International Journal of Artificial Intelligence & Applications (IJAIA), Vol.2, No.4, October 2011

FORECASTING NATURAL GAS CONSUMPTION USING PSO OPTIMIZED LEAST SQUARES SUPPORT VECTOR MACHINES

Hossein Iranmanesh[1], Majid Abdollahzade[2] and [3]Arash Miranian

[1] Department of Industrial Engineering, "University of Tehran" & "Institute for International Energy Studies", Tehran, Iran
hiranmanesh@ut.ac.ir
[2] Department of Mechanical Engineering, "K.N.Toosi University of Technology" & "Institute for International Energy Studies", Tehran, Iran
m.abdollahzade@gmail.com
[3] School of Electrical and Computer Engineering, "University of Tehran" & "Institute for International Energy Studies", Tehran, Iran
ar.miranian@gmail.com

ABSTRACT

This paper proposes an effective model based on the least squares support vector machines (LS-SVM) and the particle swarm optimization (PSO), termed PSO-LSSVM, for prediction of natural gas consumption, as an important energy resource. The salient feature of mapping nonlinear data into high dimension feature space, distinguishes LS-SVM as a powerful approach for forecasting and estimation. Optimization of the model's parameters by a fast and efficient PSO algorithm results in an optimized model which is employed for prediction of annual natural gas consumption in Iran and Unites States. Promising results were obtained for prediction of Iranian gas consumption from 1998 to 2006 and U.S. gas consumption from 2001 to 2005. Besides, comparison to an optimized multi-layer preceptron (MLP) network, using error indices of MAPE and NMSE demonstrated the superior performance of the proposed PSO-LSSVM approach.

KEYWORDS

Least square support vector machines, particle swarm optimization, natural gas consumption, forecasting

1. INTRODUCTION

Natural gas as a clean and efficient fossil fuel accounts for a considerable portion of world energy consumption. In 2008, 20.0% of total energy consumption of the world was supplied by natural gas [1]. Natural gas production and consumption are experiencing an increasing trend owing to the growth in world population and the economic development worldwide.

Due to the favourable characteristics of natural gas, such as being clean, environment friendly and highly efficient as well as its strategic status, the accurate prediction of gas consumption is crucial. Hence, various methods and approaches have been developed by the researchers for this purpose, which can be identified as deterministic or stochastic, dynamic or static and linear or nonlinear models [2].

As another classification, natural gas prediction methods can be distinguished as traditional (time series) and computational intelligence (CI)-based approaches. Liu and Lin employed time series models for forecasting consumption of natural gas in Taiwan within the residential sector [3]. In their study, they explored the relationships among residential gas consumption and several

DOI : 10.5121/ijaia.2011.2405

relevant time series variables, such as temperature of the service area and gas price, and then developed the forecast model. They provided both monthly and quarterly forecast using their developed model. A logistic curve interpretation approach was presented by Siemek et al for estimation of natural gas consumption [4]. In this approach the hypothetical natural-gas demand was described based on average trend of the economy development during recent decades. In another study, Akkurt et al used different time series models for prediction of natural gas consumption in Turkey [5]. They proposed different models such as such as exponential smoothing, winters' forecasting and Box-Jenkins methods, to forecast natural gas consumptions of Turkey in different time periods. A system dynamics model has been developed by Li et al for Forecasting the growth of Chinese natural gas consumption [6]. They applied this model to provide an outlook for Chinese gas consumption until 2030. Stochastic Gompertz innovation diffusion model, which is a statistical model, was used by Gutiérrez et al to forecast Spain natural gas consumption [2]. This approach is based on obtaining the probability density function of the process and then forecasting the future values of the process.

There are many uncertain factors influencing natural gas consumption which make gas consumption series highly complex and nonlinear [7]. Therefore, traditional linear models and statistical approaches such as linear regression or the method one proposed in [2], are not suitable for gas consumption prediction. Computational intelligence (CI) based models, including fuzzy logic, neural networks (NN) and support vector machines (SVM) are elaborate models which are effective in dealing with highly nonlinear and complex processes [8]. The CI-based models have been used for energy demand predictions to a great extent [9, 10]. Prediction of daily natural gas consumption by combination of artificial neural-network forecasters has been also carried out [11]. In this study, Khotanzad et al proposed a two-stage system with the first stage containing two NN forecasters. The second stage consisted of a combination module to mix the two individual forecasts produced in the first stage. They implemented their approach on real data from six different gas utilities.

Support vector machines, established based on the statistical learning theory, exhibit distinctive advantages to solve complex problems [12, 13]. In this paper we propose the idea of optimizing least squares support vector machines (LS-SVM) parameters using the fast and efficient algorithm of particle swarm optimization. The developed PSO-LSSVM will be used for prediction of annual natural gas consumption in Iran and United States.

2. LEAST SQUARES SUPPORT VECTOR MACHINES

Support vector machines have been developed based on the statistical learning theory by Vapnik [14]. The main theme of SVMs lies in mapping the input space into a higher dimensional feature space, and then performing the linear regression using support vector regression (SVR). The less adjustable parameters of the SVMS compared to neural networks, has made them popular for prediction, control and signal processing applications [15], [16]. Furthermore, SVMs training involves optimization of a quadratic problem with a unique solution; therefore random initialization of the model's weighting factors is prevented.

Various applications have been reported for SVMs, including pattern recognition, classification and regression analysis [17]-[19]. In case of time series prediction, SVR estimates a function using observed data and the SVMs are trained. In the rest of this paper, we restrict our attention to the mathematical formulations of the SVMs for the purpose of time series forecasting.

Consider time series $x(t)$ defined at $t = 0,1,...,N-1$ and $y(N+\Delta)$ as the predicted values in the future. The prediction function $f(x)$, defines the predicted output based on the m previous observations,

$$y(N+\Delta) = f\left(x(N-a_1), x(N-a_2), ..., x(N-a_m)\right) \tag{1}$$

where a_1, ..., a_m are time lags. By applying regression analysis the prediction function for nonlinear regression applications is defined as below,

$$f(x) = (w \cdot \phi(x)) + b \tag{2}$$

where $\phi(x)$ is the kernel function, w is the vector of weights and b is the bias. The nonlinear regression in (2) maps the input space into a higher dimension feature space by means of the kernel function and then a linear regression is performed [14]. Next, the optimal weights w and the bias b must be found through an optimization procedure, considering the proper optimization criteria, namely the flatness of the weights, measured by the Euclidean norm $\|w\|^2$ and the estimation error, defined by a loss function. Two commonly used loss functions for SVMs are ϵ-sensitive and quadratic loss functions. The latter is associated with the least squares support vector machines (LS-SVM), employed in this paper. The mathematical representation of the optimization problem for LS-SVM, given N pairs of training data (x_i, y_i), $i = 1,...,N$ is as follows,

$$\text{Minimize} \quad \frac{1}{2} w^T w + \lambda \sum_{i=1}^{N} \xi_i \tag{3}$$

$$\text{Subject to} \quad y_i - \left[w^T \phi(x_i) + b \right] = 1 - \xi_i \tag{4}$$

where λ is referred to as the regularization constant and determines the penalties to the estimation error and ξ_i are slack variables which allow for some errors in the optimization problem.

By using the Lagrange multipliers and considering the Karush-Kuhn-Tucker (KKT) conditions, the following is obtained,

$$\begin{cases} w = \sum_{i=1}^{N} \alpha_i y_i \phi(x_i) \\ \sum_{i=1}^{N} \alpha_i y_i = 0 \\ \alpha_i = \gamma \xi_i \\ y_i - \left[w^T \phi(x_i) + b \right] - 1 + \xi_i = 0 \end{cases} \tag{5}$$

Let us also define the $K(x_i, x_j)$ as the inner product of $\phi(x_i)$ and $\phi(x_j)$ vectors (called Kernel function) and consider a set of other definitions stated below:

$$\begin{cases} Z = \left[\phi(x_1)^T \ y_1, \phi(x_2)^T \ y_2, ..., \phi(x_i)^T \ y_i \right] \\ Y = [y_1, y_2, ..., y_i] \\ \xi = [\xi_1, \xi_2, ..., \xi_i]^T \\ \alpha = [\alpha_1, \alpha_2, ..., \alpha_i]^T \end{cases} \tag{6}$$

By eliminating w and γ and using (5) and (6), the following equation is obtained:

$$\begin{bmatrix} 0 & Y^T \\ Y & ZZ^T + \gamma^{-1} I \end{bmatrix} \begin{bmatrix} b \\ \alpha \end{bmatrix} = \begin{bmatrix} 0 \\ \bar{I} \end{bmatrix} \tag{7}$$

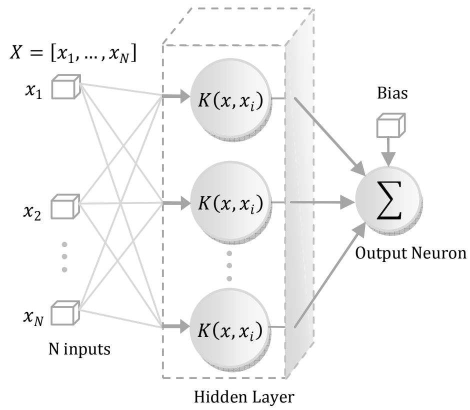


Figure 1. Structure of SVM

where $\bar{I} = [1,1,\ldots,1]^T$. By applying Mercer's condition [15] within the ZZ^T matrix, each element in this matrix will have the following form:

$$\left(ZZ^T\right)_{ij} = y_i y_j \phi(x_i)^T \phi(x_j) \tag{8}$$

Finally the resulting LS-SVM model can be represented as:

$$f(x) = \sum_{i=1}^{N} \left(\alpha_i - \alpha_i^*\right) K(x_i, x) + b \tag{9}$$

where α_i and α_i^* are Lagrange multipliers. It is noticeable that with the aforementioned definitions, there is no need to compute $\phi(x_i)$. The Kernel function which is inner product of two $\phi(x)$ functions is instead incorporated in computations. Some common Kernel functions are introduced in (10)-(13). The RBF Kernel is used in this paper.

Dot Product Kernel:

$$K(x_i, x) = \left(x_i^T, x\right) \tag{10}$$

Polynomial Kernel:

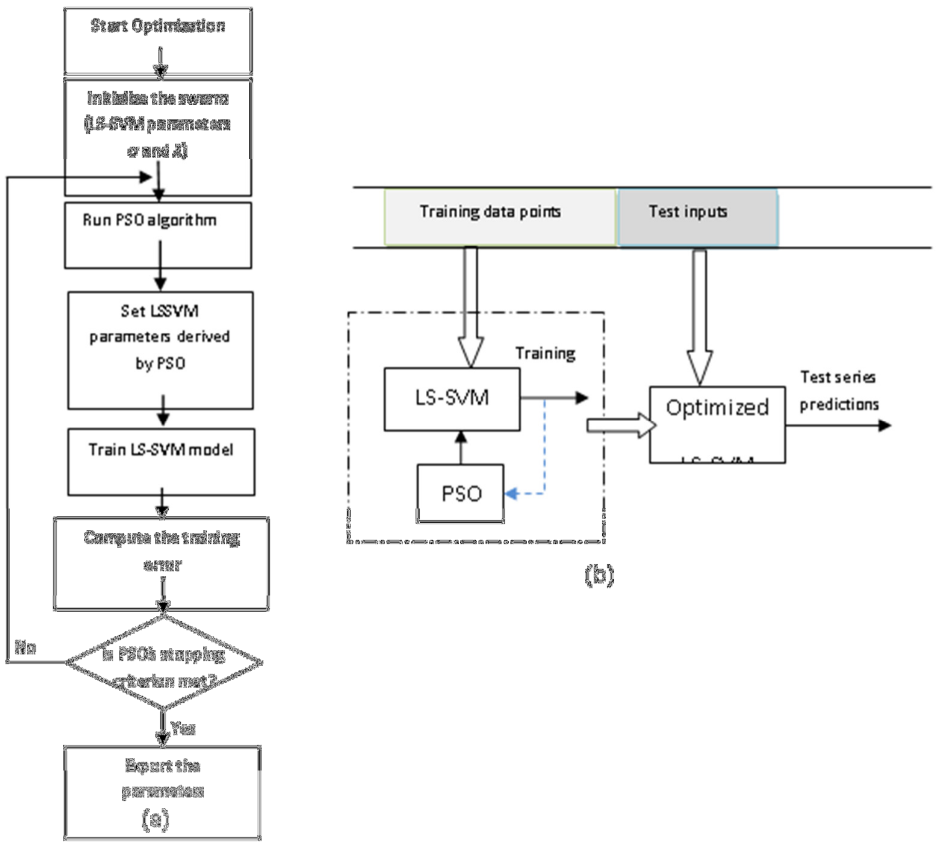
$$K(x_i, x) = \left[\left(x_i^T, x\right) + 1\right]^d \tag{11}$$

MLP Kernel:

$$K(x_i, x) = \tanh\left[\left(x_i^T, x\right) + b\right] \tag{12}$$

RBF Kernel:

$$K\left(x_i, x\right) = \exp\left(\frac{\left\|x_i - x\right\|^2}{2\sigma^2}\right) \tag{13}$$



Based on the presented description, γ and σ are the only parameters of the LS-SVM which should be optimally tuned. The PSO algorithm is presented in next section and will be further used for optimal selection of LS-SVM's parameters.

3. PROPOSED FORECAST FRAMEWORK

The LS-SVM model, described in previous section, contains two adjustable parameters which have a key role in the accuracy of predictions of the produced of the model. Various optimization techniques, such as genetic algorithms (GA), simulated annealing (SA) and particle swarm optimization (PSO) can be utilized for fine tuning of the LS-SVM's parameters. In this paper, PSO algorithms, due to the speed of convergence, simplicity of implementation and less susceptibility of being trapped in local optima, are preferred [20].

In PSO, particles flow in a multi-dimensional search space and the position of each particle is tuned based on the experiences gained by him and his neighbours. In this paper we adopt a gbest PSO algorithm. In gbest algorithm the new position of the particle is found by adding the velocity component, as following:

$$x_i(t+1)=x_i(t)+v_{ij}(t+1)$$
$$v_{ij}(t+1)=v_{ij}(t)+c_1r_{1j}(t)\left[y_{ij}(t)-x_{ij}(t)\right]+c_2r_{2j}(t)\left[\hat{y}(t)-x_{ij}(t)\right] \tag{14}$$

Table 1. Training and testing data

Case Study	Input features	Set	Period	Length
Iran	Previous gas consumptions and population	Training	1967-1997	31
		Test	1998-2006	9
U.S.	Previous gas consumptions, population and GDP per capita	Training	1985-2004	20
		Test	2005-2009	5

where, $x_i(t)$ is the position of particle i at time t, $v_{ij}(t)$ is velocity of particle i at dimension j at time t, $y_i(t)$ is the best position found by particle i, $\hat{y}(t)$ is the best position found by swarm, c_1 and c_2 are acceleration constants and $r_1(t)$ and $r_2(t)$ are uniformly distributed number in [0, 1].

For optimal selection of the LS-SVM's parameters, i.e. σ and λ, two dimensional particles are randomly distributed in the search space. The overall procedure of the LS-SVM optimization by the PSO algorithm is illustrated in Fig. 2(a).

The framework of the proposed forecast approach is shown in Fig 2(b). This figure illustrates how the LS-SVM model is optimized by PSO algorithm through training data and the optimized model is employed for prediction of the test data. In next section, the proposed PSO optimized LS-SVM model will be applied to prediction of gas consumption in Iran and U.S.

4. Gas Consumption Prediction

In this section the annual gas consumption of Iran and the United States will be forecasted using the proposed PSO-LSSVM model. The consumption in previous period as well as the population to the last point are the standard input variables for the prediction. The training and test data set for each case are presented in Table 1. These data are collected from Institute for International Energy Studies (IIES) webpage, World Bank Development Indicator datasets and the U.S. energy information administration website [20-22]. For evaluation of the performance of the PSO-LSSVM model, the following error measures are computed,

Mean absolute percentage error (MAPE):

$$MAPE=\frac{1}{T}\sum_{t=1}^{T}\frac{\left|y_t-\hat{y}_t\right|}{y_t}\times 100 \tag{15}$$

Absolute percentage error (APE):

$$APE=\frac{\left|y_t-\hat{y}_t\right|}{y_t}\times 100 \tag{15}$$

Normalized mean squares error (NMSE):

$$NMSE=\frac{\sum_{t=1}^{T}\left(y_t-\hat{y}_t\right)^2}{\sum_{t=1}^{T}\left(y_t-\bar{y}\right)^2} \tag{16}$$

$$\bar{y} = \frac{1}{T}\sum_{t=1}^{T} y_t$$

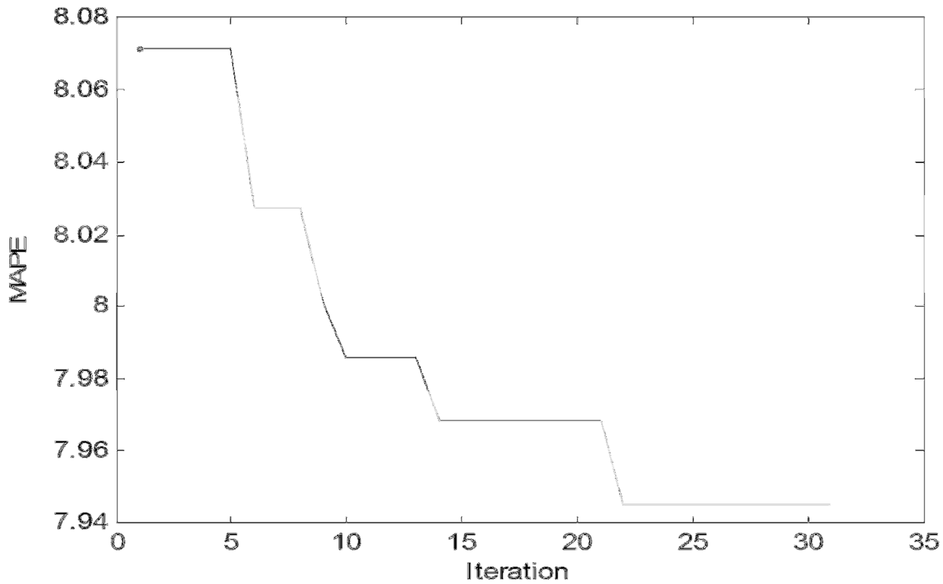


Figure. 3 Training MAPE versus PSO iterations

where, y_t and $\hat{y}_t$ are the actual and predicted consumptions at period *t*, respectively.

4.1. Prediction of Gas Consumption in Iran

In this case study, the annual consumption of natural gas in Iran form 1998 to 2006 will be forecasted. For this purpose the following input-output sets will be used for training the proposed model.

Input vector	Output	
$\{y(t-1), x(t-4), x(t-3), x(t-2), x(t-1)\}$	$y(t)$	(17)

where, $y(t)$ and $x(t)$ are gas consumption and population at time *t*, respectively. Furthermore, for the purpose of comparison, a multi-layer perceptron (MLP) network was trained and optimized using training data.

First, the optimization of the LS-SVM model was carried out by PSO algorithm for 30 iterations. The number of particles, dimension of each particle, c_1,c_2 for the PSO algorithm are set as 30, 2, 2, 2 respectively. The MAPE for the training data was selected as the fitness function in PSO. The fitness value for PSO iterations is shown in Fig. 3. The actual and forecasted values of Iranian gas consumption for training and test data are illustrated in Fig. 4, revealing the remarkable performance of the proposed model in estimating oil consumption series. The actual and forecasted gas consumption for test period is shown in Table 2. A comparison between performance of the proposed method and the MLP network is presented in Table 3. According to this table the APE ranges from 0.28% to 27.5%. The maximum APE% was occurred in 2000 when an abrupt change happens in the gas consumption series, as shown in Fig. 4, and the PSO-

LSSVM failed to capture this change in consumption. The optimized MLP network has one hidden layer with 4 neurons. The surpassing performance of the proposed method is evident.

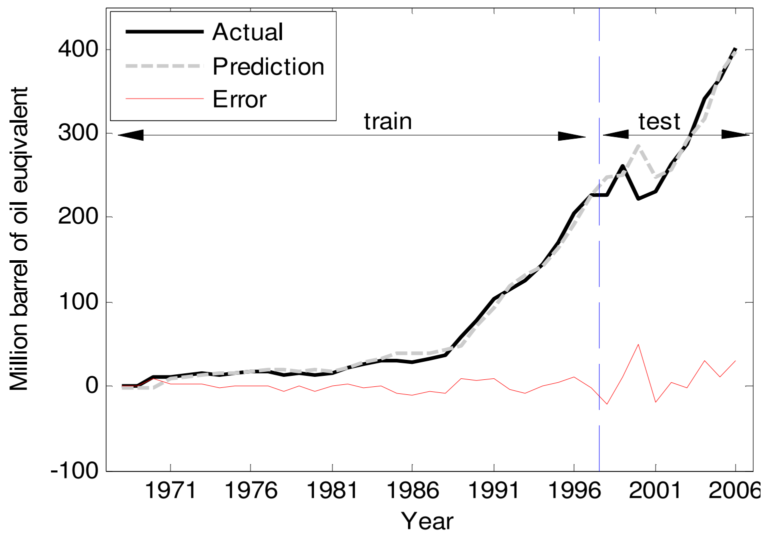


Fig. 4 Actual and predicted values for train and test data for Iran Gas consumption

4.2. Prediction of Gas Consumption in U.S.

Due to unavailability of the data for U.S. gas consumption prior to 1980, training and test data different to the previous case study are employed here. As presented in Table 1, there are 20 training data points, while the test data contain 5 samples. For single step ahead prediction of the U.S. gas consumption, following input features are considered,

Input vector Output (18)

Table 2. Actual and forecasted gas consumption for case study 1

Year	Actual	Forecast	APE%
1998	227.3	249.5	9.77
1999	262.3	251.7	4.04
2000	223.4	285	27.57
2001	230.8	250.1	8.36
2002	263.3	258.1	1.97
2003	288.1	288.9	0.28
2004	341.9	311.3	8.95
2005	364.9	354.1	2.96
2006	401.9	370.7	7.76

Table 3. Compsrison between the PSO-LLNF and MLP models for case study 1

Method	Data	MAPE%	NMSE
MLP	Training	7.33	0.26
	Test	8.45	0.33
PSO-LSSVM	Training	6.12	0.15
	Test	7.96	0.19

$$\{y(t-1), x(t-1), w(t-1)\} \qquad\qquad y(t)$$

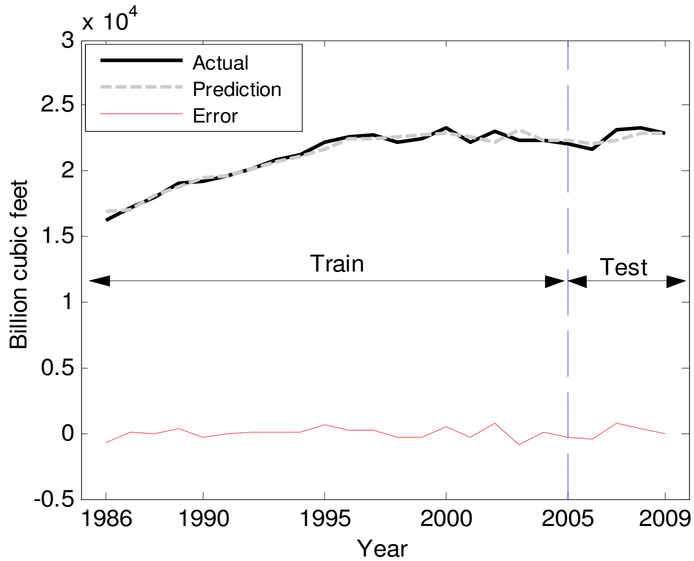


Fig. 5 Actual and predicted values for train and test data for U.S. Gas consumption

Table 4. Actual and forecasted gas consumption for case study 2

Year	Actual	Forecast	APE%
2005	22011	22364.2	1.60
2006	21685	22109.7	1.96
2007	23097	22305.4	3.43
2008	23227	22857.9	1.59
2009	22816	22819.2	0.01

Table 5. Compsrison between the PSO-LLNF and MLP models for case study 2

Method	Data	MAPE%	NMSE
MLP	Training	1.83	0.64
	Test	2.2	0.95
PSO-LSSVM	Training	1.08	0.26
	Test	1.72	0.57

where, $y(t)$, $x(t)$ and $w(t)$ are gas consumption and population and GDP per capita at time t, respectively.

Again, the LS-SVM model was optimized using PSO algorithm and MAPE as the fitness function. Similar to the previous case, an MLP network was also optimized with 3 neurons for making a comparison to the results obtained by the proposed PSO-LSSVM. The predictions of the PSO-LSSVM as well as the actual gas consumptions and the forecast error are depicted in Fig. 5. The remarkable forecast performance and accuracy of the proposed method is obvious in this figure. The actual and forecasted values of the test data of the U.S. gas consumption are given in Table 4. The minimum and maximum values of APE% are 3.43% and 0.01%, respectively. Besides, the performance of the PSO-LSSVM and the MLP network in terms of error indices MAPE and NMSE are presented in Table 5. The results in this table demonstrate the noteworthy performance of the proposed method as well as its superiority over the optimized MLP network.

Table 6. Comparison of Min. and Max. APE% for both case studies

Year	Min. APE%		Max. APE%	
	PSO-LSSVM	MLP	PSO-LSSVM	MLP
Case study 1	0.28	1.96	27.57	33.12
Case study 2	0.01	0.89	3.43	4.57

Table 7. Improvement in MAPE and NMSE

Year	MAPE improvement	NMSE improvement
Case study 1	5.8%	73.7%
Case study 2	21.8%	40%

$$\text{Improvement} = \frac{\text{Error index}_{MLP} - \text{Error index}_{PSO\text{-}LSSVM}}{\text{Error index}_{MLP}} \times 100$$

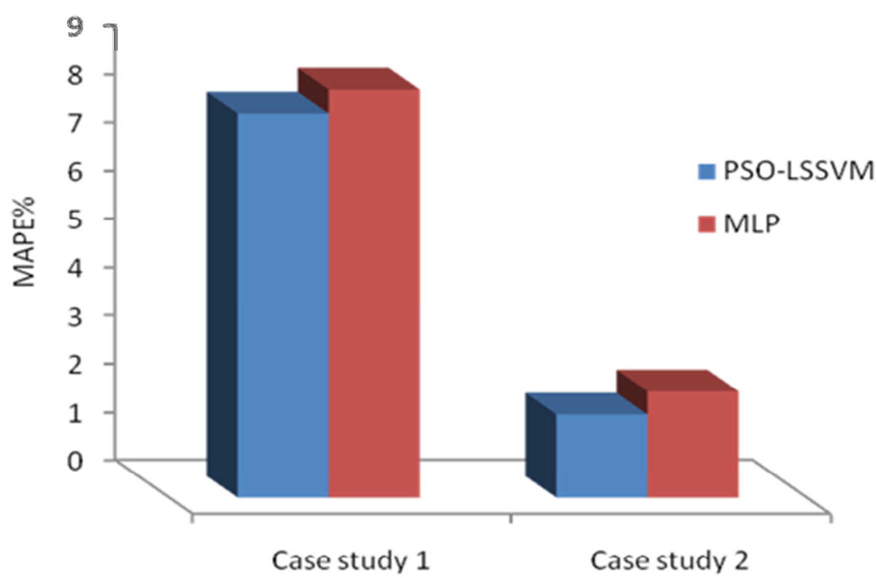


Fig. 6 overall comparison of the forecast models for both case studies

4.3. Comparison of the Results

More detailed discussion on the prediction results is presented in this sub-section. For this purpose, the maximum and minimum values of APE, achieved through the proposed method and the MLP model, for both case studies is are shown in Table 6. For both case studies, there is a considerable difference between the minimum and maximum APEs associated with the PSO-LSSVM method and the MLP model. For instance, the minimum APE of the PSO-LSSVM in the first case study is 0.28%, while this value for the MLP model is 1.96% (clearly 7 times that of PSO-LSSVM).

To thoroughly analyze the superiority of the proposed method over MLP, the improvement in error indices, i.e. MAPE and NMSE, is computed and summarized in Table 7. Clearly, significant improvement has been achieved in both case studies by employing the proposed PSO-LSSVM mode.

Furthermore, an overall comparison between the proposed approach and the MLP network for both case studies is provided by Fig. 6. Obviously, the proposed PSO-LSSVM model has outperformed MLP network in both case studies. As another finding from Fig. 6, both forecast models had better accuracy for the second case study. The reason for this can be understood by comparing actual gas consumption series of Iran and the U.S. in Figs. 4 and 5, respectively. As shown in these figures. the Iranian gas consumption series exhibits more changes and fluctuations through the time. Hence it is less predictable in comparison the U.S. gas consumption series, which is a more smooth series.

5. CONCLUSION

Accurate forecasting of natural gas consumption, due to its large contribution in providing the world energy demand, needs special attention. This paper proposed a PSO optimized LS-SVM approach for prediction of natural gas consumption in Iran and United States. Support vector machines show noticeable forecast and estimation capabilities owing to mapping nonlinear data into high dimensional feature space and then performing linear regression. Optimization of parameters of the LS-SVM model by a simple bust fast and efficient PSO algorithm resulted in a hybrid model, applied for gas consumption prediction. Finally, two different case studies were considered for evaluating the performance of the proposed PSO-LSSVM approach. Forecasting annual gas consumption in Iran, as one of the world largest gas producers, and the U.S., as one the world largest gas consumers, revealed the promising forecast ability of the method. Assessing performance of the proposed approach in terms of NMSE and MAPE and comparison to an optimized MLP network showed the superior performance of the PSO-LSSVM model.

APPENDIX: RAW DATA FOR PRESENTED CASE STUDIES

Table A1: Raw data for Iran

Year	Gas consumption (million barrel of equivalent)	Population (million people)
1967	0.7	26.07
1968	0.8	26.82
1969	1	27.6
1970	10.2	28.43
1971	12	29.35
1972	13.1	30.27
1973	15.4	31.2
1974	14.2	32.17
1975	15	33.21
1976	16.7	34.28
1977	17.1	35.39
1978	13.4	36.55
1979	15.3	37.79
1980	12.9	39.12
1981	15.9	40.54
1982	22	42.02
1983	25.2	43.6
1984	31.2	45.28
1985	30.3	47.1
1986	28.7	48.82
1987	32.9	50.42
1988	36.3	51.9
1989	57.6	53.23
1990	78.9	54.4
1991	103.3	55.28
1992	115.4	56.18
1993	125.4	57.09
1994	145.1	58.01
1995	171.2	58.95
1996	204.3	59.88
1997	226.1	60.8
1998	227.3	61.85
1999	262.3	62.9
2000	223.4	63.94
2001	230.8	64.98
2002	263.3	66.01
2003	288.1	67.04
2004	341.9	68.07
2005	364.9	69.09
2006	401.9	70.1

Table A2: Raw data for the U.S.

Year	Gas consumption (billion cubic feet)	Population (million people)	GDP per capita
1985	17281	17588.81	3.21
1986	16221	18427.29	2.43
1987	17210.81	19394.19	1.95
1986	18030	20698.24	1.84
1987	19119	22038.82	1.82
1988	19174	23053.97	1.94
1990	19562	23492.67	1.83
1991	20228	24526.93	1.85
1992	20789.51	25447.54	2.03
1993	21247	26719.14	1.87
1994	22207	27637.66	1.49
1995	22609	28894.11	1.97
1996	22737	30363.79	2.17
1997	22246	31687.05	1.97
1998	22405	33332.14	2.24
2000	23333	35080.73	3.95
2001	22239	35898.09	4.43
2002	23007	36796.57	3.15
200	22277	38195.68	5.17
2004	22389	40308.69	5.81
2005	22011	42534.48	8.12
2006	21685	44663.47	6.88
2007	23097	46627.1	6.87
2008	23227	47208.54	8.7
2009	22816	45989.18	4.19

REFERENCES

[1] International Energy Agency (IEA) [online] Available: www.iea.org

[2] Gutiérrez, A., Nafidi, A., Gutiérrez Sánchez, R., 2005. Forecasting total natural-gas consumption in Spain by using the stochastic Gompertz innovation diffusion model. Applied Energy; 80(2), 115–124.

[3] Liu, L.M, Lin, M.W., 1991. Forecasting residential consumption of natural gas using monthly and quarterly time series, International Journal of Forecasting; 7(1), 3-16.

[4] Siemek, J., Nagy, S., Rychlicki, S., 2003. Estimation of natural-gas consumption in Poland based on the logistic-curve interpretation. Applied Energy; 75(1-2), 1-7.

[5] Akkurt, M., Demirel, O.F., Zaim, F., 2010. Forecasting Turkey's natural gas consumption by using time series methods. European Journal of Economic and Political Studies; 3(2).

[6] Li, J., Dong, X., Shangguan, J., Höök, M., 2001. Forecasting the growth of Chinese natural gas consumption. Energy; 36(3), 1380-1385.

[7] Sánchez-Úbeda, E.F., Berzosa, A., 2007. Modeling and forecasting industrial end-use natural gas consumption. Energy Economics; 29(4), 710-742.

[8] Jang, J.S., Sun, C.T., Mizutani, E., 1997. Neuro-Fuzzy and Soft Computing: A Computational Approach to Learning and Machine Intelligence. Prentice-Hall International, Inc.

[9] Hippert, H.S., Pedreira, C.E., Souza, R.C. 2001. Neural networks for short-term load forecasting: a review and evaluation. IEEE Transactions on Power Systems; 16(1), 44–55.

[10] Lauret, P., Fock, E., Randrianarivony, R.N., Manicom-Ramasamy, J.F., 2008. Bayesian neural network approach to short time load forecasting. Energy Conversion and Management;49(5), 1156–66.

[11] Khotanzad, A., Elragal H., Lu, T,L., Dallas, T.X., 2000. Combination of artificial neural-network forecasters for prediction of natural gas consumption. IEEE Transactions on Neural Networks; 11(2), 464 – 473.

[12] Cristianini, N., Shawe-Taylor, J. An Introduction to Support Vector Machines and Other Kernel-Based Learning Methods. Cambridge, U.K.: Cambridge Univ. Press, 2000.

[13] Suykens, J.A.K., Van Gestel, T., De Brabanter, J., De Moor, B., Vandewalle, J. Least Squares Support Vector Machines. Singapore: World Scientific, 2002.

[14] Vapnik, V. N. Statistical Learning Theory. New York: Wiley, 1998.

[15] Sapankevych, N., Snakar, R. 2009. Time series prediction using support vector machines. IEEE computational intelligence magazine, 24-28.

[16] Lee, M.C., To, C., 2010. Comparison of Support Vector Machine and Back Propagation Neural Network in Evaluating the Enterprise Financial Distress. IJAIA; 1(3), 31-43.

[17] Kim, S.K., Park, Y.J., Toh, K.A., Lee, S., 2010. SVM-based feature extraction for face recognition. Pattern Recognition; 43(8), 2871-2881.

[18] Mehta, S.S., Lingayat, N.S., 2008. Development of SVM based classification techniques for the delineation of wave components in 12-lead electrocardiogram. Biomedical Signal Processing and Control; 3(4), Pages 341-349.

[19] Xu, S.Q., Luo, Q.Y., Xu, G.H., Z., Lei. 2009 .Asymmetrical interval regression using extended ε-SVM with robust algorithm. Fuzzy Sets and Systems; 160(7), 988-1002.

[20] Engelbert, A. P. Computational Intelligence- An Introduction, Wiley, 2007.

[21] Institute for International Energy Studies. [online] Available: http://iies.ac.ir

[22] World Bank, 2006. World Development Indicators 2006. World Bank, Washington, DC.

[23] U.S. Energy Information Administration. [online] Available: http://www.eia.gov/.