Forecasting of Turkey’s Electricity Consumption with Support Vector Regression and Chaotic Particle Swarm Algorithm*

Oğuz KAYNAR** & Halil ÖZEKİÇİOĞLU*** & Ferhan DEMİRKO PARAN****

Abstract:
Energy is a very important factor in terms of sustaining the economic development for developing and industrialized countries. Electricity is one of the most important forms of energy for industrialization and improvement of living standards. The estimation and modeling of electricity consumption has a special importance in Turkey which is a foreign-dependent country in energy. In this study, a forecasting application is made by using Turkey’s electricity consumption, population, import, export and gross domestic product between 1975-2014 employing support vector regression methods. Chaotic particle swarm optimization algorithm (CPSO) is used to choose the parameters of SVR.

Keywords: Electricity consumption, Support Vector Regression, Chaotic Particle Swarm Optimization Algorithm, Prediction

* This paper presented as an extended abstract “2nd International Annual Meeting of Sosyoekonomi Society” where held in October 28-29, 2016 in Amsterdam NL.

** Assoc.Prof., Department of MIS, Cumhuriyet University, Turkey, okaynar@cumhuriyet.edu.tr

*** Assoc.Prof., Department of Economics, Cumhuriyet University, Turkey, bozekicioglu@cumhuriyet.edu.tr

**** Res. Assist., Department of MIS, Cumhuriyet University, Turkey, fdemirkoparan@cumhuriyet.edu.tr
1. INTRODUCTION

Energy is an important factor in terms of sustaining economic development for developing and industrialized countries. Worldwide energy consumption is rapidly increasing for reasons such as population growth, the importance given to large scale industrialization and maintenance of positive economic growth rate\(^1\). Rising energy prices, global warming and climate change, increase in worldwide energy demand, dependence of fossil fuel depleted rapidly, the lack of new energy technologies meet the increasing energy demand commercially cause countries to be concerned about supply security\(^2\). Accurate predictions of energy consumption affect not only capital investments, environmental quality, income analysis and research management but also maintain supply security and provide the energy policies to be implemented effectively\(^3\).

Energy plays an active role in the demand and supply in the economy. In terms of demand, energy is a product that consumers buy to maximize their benefits. In terms of supply, energy is a production factor like labor and capital. Also, because of being an essential input in the most production and consumption activities, energy has a decisive role in the realization of economic growth and development\(^4\).

Electrical energy is also one of the most important form of energy for industrialization and improvement of living standards. Dependence on electrical energy worldwide is increasing in parallel with energy. According to IEA (International Energy Agency), the proportion of electricity in total energy demand of the world will increase in medium term and electric will be the fastest growing form of energy for end-users\(^5\). Also in the light of developments in information and communication technologies electricity is seen as the main source of energy for the countries towards becoming a digital society and plays a vital role in scientific developments. Electricity consumption in Turkey is also increasing rapidly in parallel with the energy consumption. Electricity consumption has a chaotic and nonlinear trend in Turkey which is a foreign dependent country due to an unstable economy having an extremely sensitive nature against domestic and foreign developments in politics, economics and market\(^6\). Therefore, modeling and estimation of electricity consumption in Turkey has a special importance.

There are a lot of studies that forecast electricity consumption in literature. Gürbüz et al. forecasted electricity consumption of Turkey with three different scenarios

---

by using traditional and meta-heuristic artificial neural networks with economic indicators like population, import, export and GNP. Akay and Atak proposed grey prediction method with rolling mechanism to forecast total and industrial electricity consumption of Turkey. Hamzaçebi forecasted electricity consumption of Turkey between 2005-2020 by using ANN and compared the results with MAED method used by Turkey Ministry of Energy and Natural Resources. Sözen and Arcaklioğlu created three different ANN models by using GNP, GDP and population indicators to forecast net energy consumption of Turkey. Toksarı presented an ant colony algorithm benefited from population, GDP, import and export data to forecast energy demand of Turkey. Kavaklioğlu et al. modelled electricity consumption of Turkey as a function of economic indicators such as population, GDP, import and export and made predictions up to 2027 by using the data between 1975-2000. Küçükali and Barış used fuzzy logic method to make predictions of annual gross electricity demand of Turkey.

Song et al. proposed tent-map-based chaotic PSO and expressed that it has higher iterative speed than logistic map. Alataş et al. presented 12 different PSO that employed chaotic map used sequences generated from different chaotic systems.

In the studies not only past values of electricity consumption but also economic and non-economic indicators such as population, GDP, GNP, import, export, price, added-value, number of customers, CO₂ emission, installed capacity, climate, temperature and relative humidity are used.

Recently, among prediction models Support Vector Regression (SVR) has attracted researchers’ attention. Hu et al. proposed firefly based memetic algorithm (FA_MA) to determine the parameters of SVR in their study which electricity load prediction is made. The proposed method made more accurate predictions than not only 4 well knowned evolutionary algorithm and 3 prediction model but also outperformed hybrid algorithms in literature. Fan et al. presented a novel method to forecast electricity load. First, they clustered input data unsupervisedly and then

---

they used SVR groups to fit training data of each subset supervisedly. Hong utilised artificial immune algorithm to adjust SVR parameters for making predictions of regional electricity load in Taiwan. Hong used chaotic particle swarm algorithm to set the parameters of SVR and expressed that used method outperformed GA and simulated annealing algorithm. Hong used chaotic ant colony algorithm to determine SVR parameters in another study. Li et al. made short-term load forecasting by using PSO integrated with chaotic search process to adjust the parameters of SVM. Wu proposed a new v-support vector machine method used Gaussian loss function trained with chaotic PSO. Kavaklioğlu used ε-SVR method to forecast electricity consumption of Turkey. Population, GNP, import and export variables were used to model consumption function. In another study which electricity consumption of Turkey was forecasted, ANN and SVR methods were used together. Kavaklioğlu used multivariable regression method to forecast annual electricity consumption of Turkey in another study. He utilised Singular Value Decomposition (SVD) method to reduce the size of the problem and increase the robustness of predictions.

Accurate predictions of energy are important and required for not only planning expansion of capacity but also monitoring the environmental problems, setting the taxes and demand management planning. Low estimates led to cuts that paralyzes economy and daily life, high estimates cause unnecessary and wasted capacity which mean waste of financial resources. In addition, all possible variables effect the output data should be included in the model to make accurate predictions. When correlation and casualty between electricity and some economic indicators are con-

---

considered, it can be seen a lot of studies exist in literature (27, 28, 29, 30, 31, 32, 33). Main indicators are GDP and economic growth.

The studies examine the causality between electricity consumption and GDP vary according to their results. Variety of obtained results is derived from used method, data, examined country being developed or developing, different countries having different characteristics such as having local energy resources political order, culture, energy politics, usage rate of electricity (34, 35). However, there is a strong relationship between electricity consumption and GDP and accordingly economic growth. For electric energy is basic input for production, while GDP increases electricity demand also increases because production activities in industrial sectors like construction, manufacturing and transportation requires sufficient level of electricity resource. So restricted infrastructure of electricity is a factor that could prevent the economic growth. Electricity consumption is closely related to national wealth and also an indicator of socio-economic development.

Turkey is one of the biggest countries in Europe. Annual population growth rate have the biggest value in IEA countries which is 1.6% 38. It is expected energy demand will increase annually and increasingly continue in parallel with economic growth and rapid population growth. Due to the lack of energy resources, Turkey is dependent on imported fossil fuel for electricity generation. 60% of energy need in Turkey is met by import and share of energy in import is increasing every year.

---

It is seen that electricity consumption is closely related to economic indicators like GDP, population, import and export. Considering this relationship, in this study, a forecasting application is made by using electricity consumption, population, GDP, import and export data of Turkey between 1975-2014.

2. METHODOLOGY

2.1. Support Vector Regression (SVR)

Support vector regression proposed by Vapnik⁴⁰ is a powerful machine learning technique based on developments in statistical learning theory and used for classification and regression. SVR, which is built on the principle of risk minimization that estimates the upper limit of generalization error by minimizing it, is resistant to overfitting problem and it is shown that it achieves a high generalization performance eventually while solving forecasting problems of various time series⁴¹. This method eliminates some disadvantages like trapping local minimum of ANN that minimize errors by empirical risk minimization.

The principle of SVR is mapping data to a feature space having higher dimension. Learning ability of SVR is independent from dimension of feature space, so it performs well⁴².

For a given dataset \( \{ (\mathbf{x}_i, d_i) \}_{i=1}^N \), let \( \mathbf{x}_i \) be input vector, \( d_i \) actual value and \( N \) the number of data. So SVR function is:

\[
y = f(x) = w \psi(x) + b
\]  

(1)

where \( w \psi(x) \) is the feature mapped from non-linear input space \( x \). In this method non-linear kernel functions \( (\psi(x)) \) are let to be used from input space to feature space. \( w \) and \( b \) are predicted coefficients by minimizing regularized risk minimization. SVR method need model have a good generalization performance and \( w \) be smooth as far as possible. So norm of \( w \) vector \( (\|w\|) \) must be minimized for each data point.

\[
R(C) = \frac{1}{(C / N)} \sum_{i=1}^{N} L_e(d_i, y_i) + \frac{\|w\|^2}{2}
\]  

(2)

\[ L_\varepsilon(d, y) = \begin{cases} 
0, & |d - y| \leq \varepsilon \\
|d - y| - \varepsilon, & \text{otherwise} 
\end{cases} \] (3)

where \( C \) and \( \varepsilon \) are user specified prescribed parameters. \( L_\varepsilon(d, y) \) is \( \varepsilon \)-insensitive loss function. If the forecasted value is within the \( \varepsilon \)-tube, the loss equals 0. The second term in equation 3, \( \frac{1}{2} w^T w \), measures the flatness of the function. \( C \) is used to determine the trade-off between empirical risk and flatness of the model. \( C \) and \( \varepsilon \) are user specified parameters. \( \xi^i \) and \( \xi^i^* \) are slack variables that figure the distance between actual value and corresponding boundary values of \( \varepsilon \)-tube. So equation 2 is transformed into43:

Minimize;

\[ R(w, \xi, \xi^*) = \frac{1}{2} w^T w + C \left( \sum_{i=1}^{N} (\xi_i + \xi_i^*) \right) \] (4)

with the constraints;

\[
\begin{align*}
\sum_i \alpha_i \psi(x_i) + b_i - d_i &\leq \varepsilon + \xi_i^* \\
d_i - \sum_i \alpha_i \psi(x_i) &\leq \varepsilon + \xi_i^* \\
\xi_i, \xi_i^* &\geq 0
\end{align*}
\]

where \( C \) parameter provide the balance between the flatness of \( w \) vector and penalty of errors higher than \( \varepsilon \). So it turns into a optimization problem that forecasts \( w \) and \( b \) parameters minimizing the cost. Constrained optimization problem is solved by Lagrangian form:

\[
L(w, b, \xi, \xi^*, \alpha, \alpha^*, \beta, \beta^*) = \frac{1}{2} w^T w + C \left( \sum_{i=1}^{N} (\xi_i + \xi_i^*) \right) - \sum_i \beta_i \left[ \sum_{i=1}^{N} \psi(x_i) + b_i - d_i + \varepsilon + \xi_i^* \right] \\
- \sum_{i=1}^{N} \beta_i \left[ \sum_{i=1}^{N} \psi(x_i) - b_i + \varepsilon + \xi_i^* \right] - \sum_{i=1}^{N} (\alpha_i \xi_i + \alpha_i^* \xi_i^*)
\] (5)

\[ ^{43} \text{Hong, W. C. (2009a). Electric load forecasting by support vector model. Applied Mathematical Modelling, 33(5), 2444-2454.} \]
This equation is minimized according to basic variables $w$, $b$, $\zeta$ and $\zeta^*$ and maximized according to non-negative Lagrange multipliers $\alpha_i$, $\alpha_i^*$, $\beta_i$ and $\beta_i^*$. To find minimum, function differentiate according to $w$, $b$, $\zeta$ and $\zeta^*$ separately and equalize 0. Karush-Kuhn-Tucker conditions are applied to regression and dual Lagrangian is obtained while Kernel function is $K(x_i, x_j) = \psi(x_i)\psi(x_j)$:

\[
\mathcal{L}(\beta_i, \beta_i^*) = \sum_{i=1}^{N} d_i (\beta_i - \beta_i^*) - \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} (\beta_i - \beta_i^*) (\beta_j - \beta_j^*) K(x_i, x_j)
\]

with constraints:

\[
\sum_{i=1}^{N} (\beta_i - \beta_i^*) = 0
\]
\[0 \leq \beta_i \leq C, \quad i=1, 2, \ldots, N\]
\[0 \leq \beta_i^* \leq C\]

In equation 6 Lagrange multipliers proves $\beta_i \ast \beta_i^* = 0$. Lagrange multipliers $\beta_i$ and $\beta_i^*$ are computed and optimum weight vector of regression hyperplane is:

\[
w^* = \sum_{i=1}^{N} (\beta_i - \beta_i^*) \psi(x)
\]

So regression function would be:

\[
f(x, \beta, \beta^*) = \sum_{i=1}^{N} (\beta_i - \beta_i^*) K(x, x_i) + b
\]

$K(x, x_i)$ is kernel function. Kernel value equals inner product of $x$ and $x_i$ vectors in $\psi(x)$ and $\psi(x_i)$ feature space. The most common kernel functions are below.

Polinomial kernel function:

\[
K(x, x_i) = (x^T x_i + a_1)^d
\]

where $d$ is order, $a_1$ and $a_2$ are coefficients.

Multilayer perceptron kernel function:

\[ K(x_i, x) = \tanh(x_i^T x - b) \]  

(10)

where \( b \) is a constant.

Gaussian RBF kernel function:

\[ K(x_i, x) = \exp(-\|x_i - x\|^2 / 2\sigma^2) \]  

(11)

In SVR method, most important point is determining optimum user-specified parameters which are error (\( \varepsilon \)), constant (\( C \)) and width of Gaussian function \( \sigma \). The selection of these three parameters directly affect the performance of SVR model\(^44\). If \( C \) is very high (approximates infinite), empirical risk must be minimized. High value of \( \varepsilon \) causes regression forecasting function be flat. \( \sigma \) adjusts the width of Gaussian function so it presents range of \( x \) values in training set. Hence, all three parameters effect the structure of model in a different way. Many methods have proposed to determine \( C \) and \( \varepsilon \) parameters. Some of them are choosing them according to user experience, cross validation, asymptotic optimization and evolutionary algorithms. In this study chaotic particle swarm optimization algorithm is used to determine the parameters of SVR.

2.2. Chaotic Particle Swarm Optimization (CPSO)

Particle Swarm Optimization algorithm, first introduced by Eberhart and Kennedy\(^45\) is inspired by fish and bird’s behaviour of foraging. It is initialized by random solutions which are particles in a swarm. Each particle in the population represents a potential solution. Each particle have velocity, position and best position that are defined as 

\[ X_{(k)} = [x_{(k),1}, x_{(k),2}, \ldots, x_{(k),n}] \]

\[ V_{(k)} = [v_{(k),1}, v_{(k),2}, \ldots, v_{(k),n}] \]

and

\[ P_{(k)} = [p_{(k),1}, p_{(k),2}, \ldots, p_{(k),n}] \]

respectively where \( k=C, \varepsilon, \sigma \) and \( i=1, 2, \ldots, N \).

The global best position in whole population called global best is represented as 

\[ P_{(g)} = [p_{(g),1}, p_{(g),2}, \ldots, p_{(g),n}] \]

where \( k=C, \varepsilon, \sigma \) and \( g=1, 2, \ldots, N \). The position and velocity of each particle are updated iteratively in the search process by the equations below:

\[ V_{i}[t + 1] = \alpha V_{i}[t] + c_1 \text{rand}(.) (P_{i} - X_{i}) + c_2 \text{Rand}(.) (P_{g} - X_{i}) \]  

(12)

\[ X_{i}[t + 1] = X_{i}[t] + V_{i}[t + 1] \]  

(13)


where $w$ is inertia weight which determines how much previous velocity impacts the current velocity. $c_1$ and $c_2$ are positive constants called acceleration coefficients. $\text{Rand}(.)$ and $\text{rand}(.)$ are independent random variables which distributes uniformly in the range of $[0,1]$. To prevent the particles go out of the search space every components in $V(k)i$ are limited in the range of $[-v_{max}, v_{max}]$. This iterative search process continues until stopping criteria is satisfied.

Differently from other metaheuristic algorithms, PSO has a memory that remembers all the good solutions of particles. Moreover, particles cooperate and share information. PSO is a simple, efficient and easy to implement. It can solve many different problems in different areas, find global optimum region, at least a good local optimum. It is inexpensive in terms of CPU and memory. However it convergences very fast and is not able to improve the quality of solutions while iteration number increases$^{46}$. The performance of PSO largely depends on its parameters and it can easily trap in local optimum. To eliminate these drawbacks chaos and chaos-based searching algorithms have draw attention recently due to being easy to implement and to be able to avoid trapping local optimum. Chaos have three characteristics which are randomness, ergodicity and regularity and it is highly sensitive to initial conditions. One of the critical factors effects the performance of chaotic optimization algorithms is chaotic mapping function. Most common chaotic mapping function used for generating chaotic sequences is logistic chaotic function$^{47}$:

$$x_{n+1} = \mu x_n (1-x_n) \tag{14}$$

where $x_n$ is the iteration value of $x$ in time $n$ and $\mu$ is control parameter. When $\mu=4$, system is completely chaotic and $x_0$ can take any value in the range of $(0, 1)$ except $\{0.25, 0.5, 0.75\}$.

Liu et al. proposed a hybrid chaos and revised PSO algorithm. They used adaptive inertia weight to provide the trade-off between global exploration and local exploitation. They combined PSO with chaotic search and enhanced the searching efficiency and improved the quality of search$^{48}$. In this study, CPSO algorithm is used to determine three parameters of SVR due to the ability to escape trapping local optimum.

3. EMPIRICAL RESULTS
1975-2014

In this study electricity consumption of Turkey between the years of 1974-2014 is modelled by using GDP, import, export and population variables. A prediction application is made by employing this model and SVR method. 60% of data is used as training data, 40% of data is used as test data.

---


The optimum values of user specified parameters of SVR method which are C (penalty), \( \sigma \) (width of Gaussian RBF function) and \( \varepsilon \) (error) are determined by chaotic particle swarm optimization. Firstly, data is normalized to eliminate the effects of different scaling. Due to confine the solution space, the parameters of C, \( \sigma \) and \( \varepsilon \) are limited between 1-1000, 0-1 and 0-0.1 respectively. The parameter values of the model are given in Table 1.

**Table 1: Parameter values of model determined by CPSO**

<table>
<thead>
<tr>
<th></th>
<th>C (Penalty)</th>
<th>( \sigma ) (width of Gaussian RBF function)</th>
<th>( \varepsilon ) (error)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Value</td>
<td>90.9094</td>
<td>0.4026</td>
<td>0.00001</td>
</tr>
</tbody>
</table>

**Table 2: Error values of the model**

<table>
<thead>
<tr>
<th></th>
<th>Test Data</th>
<th>All Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>MAPE</td>
<td>% 3.66</td>
<td>% 1.46</td>
</tr>
</tbody>
</table>

The error of model according to MAPE metric 3.66% in test data and 1.46% in general as it can be seen in Table 2. Actual and predicted values of the model can be seen in Figure 1.

**Figure 1: The actual and predicted values of electricity consumption in train, test and all data set.**
CONCLUSION

The consumption of energy, especially electrical energy follows an increasing trend in Turkey like other developing countries. So, future predictions are really important in terms of economical, environmental, industrialization and management problems concerning electricity consumption. Electricity consumption can be influenced by many economic indicators, especially for a country like Turkey that is foreign-dependent in terms of energy and owns very unstable economy having an extremely sensitive nature against domestic and foreign developments in politics, economics and market. The most important ones of these economic indicators are import, export, GDP and population. With so many effecting factors, it is so difficult to plan energy policies, capital investments, energy activities and income.

SVR hybridized with chaotic methods which provide a simple and effective way of prediction is seen that increases the accuracy and improves the quality of predictions in many studies. The annual electricity consumption of Turkey is modelled by using mentioned economic indicators. SVR method trained by chaotic particle swarm algorithm is used. The SVR method trained by CPSO with error performance of %3.66 indicates that this model can be used an alternative method to classical regression and artificial neural networks.
REFERENCES


