Appraisal of the support vector machine to forecast residential heating demand for the District Heating System based on the monthly overall natural gas consumption

Nima Izadyar a, Hossein Ghadamian b, Hwai Chyuan Ong a, **, Zeinab moghadam c, Chong Wen Tong a, Shahaboddin Shamshirband d, *

a Department of Mechanical Engineering, Faculty of Engineering, University of Malaya, 50603 Kuala Lumpur, Malaysia
b Department of Energy, Materials and Energy Research Center (MERC), P.O. Box: 14155-4777, Tehran, Iran
c School of Engineering, Edith Cowan University, Joondalup, Perth, WA 6027, Australia
d Department of Computer System and Information Technology, Faculty of Computer Science and Information Technology, University of Malaya, 50603 Kuala Lumpur, Malaysia

A R T I C L E    I N F O

Article history:
Received 24 April 2015
Received in revised form 2 October 2015
Accepted 6 October 2015
Available online 19 November 2015

Keywords:
Residential natural gas demand
DHS (District heating system)
Estimation
Wavelet and firefly algorithms (FFAs)
SVM (Support vector machine)

A B S T R A C T

DHS (District Heating System) is one of the most efficient technologies which has been used to meet residential thermal demand. In this study, the most accurate forecasting of the residential heating demand is investigated via soft computing method. The objective of this study is to obtain the most accurate prediction of the residential heating consumption to employ forecasting result for designing optimum DHS system as a possible substitute of a pipeline natural gas in BAHARESTAN Town. For this purpose, three Support Vector Machine (SVM) models namely SVM coupled with the discrete wavelet transform (SVM-Wavelet), the firefly algorithm (SVM-FFA) and using the radial basis function (SVM-RBF) were analyzed. The estimation and prediction results of these models were compared with two other soft computing methods (ANN (Artificial Neural Network) and GP (Genetic programming)) by using three statistical indicators i.e. RMSE (root means square error), coefficient of determination ($R^2$) and Pearson coefficient ($r$). Based on the experimental outputs, the SVM-Wavelet method can lead to slightly accurate forecasting of the monthly overall natural gas demand.

© 2015 Elsevier Ltd. All rights reserved.

1. Introduction

Nowadays, residential heating demand including heating and domestic hot water is an important section of the total natural gas consumption in the world. In Iran, the proportion of the residential heat demand is about 20 percent of the overall natural gas consumption [1], and it is also about four times higher than the global average of the residential natural gas consumption. Therefore, finding the alternatives for generating residential heating demands is to replace the national grid of natural gas. DHS (District Heating System) is one of this alternatives for producing residential heat demand which includes several important benefits such as high efficiency (savings energy up to 30%), cost-effective, clean (CO$_2$ emission reduction) and compatible in terms of environment. For designing DHS, the first step is to forecast residential heating demand of the case study, because a more accurate forecasting of the DHS demand can lead to more precise DHS simulation and design. So far several authors have investigated demand forecasting as a crucial part of optimizing the DHS. Ma et al. [2] have proposed a methodology which considers time and type of the buildings as two key factors of the energy consumption patterns. Based on this method, a Gaussian mixture model has been developed to scrutinize the influence of the time and type of the buildings on the heat consumption patterns and then the procedure has been validated by data measuring and a real DHS. The results of this study indicate that in contradiction with time factor, building classification is the effective key factor in the heat consumption patterns. Another study has been accomplished by Park et al. [3] who have applied three forecasting techniques to predict heat demand during one week (short period) for effective management of the DHS. Two months heat consumption and ambient temperature data have been employed to forecast short term data by using Support Vector Regression (SVR), Partial Least Squares (PLS), and Artificial Neural
Network (ANN), and further the prediction of heat consumption was compared with real data for validating models. In the same year, an online short-term prediction technique has been proposed by Grosswindhager et al. [4] to forecast heat load for the DHS. In this research, the SARIMA (Seasonal Autoregressive Integrated Moving Average) has been applied for the short-term forecasting and then the accuracy of this method was validated by the standard regression diagnostics. Other factors such as ambient temperature and social component were also simulated by the SARIMA. Results of this research show that the proposed method can be applied as an appropriate procedure for the online short-term forecasting.

In the context of natural gas demand forecasting, an accurate short demand forecasting model via a new procedure has been introduced by Lai et al. [5]. In that study, a combination of the SVM-Wavelet and Genetic Algorithm (GA) named (GA Wv-SVM) has been applied to predict gas demand, and GA was applied to obtain the Wv-SVM parameters. The outcome of this proposed model demonstrates that this model is more accurate than ANN and Wv-SVM. The wavelet transform has been also combined with other soft computing methods to predict future demand of energy, as an example, wavelet transform, evolutionary algorithm and neural network was combined by N. Amjady et al. [6] to obtain short-term hourly forecasting of the load in the power system. Furthermore, some other soft computing algorithms have been employed to obtain the accurate forecasting of the natural gas consumption in various fields, for example, Miha Kovacic et al. [7].

The dependency of the daily gas demand and ambient temperature has been investigated by Jinsoo Park [8]. For this purpose, two forecasting natural gas demand methods including an improved autoregressive model and a model based on the changing in the ambient temperature were proposed in the first step. In the second step relation between temperatures and heating demand was studied to forecast the daily gas demand. However, the proposed models did not meet the constraints; therefore the weighted average model including combination of these models has been proposed to obtain the daily demand forecasting. In addition, this relevance of gas demand and ambient temperature has been investigated in another research by Jolanta Szoplik et al. [9] that used Multilayer Perceptron of the Artificial Neural Networks (MLP-ANNs) to obtain the accurate forecasting for the natural gas consumption in a case study. In this research, calendar parameters and ambient temperature have been considered and used as two factors which affect the natural gas demand. The results show the accuracy of the forecasting demand is influenced by changing the number of neurons in the hidden layer, so more neurons lead to higher accuracy in the forecasting result. The MLP-ANNs was also employed by Taspinar et al. [10] to obtain short-term demand forecasting with time series approach and four years meteorological data (such as temperature, moisture and etc.). The outcomes of this study indicate that the forecasting method has adequate accuracy.

Three various AI (Artificial Intelligence) including ARIMA (Autoregressive Integrated Moving Average), ANFIS (Adaptive Neuro Fuzzy Inference System) (which combines ANN and Fuzzy Inference System), and MLP-ANN and Radial Basis Function Network ANN (RBFN-ANN) were used to predict weekly natural gas demand [11]. A comparison between validation results of these models indicates that ANFIS and MLP-ANN are the most accurate techniques among these models respectively.

In summary, it can be seen from the previous studies that soft computing methods are usually employed to obtain the short-term (weekly, daily, hourly or even less than hour) prediction of the natural gas consumption. This issue has been considered for the sake of the change in the different weather data (especially ambient temperature) which leads to change in the heat and thermal power demand amount. In the current study, the fluctuations of the average monthly outdoor temperature have been considered as a part of the forecasting methods and the relation between natural gas consumption and ambient temperature has been practiced.

There is a requisite for putting developed procedure for estimating the residential heating via previous natural gas demand data of the case study. This procedure should be simpler and more accurate than conventional approaches. Consequently, soft computing methods such as SVM are potential to be applied to forecast an overall monthly natural gas consumption model as a residential heating demand of the DHS. The SVM is a novel learning algorithm which was broadly applied in the various areas [12–15] and for various purposes such as the pattern recognition, forecasting, categorization and regression analysis [16–18]. The nature of the experimental data affects the adoption of the kernel function, but the main employed kernels contain linear, polynomial inner product functions and the RBF (Radial Basis Function) [19]. Biological inspired metaheuristic optimization algorithms like GA (Genetic Algorithm), ACO (Ant Colony Optimization), and PSO (Particle Swarm Optimization) have been extensively used in different fields of science [20–31]. These techniques are obtained from the approach of adopting the most suitable characteristics in biological systems. The latest methods in biology provoked from metaheuristic optimization techniques are FFA (Firefly Algorithm) (FFA) which has been described by Yang [32]. Among the biological based optimization methods, the FFA is one of the most practical and appropriate methods which can be applied to find local and global optimum comparison [32–40]. The prediction accuracy of the SVM model highly depends on the definition of the suitable model parameters, therefore, selecting procedure and arrangement of the model parameters are two requirements for the model construction, and hence the FFA is used for determination of SVM parameters.

In the current study, the SVM was coupled with DWT (Discrete Wavelet Transform) which has many useful basic functions to select from depending on the signal that is being analyzed. The input of the SVM method is obtained by using Wavelet analysis which decomposes the data time series into its different elements. Over the past few years, this technique has received great interest for engineering applications [41,42]. The objectives of the current study are to construct, develop and appraise the results of Wavelet, the firefly algorithm and the radial basis function of SVM (called SVM-Wavelet, SVM-FFA and SVM-RBF respectively) for predicting monthly natural gas consumption. An attempt is made to retrieve correlation for the two outputs: months of the year and average temperature in each month.

2. Methodology

2.1. Experimental setup

In this study, BAHARESTAN Town (located in the HALJROOD area of KAMALSHAHR in the north east of KARAJ city along the KARAJ-QAZVIN highway) is adopted as a case study. The main portion of the natural gas demand of this town is the industrial demand consumed by several factories located in this location, while this location was chosen based on some reasons which have been listed as below:

> Heating resource for generating thermal energy is required for the town which would be based on the DHS system. This place is located near some power plants that their thermal energy waste can be employed as a source of thermal energy production of the DHS. Clearly, the distance of the supply side and demand side in the DHS should be near in order to reduce pipeline costs and waste of energy.
The transmission pipeline should be designed in the best and evaluated route. Thanks to KAMALSHAHR gas department, we have learned about the map pipeline of the BAHARESTAN town and all of the units in this location, so we could design the DHS pipeline according to this route and other precise information of the units.

Geographic coordinates of the BAHARESTAN Town is (35° 53' 33''N 50° 53' 34''E) as demonstrated in Fig. 1.

Case study area categorizations are summarized in Table 1.

Case study area areas are categorized in Table 1.

The current thermal demand of this town has been met by the means of urban natural gas pipeline for two various purposes including industrial and residential consumptions. Because of the different levels of consumption and also price of the natural gas in the industrial and residential sections of Iran, there are two separate bills for each unit per month. Accordingly, the previous residential heating demands of the units are obtained via gathering residential natural gas bills through six years (72 months). Total residential gas consumption of the town has been obtained by gathering monthly natural gas bill of each unit for six years (Jan 2006 to end of Dec 2011) and summation of all units for monthly demand of the case study. The overall monthly consumption of the case study is illustrated in Fig. 2.

As mentioned in the introduction, meteorological data like ambient temperature has influence on the residential demand. Natural residential gas consumption is usually divided into two parts including domestic (hot water for shower or gas for cooking) and heating consumption. The domestic natural gas demand is constant approximately through the year, but heating consumption increased in the cold months. As a result, temperature has a significant impact on the overall residential natural gas demand. The monthly temperature has been gathered for the same period of time as indicated in Fig. 3.

2.2 Input and output variables

The input and output variables used to predict monthly consumption of the natural gas have been determined and gathered in Table 2.

<table>
<thead>
<tr>
<th>Sector</th>
<th>Area (m²)</th>
</tr>
</thead>
<tbody>
<tr>
<td>User space</td>
<td>837,408</td>
</tr>
<tr>
<td>Service area</td>
<td>264,177</td>
</tr>
<tr>
<td>Green space</td>
<td>311,700</td>
</tr>
<tr>
<td>Passages area</td>
<td>317,500</td>
</tr>
<tr>
<td>Total area</td>
<td>1,730,785</td>
</tr>
</tbody>
</table>

Table 1

In fact, 70%, 15% and 15% of whole data per month are considered as training, testing and validation data respectively. In the current study, all month’s data are included in the training data of the models firstly and then testing and validation data in order to cover all possible situations.

2.3 Soft computing methodologies

2.3.1 SVM (Support vector machine)

Concept of the SVM method is based on statistical learning procedure and minimization of the structural risk [43]. By this concept, the upper boundary extension error is minimized by SVM, while the local training error is usually minimized by the other common machine learning methodologies. In addition, the SVM method has some other benefits compared with other regular algorithms such as, unique solution and employing multi-dimensional spaced kernel functions’ collection which comprises wise non-linear transformation.

In the SVM structure, a collection of data is introduced by \( (x_i, d_i) \) \(^n\), where \( x_i \) is a vector which contains the input sample data, \( d_i \) gives the target value and \( n \) represents the data size. The mathematical structure of the SVM can be indicated by the following equations.

\[
f(x) = w\phi(x) + b \tag{1}
\]

\[
R_{SVM}(C) = \frac{1}{2}||w||^2 + C \frac{1}{n} \sum_{i=1}^{n} L(x_i, d_i) \tag{2}
\]
In Equation (1), $x$ represents input vector which is charted by $\phi(x)$ that represents a multi-dimensional space specification, $w$ denotes a regular vector and $b$ represents a scalar. In the second formula, $C_i \sum_{i=1}^{n} L(x_i, d_i)$ signifies the experimental error. Two parameters including $w$ and $b$ can be assessed by minimizing risk function and after the slack variables ($\xi_i$ & $\xi_i^*$) definition as can be seen in Equation (3). The slack variables including $\xi_i$ and $\xi_i^*$ represent upper and lower permitted deviation boundaries, respectively.

$$\text{Minimize } R_{\text{SVM}}(w, \xi^*) = \frac{1}{2} \|w\|^2 + C \sum_{i=1}^{n} (\xi_i + \xi_i^*)$$

Subject to

$$d_i - w \phi(x_i) + b_i \leq \epsilon + \xi_i$$

$$w \phi(x_i) + b_i - d_i \leq \epsilon + \xi_i^*$$

$$\xi_i, \xi_i^* \geq 0, \ i = 1, \ldots, l$$

Where, $\frac{1}{2} \|w\|^2$ represent the adjustment parameter, $C$ denotes the error forfeit parameter which is employed to control the difference between the adjustment parameter and experimental error. In addition, $\epsilon$ signifies loss function which equals with training data guesstimate accuracy, lastly number of training data set is indicated by $l$.

Equation (1) can be dissolved via the Lagrange multiplier and optimum restrictions, therefore a generic function can be obtained by using the following relation:

$$f(x, a_i a_i^*) = \sum_{i=1}^{n} (a_i - a_i^*) K(x, x_i) + b$$

where, $K(x, x_i)$ is entitled the kernel function obtained by $K(x, x_i) = \phi(x_i) \phi(x_i)$, which is created by two internal vector including $x_i$ and $x_i$ in the sub space $\phi(x_i)$ and $\phi(x_i)$.

In the SVM procedure, four kernel functions are supplied which called linear, sigmoid, polynomial and Radial Basis Function (RBF). Among these kernel functions, RBF is the best kernel in terms of computational proficiency, easiness, simple adoption, reliability and compatibility in handling intricate parameters [44,45]. The non-linear radial basis kernel function is defined as:

$$K(x_i, x_j) = \exp\left(-\gamma \|x_i - x_j\|^2\right)$$

where, variable $x_i$ and $x_j$ represents the input vectors, such as features vectors which calculated from testing data set, $\gamma$ regulates the tradeoff between minimization of fitting error and smoothness of the appraised function.
The forecasting accuracy via RBF depends on the subparameters containing \((\gamma, \varepsilon, C)\). In the current study, the optimum values of the mentioned parameters are obtained by means of firefly optimization algorithm.

2.3.2. SVM firefly optimization procedure

The FFA procedure is founded on certain patterns, i.e. the fireflies flashing characteristic. A firefly is the name of an insect which employs bioluminescence to entice quarry and it causes that the other fireflies can trail its track for searching their hunt. Accordingly, the firefly algorithm can be developed based on the concept of this luminance generation.

Formulation of the objective and the light intensity disparities are the most important matters in terms of the FFA progress. For instance, in optimum design which includes objective maximization, the fitness function depends on the illumination or the light’s amount which was released by the firefly. Accordingly, the light strength reduction leads to remoteness among the fireflies due to intensity variations of and so scamps the desirability among them. The following relation indicates the relation between different distances \((r)\) and the light intensity \((I)\).

\[
I(r) = I_0 \exp(-\gamma r^2)
\]  

(6)

where, \(I\) is the strength of light at distance \(r\) from a firefly, \(I_0\) signifies initial strength of glory in zero distance and \(\gamma\) indicates the light sorption factor. The attractiveness of the firefly \((\beta)\) in the distance of \(r\) is computed by the following formula:

\[
\beta(r) = \beta_0 \exp(-\gamma r^2)
\]  

(7)

Clearly, \(\beta_0\) signifies the attractiveness at \(r = 0\).

In addition, the Cartesian distance \((r_{ij})\) between every two fireflies \(i\) and \(j\) is calculated by the following equation.

\[
r_{ij} = \left\| x_i - x_j \right\| = \sqrt{\sum_{k=1}^{d} (x_{i,k} - x_{j,k})^2}
\]  

(8)

The amount of firefly \(i\) motion to the firefly \(j\) with more luminance can be signified by the following relation.

\[
\Delta x_i = \beta_0 e^{-\gamma r^2} (x_j - x_i) + \alpha e_i
\]  

(9)

As can be seen from Equation (9), it comprises two parts which first represents the attraction which has been explained in the current section and the second part indicates the randomization. Accordingly, \(\alpha\) represents the randomization factor and \(e_i\) denotes the random number vector which can be concluded from a Gaussian repartition. The subsequent motion of the firefly \(i\) is concluded as follow:

\[
x_i^{t+1} = x_i + \Delta x_i
\]  

(10)

The rudimentary steps of the FFA evolvement are illustrated by following algorithm which is indicated in Fig. 4.

2.3.3. Discrete wavelet transform (DWT)

Wavelet transform is a signal step processing method which is extended by using the Fourier. The wavelet transform algorithm signifies a mathematical explanation which divides a time-series frequency signal into various parts. This method contains some significant benefits, for instance a full analysis of the concluding decomposed parts with high accuracy is achieved via Fourier transform. It helps to enhance capacity of the prototype model, because the required information is obtains at various levels [46].
**Firefly Algorithm**

**start**

- Define the objective function, \( f(x), x = (x_1, \ldots, x_d) \)
- Generate initial population of fireflies \( x_i, (i = 1, 2, \ldots, n) \)
- Determine light intensity \( I_i \), at \( x_i \) from \( f(x_i) \)
- Define light absorption coefficient \( \gamma \)
- while \( t < \text{Max Generation} \)
- Make a copy of population for movement function
  - for \( i = 1 : n \) all \( n \) fireflies
    - if \( (t_i > t) \)
      - Move fireflies \( i \) and \( j \) in \( d \)-dimension;
    - end
  - end
- Attractiveness varies with distance \( r \) via \( \exp(-r \gamma) \)
- Evaluate new solution and update light intensity
- end
- Rank the fireflies and find the current best
- end
- Post process results and visualization

**end**

Fig. 5. Pseudo-code for firefly algorithm.

This technique is appropriate for examining data in the frequency and time domains due to its capability of elude data from non-intermittent and glancing signals; therefore it is so beneficial in time-frequency localization.

Continuous wavelet transform of a signal \( f(t) \) is a signal processing time scale technique which is described as the integral of all signals within entire period multiplied by the scaled. The shifted version of the wavelet function \( \psi(t) \) is indicated by the following relation.

\[
W_s(a, b, \psi) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} f(t) \psi^* \left( \frac{t-b}{a} \right) dt.
\] (11)

Where, \( \psi(t) \) represents primary wavelet function, \( a \) denotes scale index (like reverse of the frequency), and \( b \) is the time shifting or translation factor. This algorithm (DWT) is developed after obtaining the scale index (\( a \)) and time shifting (\( b \)) parameters which can be defined as follow.

\[
a = a_0^m, \quad b = na_0^m b_0
\] (12)

Where, \( n \) and \( m \) are integer variables. When substituting \( n \) and \( m \) in the previous relation, the following relation is obtained.

\[
W_s(m, n, \psi) = a_0^{m/2} \int_{-\infty}^{\infty} f(t) \psi^* \left( a_0^m t - nb_0 \right) dt
\] (13)

In this study, wavelet analysis is employed according to Equations (11) and (12) to decay time series of the overall gas consumption data to individual components, where the decomposed components can be considered for data inputs into the SVM model expressed in Equations. (1)–(5). Fig. 6, depicts the flow chart for obtaining the optimal SVM parameters using the Wavelet algorithm.

### 2.4. Evaluating accuracy of proposed models

Predictive performances of proposed models were presented as root means square error (RMSE), coefficient of determination (R²) and Pearson coefficient (\( r \)). These statistics are defined as follows:

1) **Root Mean Square Error (RMSE)**

2) **Pearson correlation coefficient (\( r \))**

\[
r = \frac{n \left( \sum_{i=1}^{n} O_i - \bar{O_i} \right) \left( \sum_{i=1}^{n} P_i - \bar{P_i} \right)}{\sqrt{\left( n \sum_{i=1}^{n} O_i^2 - \left( \sum_{i=1}^{n} O_i \right)^2 \right) \cdot \left( n \sum_{i=1}^{n} P_i^2 - \left( \sum_{i=1}^{n} P_i \right)^2 \right)}}
\] (15)

3) **Coefficient of determination (R²)**

\[
R^2 = \frac{\left( \sum_{i=1}^{n} (O_i - \bar{O_i}) \cdot (P_i - \bar{P_i}) \right)^2}{\left( \sum_{i=1}^{n} (O_i - \bar{O_i}) \cdot \sum_{i=1}^{n} (P_i - \bar{P_i}) \right)^2}
\] (16)

Where, \( P_i \) and \( O_i \) denotes the empirical and predicated values of residential heating demand of the district heating system, respectively, \( n \) represents the overall number of training data.

While some controversy regarding the validity of this error metrics do exist, it pertains to comparison of forecasting methods across different data sets [47,48], what is not the case in this work.

### 3. Results and discussion

#### 3.1. Performance evaluation of proposed SVM models

The RBF has been employed as the kernel function to forecast the surface roughness in current research. The accuracy of the SVM method is affected by selecting the parameters (\( C, \gamma \) and \( \varepsilon \)), so the optimum amounts of these parameters are evaluated for the proposed method and summarized in Table 3. The average computational time for the SVM-Wavelet modeling was around 321 s using a PC with Intel Core Duo CPU E7600 @3.06 GHz and (2 GB) RAM.

![Fig. 6. Flow chart of proposed Wavelet-based parameter determination approach for the SVM.](Image)

### Table 3

User-defined parameters for SVM models.

<table>
<thead>
<tr>
<th>SVM models</th>
<th>Parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>SVM-Wavelet</td>
<td>( C = 1.45; \gamma = 0.342; \varepsilon = 0.34 )</td>
</tr>
<tr>
<td>SVM-RBF</td>
<td>( C = 2.47; \gamma = 0.67; \varepsilon = 0.62 )</td>
</tr>
<tr>
<td>SVM-FFA</td>
<td>( C = 1.74; \gamma = 0.47; \varepsilon = 0.27 )</td>
</tr>
</tbody>
</table>
Fig. 7. Scatter plots of actual and predicted values of experimental data using (a) SVM-Wavelet, (b) SVM-RBF and (c) SVM-FFA method.
Fig. 8. Plots of actual and predicted values of experimental data using (a) SVM-Wavelet, (b) SVM-RBF and (c) SVM-FFA method.
Average computational time for the SVM-FFA and SVM-RBF models were around 332 and 432 s, respectively, using the same PC with the same performances.

To evaluate SVM model performance, measured gas consumption was schemed besides the forecasted results. Fig. 7(a) presents the accuracy of developed SVM-Wavelet for the monthly natural gas consumption of the case study, predictive model, while Fig. 7(b) and (c) present the accuracy of developed SVM-RBF and SVM-FFA on the monthly natural gas consumption, predictive models, respectively. It also indicates that majority of the points are along the oblique line for SVM-Wavelet forecasting model. It demonstrates that forecasting outcomes are in agreement with the measured data for SVM-Wavelet model and can be validated for designation coefficient. The number of either overestimated or underestimated values produced is limited and obviously the forecasted values are in the adequate accuracy. Fig. 8 shows the prediction of the overall gas consumption in relation to experimental samples. It can be noted that there is a close prediction with SVM-Wavelet and with SVM-FFA.

4. Performance comparison of SVM algorithms

In this study, accuracy of the three proposed SVM algorithms are confirmed by comparing three SVM forecasting accuracy with each other and two other strong soft computing methods (ANN). Three common indicators including RMSE, r and R² are employed to assess statistical error and compared with these five applied methods. Accordingly, accuracy of the forecasting results test data sets are evaluated based on these three indicators and summarized in Table 4. However, training error is not mentioned in this table, because training error is not credible indicator for prediction potential of particular model. As mentioned, GP (Genetic programming) [49] and ANN (Artificial Neural Network) [50,51] results are also added as comparators for the current study. According to Table 4, all SVM models outperform ANN and GP models.

SVM-Wavelet model slightly outperform SVM-RBF and SVM-FFA models according to the results in Table 4. The predictions from the SVM models correlate highly to the actual data (r > 0.85). Accuracy results, measured in terms of r, R², indicate that SVM-Wavelet predictions are more accurate than the SVM-RBF and the SVM-FFA. Regarding the RMSE indicator, results show that SVM-Wavelet is slightly better than two other methods. And, based on these outcomes, the current study is accomplished to obtain appropriate forecasting results for designing the optimum DHS for substituting with urban gas pipeline in this case study. It is clear that the accomplishment of this substituting would be dependent on feasibility study and a comparison between the DHS and current urban natural gas cost and emission. These results can be also applied on the other modern technologies which can be employed to substitute natural gas urban pipeline.

On the other hand, another goal of the current study was to obtain more accurate forecasting results via comparison of the soft computing methods. About other advantages regarding conventional methods, we have tried to mention these points in the manuscript. The main advantage of the proposed SVM models are highlighted. For example FFA model has automatically searching parameters detection. On the other hand RBF model needs manually selection of the SVM parameters which is one of the disadvantage. SVM-wavelet model separate training data in several signals in order to analyses each part separately and after training the SVM model all wavelet branches are connected.

In addition, more evaluation is required in the future studies to assess the suitability of the SVM in the context of demand forecasting for DHS in different time scales. Also, in the future works the SVM can be hybridized with other algorithms to assess the possibility of attaining higher precision and reliability in the predictions under different time scales and conditions. Finally, more combinations of variables should be analyzed to determine the most influential set of variables in this context.

5. Conclusion

This study carried out a systematic approach to create the SVM models for the monthly natural gas consumption forecasting as a residential demand of the DHS such as SVM-Wavelet, SVM-RBF and SVM-FFA. The SVM accomplishes structural minimization whereas other conventional, soft computing models centre of minimizing errors procedure, which is much less efficient. The provided SVM-Wavelet model was obtained by combining two methods, i.e., the SVM and the Wavelet transform. The RBF has been selected as the kernel function for the SVM, while the FFA was used to obtain the SVM parameters. A comparison of the SVM-Wavelet, the SVM-RBF and the SVM-FFA was performed in order to assess the prediction accuracy.

Table 4

Comparative performance statistics of the SVM-Wavelet, SVM-RBF, SVM-FFA, ANN and GP models for the overall gas consumption prediction of the case study.

<table>
<thead>
<tr>
<th>SVM model</th>
<th>RMSE</th>
<th>R²</th>
<th>r</th>
</tr>
</thead>
<tbody>
<tr>
<td>SVM-Wavelet</td>
<td>0.18171</td>
<td>0.7385</td>
<td>0.85934</td>
</tr>
<tr>
<td>SVM-RBF</td>
<td>0.1845</td>
<td>0.7291</td>
<td>0.8539</td>
</tr>
<tr>
<td>SVM-FFA</td>
<td>0.18178</td>
<td>0.7372</td>
<td>0.85859</td>
</tr>
<tr>
<td>ANN</td>
<td>0.18504</td>
<td>0.7285</td>
<td>0.85301</td>
</tr>
<tr>
<td>GP</td>
<td>0.18504</td>
<td>0.7276</td>
<td>0.8535</td>
</tr>
</tbody>
</table>

Acknowledgments

The authors would like to acknowledge the Ministry of Higher Education, Malaysia and The University of Malaya, Kuala Lumpur, Malaysia for the financial support under HRG: UM.C/HIR/MOHE/ENG/06 (D000006-16001) and SATU: RU022H-2014.

References

[1] https://www.iea.org/statistics/statisticssearch/report/?country—Iran&20k; product—balances,