<td>243</td>
</tr>
<tr>
<td>1991</td>
<td>7551</td>
<td>4043</td>
<td>4548</td>
<td>264</td>
</tr>
<tr>
<td>1992</td>
<td>8352</td>
<td>4425</td>
<td>4803</td>
<td>292</td>
</tr>
<tr>
<td>1993</td>
<td>8781</td>
<td>4594</td>
<td>5192</td>
<td>307</td>
</tr>
<tr>
<td>1994</td>
<td>9400</td>
<td>4771</td>
<td>5352</td>
<td>325</td>
</tr>
<tr>
<td>1995</td>
<td>10,254</td>
<td>4483</td>
<td>5797</td>
<td>343</td>
</tr>
<tr>
<td>1996</td>
<td>10,719</td>
<td>4935</td>
<td>6369</td>
<td>363</td>
</tr>
<tr>
<td>1997</td>
<td>11,222</td>
<td>5061</td>
<td>6336</td>
<td>358</td>
</tr>
<tr>
<td>1998</td>
<td>11,642</td>
<td>5246</td>
<td>6318</td>
<td>397</td>
</tr>
<tr>
<td>1999</td>
<td>11,981</td>
<td>5233</td>
<td>6259</td>
<td>401</td>
</tr>
<tr>
<td>2000</td>
<td>12,924</td>
<td>5633</td>
<td>6804</td>
<td>420</td>
</tr>
</tbody>
</table>
(as in Equation 15) [25] is applied to the proposed model. From Equation 15, the actual data to the period prior to the forecast data are used as the condition attribute.

\[ S_t(X_t) = (X_{t-1}, X_{t-2}, X_{t-3}, X_{t-4}, \ldots), \]
\[ S_{t+1}(X_t) = (X_t, X_{t-1}, X_{t-2}, X_{t-3}, \ldots), \]  \( (15) \)

where \( S_t(X_t) = (X_{t-1}, X_{t-2}, X_{t-3}, X_{t-4}, \ldots) \) denotes that if \( X_{t-1} \) (electricity load in 1996) and \( X_{t-2} \) (electricity load in 1995) and \( X_{t-3} \) (electricity load in 1994) and \( X_{t-4} \) (electricity load in 1993)… and \( X_{t-16} \) (electricity load in 1981), then \( S_t \) (electricity load in 1997) (i.e., use 16 data to build one step ahead model), and \( X_{t-1} \) represents the actual value for period \( t-1 \), and \( S_t \) represents the forecast value for period \( t \).

**Step 3:** Decompose the input variable by EMD

To obtain interpretable information on the input variable, this paper uses EMD to decompose the input variable (the right part of equation (15)) into a finite set of IMFs (the residual \( r_{n+1}(t) \) is also considered an IMF). In this study, each input variable is decomposed into 2 IMFs that exhibit stable and regular variation. Thus, the interruption and coupling between information on various characteristics that are embedded in the original data have been weakened. As a result, the forecasting model is easier to build.

**Step 4:** Build SVR forecasting model

In this step, this paper applies SVR with the \( \varepsilon \)-insensitive loss function (\( \varepsilon \)-SVR) to construct the forecasting model. The RBF function is adopted as the kernel function, because it can handle nonlinear and high-dimension data. To build the forecasting model, 3 parameters are set: the loss function \( \varepsilon \), the \( \gamma \) of the RBF kernel, and the penalty coefficient \( C \). To obtain a better forecasting model, this study uses a GA to optimize the parameters \( \varepsilon \gamma \) and \( C \).

**Step 5:** Optimize parameters by GA under minimal MAPE of SVR

To increase the forecasting accuracy, this step uses a GA to optimize the parameters \( \varepsilon \), \( \gamma \), and \( C \) in the SVR model under minimal MAPE (as Eq. 16) for the training data. MAPE is defined as:

\[ MAPE = \frac{1}{n} \sum_{i=1}^{n} \left| \frac{actual(t) - forecast(t)}{actual(t)} \right| \times 100\% \]  \( (16) \)

where \( actual \) \( (t) \) denotes the real electricity load value, \( forecast \) \( (t) \) is the predicted electricity load, and \( n \) is the number of data points.

Six substeps constitute operation of the GA processes as follows:

**Step 5.1:** Initialize the parameter for the GA

The initial population, with 80 individual solutions, is randomly generated in this step. Search for the optimal values \( \varepsilon \), \( \gamma \), \( C \) by GA. \( \varepsilon \), \( \gamma \), \( C \) are encoded into a chromosome. Take the coding processes of, for example; the procedure for coding \( \varepsilon \) is as follows:

This step sets the continuous parameter \( \varepsilon \) with a precision of 1/1000. In this case, 10 binary bits are required in each chromosome to express them with 1/1000 precision \((512 = 29 < 1000 \leq 210 = 1024)\). The 10-bit binary numbers (i.e., \((1001001001)2\)) can be transformed into decimal floating numbers (i.e., \((585)10\)). The component values of the chromosome for each population are initiated into random values before the search process.

In addition, the probability of crossover, the probability of mutation, and the maximum numbers of generations are initially set to 0.8, 0.08, and 2000.

**Step 5.2:** Evaluate each chromosome’s fitness

Each chromosome is evaluated using a defined fitness function (MAPE) to determine the goodness for each solution of the population.

**Step 5.3:** Check the stop criterion

This study sets the stop rule as follows. If either of the following conditions is reached, then the GA process is terminated:
1. The maximum number of generations is achieved.
2. The optimized solution under minimal RMSE is obtained. This study sets the minimal RMSE to 0.00001.

If the criterion is not reached, then a new iterative loop (Step 5.2 to 5.5) is repeated.

**Step 5.4:** Select the parents by fitness function

Selection is a process of choosing the chromosome with greater fitness for transmission to offspring and eliminating the poor-fitness chromosome to limit its inheritance by offspring. Roulette wheel selection is applied to choose the chromosomes for reproduction.

**Step 5.5:** Perform crossover and mutation

Recombine parents to generate offspring and mutate offspring. In this step, the proposed model uses one-point crossover. Then, it randomly chooses a member of
the population and changes one randomly chosen bit in its bit string representation.

**Step 6:** Forecast testing datasets using the trained model

The ε-SVR parameters of the forecasting models are determined when the stop criterion from Step 5 is reached; then, the training forecasting model is used to forecast $S_i, S_{i+1}, \ldots$, for the target training and testing datasets, respectively.

**Step 7:** Evaluation and comparison

Calculate MAPE values in the testing datasets per Equation (16). Then, MAPE is taken as the evaluation criterion for comparison with other models.

### 4. Experiments and Comparisons

In this section, 2 case studies on forecasting electricity loads are presented to validate the superiority of the proposed model in sections 4.1 and 4.2.

#### 4.1 Case Study of Forecasting Electricity Load in Taiwan

The 4 annual regional (north, center, south, and east) electricity load datasets in Taiwan from 1981 to 2000 [13] are the experimentation datasets that are used to verify the proposed model. The subdatasets from the previous 16 years are used as for training, and the remaining datasets (1997 to 2000) are collected for testing. Therefore, 16 years of electricity load data are used as the training data to forecast electricity loads in the next year.

To compare the performance of the proposed model with a traditional time series model (AR model), this paper uses the E-Views software package to fit the time-series models for various lags and orders of annual electricity load in the 4 regions; based on the statistical test, the order of AR for the electricity load model is 2. Therefore, AR (2) is the comparison model for verify the proposed model.

After decomposing the electricity load data into IMFs by EMD, this study uses IMFs as input variables to construct the forecasting model by SVR and uses the GA to optimize the parameters for SVR. These parameters for various regional datasets are shown in Table 2.

The performance of the proposed model is compared with that of several models: the recurrent support vector machine with genetic algorithm (RSVMG) model [16], back propagation (BP) neural network model [13], regression model [45], hybrid ellipsoidal fuzzy systems for time series forecasting (HEFST) model [16], AR (2) model [23], and ε-SVR model [15]. The forecasting performance of these models for the regional electricity loads of 4 regions in Taiwan is listed in Table 3; the proposed model outperforms these models for all 4 regions. In Table 3, the last column is the average forecasting performance of the comparison models; we can see that the proposed model outperforms these models.

#### 4.2 Case Study on Forecasting Electricity Loads in Poland

In this section, datasets on electricity load in Poland [46] that contain daily average load values are used as an illustrative example to evaluate the effectiveness of the proposed model. The training size, test size, and unit of the datasets are 1400, 201, and MW (10^6 Wh), respectively.

Further, a regression model [45], ε-SVR model [15], denoised SVR [47], and Yaslan and Bican’s model [47] are compared with the proposed model. The optimized parameters of the proposed model by GA in various regional datasets are shown in Table 2.

The performance of the proposed model is compared with that of several models: the recurrent support vector machine with genetic algorithm (RSVMG) model [16], back propagation (BP) neural network model [13], regression model [45], hybrid ellipsoidal fuzzy systems for time series forecasting (HEFST) model [16], AR (2) model [23], and ε-SVR model [15]. The forecasting performance of these models for the regional electricity loads of 4 regions in Taiwan is listed in Table 3; the proposed model outperforms these models for all 4 regions. In Table 3, the last column is the average forecasting performance of the comparison models; we can see that the proposed model outperforms these models.

### Table 2. Optimal parameters of the proposed model by GA in various regional datasets

<table>
<thead>
<tr>
<th>Parameter</th>
<th>North</th>
<th>Center</th>
<th>South</th>
<th>East</th>
</tr>
</thead>
<tbody>
<tr>
<td>ε</td>
<td>0.02</td>
<td>0.01</td>
<td>0.01</td>
<td>0.03</td>
</tr>
<tr>
<td>γ</td>
<td>22</td>
<td>0.91</td>
<td>0.31</td>
<td>31</td>
</tr>
<tr>
<td>C</td>
<td>91</td>
<td>81</td>
<td>3</td>
<td>53</td>
</tr>
</tbody>
</table>

### Table 3. Comparison of models for various regional electricity loads in Taiwan

<table>
<thead>
<tr>
<th>Models</th>
<th>North</th>
<th>Center</th>
<th>South</th>
<th>East</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>RSVMG</td>
<td>0.76</td>
<td>1.31</td>
<td>1.71</td>
<td>1.87</td>
<td>1.41</td>
</tr>
<tr>
<td>ANN</td>
<td>1.06</td>
<td>1.34</td>
<td>2.34</td>
<td>3.23</td>
<td>1.99</td>
</tr>
<tr>
<td>Regression</td>
<td>2.54</td>
<td>8.52</td>
<td>8.29</td>
<td>4.1</td>
<td>5.86</td>
</tr>
<tr>
<td>HEFST</td>
<td>0.35</td>
<td>0.94</td>
<td>1.85</td>
<td>1.46</td>
<td>1.15</td>
</tr>
<tr>
<td>AR(2)</td>
<td>3.87</td>
<td>3.17</td>
<td>3.53</td>
<td>5.14</td>
<td>3.93</td>
</tr>
<tr>
<td>ε-SVR</td>
<td>1.92</td>
<td>2.27</td>
<td>6.80</td>
<td>3.86</td>
<td>3.71</td>
</tr>
<tr>
<td>Proposed model</td>
<td>0.06*</td>
<td>0.05*</td>
<td>0.04*</td>
<td>0.07*</td>
<td>0.06*</td>
</tr>
</tbody>
</table>

Note: *Best performing among the 7 models.
parameters ε, γ, C of the proposed SVR model by GA in the training datasets are 0.03, 19, and 61, respectively. Table 4 shows the forecasting performance of the comparison models for electricity loads in Poland; the proposed model outperforms these models with regard to forecasting electricity loads.

5. Conclusions

Electricity is a core requirement for industries and families. Forecasting electricity loads is fundamental for supplying electricity at minimal cost. This paper has presented a load forecasting model by integrating one-step-ahead, EMD, SVR, and GA. The main contribution of the paper is that it proposes a novel method and a simple approach for making stable predictions with fluctuating data. The proposed method builds a one-step-ahead time-series model and then decomposes raw data into more stationary and regular components (IMF) using the EMD technique. Further, the corresponding GA-SVR model for each divided component is easier to construct.

This paper has compared the proposed method with several tools, using MAPE as its criteria. Based on the experimental results, the proposed EMD-SVR model generates the lowest forecasting error in the datasets. The proposed model surpasses ε-SVR, the regression model, and other forecasting models. Therefore, the proposed method is an efficient method for forecasting electricity loads, especially with nonlinear and noisy data. Further, our results are applicable for planning and operating energy systems. In subsequent research, artificial intelligence methods can be applied to the proposed model to improve its forecasting performance.

Table 4. Comparisons of models for electricity loads in Poland

<table>
<thead>
<tr>
<th>Models</th>
<th>MAPE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regression</td>
<td>7.19</td>
</tr>
<tr>
<td>ε-SVR</td>
<td>3.78</td>
</tr>
<tr>
<td>Denoised SVR</td>
<td>3.36</td>
</tr>
<tr>
<td>Yaslan and Bican’s model</td>
<td>2.83</td>
</tr>
<tr>
<td>Proposed model</td>
<td>2.71*</td>
</tr>
</tbody>
</table>

Note: * Best performing among the 5 models.

References


[10] Pino, R., Parreno, J., Gomez, A. and Priore, P., “Forecasting Next-day Price of Electricity in the Spanish...


Manuscript Received: Aug. 24, 2016
Accepted: Aug. 15, 2017