Forecasting Electricity Consumption In Bulgaria By Studying Its Dependence On Socio-Economic And Demographic Variables

Kostadin Yotov, Emil Hadzhikolev, Stanka Hadzhikoleva

Abstract: The tendency of continuous increase in electricity consumption cannot continue forever. Many countries are actively conducting policies to reduce electricity consumption and increase energy efficiency. Their efforts are targeted in several directions - renovation of residential and industrial buildings, use of energy-saving appliances in households, optimization of electricity consumption in industry, etc. The application of new technologies and the energy generation from renewable sources have created a mix of factors that have an unpredictable impact on electricity consumption. This greatly complicates the estimation of its dependence on the other factors as well as its long-term prediction. The article presents the results of conducted research on the relationship between electricity consumption in Bulgaria and six socio-economic and demographic factors – GDP, Energy intensity, Population, Annual income, Electricity price for the industry and Electricity price for households. The results presented are part of a larger study to create a comprehensive model for forecasting electricity consumption in Bulgaria.

Index Terms: electricity consumption prediction, forecasting electricity consumption, relationship between electricity consumption and socio-economic and demographic factors.

1. INTRODUCTION

One of the most important factors for the effective planning and management of a country's energy system is the forecasting of electricity consumption. The increasing global energy demand and the high and volatile energy prices create dependence on energy importers. Many countries are developing projects aimed at diversifying energy sources and suppliers, integration and interconnection of energy markets. Policies and strategies are directed at improving energy efficiency, using renewable energy sources, reducing reverse climate impacts, etc. [1]. In 2012, an ambitious target was set in the European union – to achieve a 20% reduction in energy consumption within the next 8 years. In this regard, all European countries had to take specific measures and legal regulations, to periodically monitor their compliance and the achievement of the intended results. In 2018, the agreement was revised, and a new goal was set – to improve the energy efficiency by at least 32.5% by 2030 [2].

The active implementation of policies to reduce electricity consumption has many aspects – renovation of residential and industrial buildings, use of energy-saving appliances in households, optimization of electricity consumption in the industry by means of new technologies, and so on. The application of new technologies and energy generation from renewable sources has created a mix of factors that have an unpredictable impact on energy consumption. This greatly complicates the estimation of its dependence on the other factors and its long-term forecasting. A few years ago, it was widely accepted that GDP growth led to increased electricity consumption. However, the reality in Bulgaria over the last 2 years disproves this claim. In 2018, electricity consumption was 2.84% lower than it was in 2017, and in 2019 it also dropped compared to 2018 – by 1.07%. The reasons for this are numerous and varied. The favorable climate over the last two years, characterized by milder winters and cooler summers, has led to decreased need for heating and cooling. Renovation of residential buildings, which has recently been funded with government support and various energy efficiency programs, has also had a beneficial effect on households. The use of new energy-efficient technologies in industrial production and the increasing supply of energy-saving appliances (stoves, refrigerators, washing machines, lighting, etc.) is another important detail on the overall energy picture. Undoubtedly, predicting power consumption is a complex task. There are different approaches to solving it – through statistical analysis methods, numerical approximation techniques, use of artificial intelligence methods, and more.

The object of our research is the electricity system of the Republic of Bulgaria, and the main objective – to effectively forecast electricity consumption by using a software system that integrates different methods. The article presents the results obtained by using numerical methods. The methods are applied to study the relations between electricity consumption and six factors – GDP, Energy Intensity, Population, Annual Income, Electricity price for the industry and Electricity price for households. Studies on the relations between power consumption and the various pairs of factors listed are also presented.

2 MAIN FACTORS IN FORECASTING POWER CONSUMPTION

Each company in the industrial sector, every farmer and even every single household strive to reduce their electricity consumption, even as production increases and quality of life rises. The prediction of power consumption in a national energy system is directly related to the correct assessment of the factors affecting it. This includes a number of macroeconomic and demographic indicators, social parameters, climate conditions, and so on. [3], [4]. The following indicators are mainly taken into account:

- Gross domestic product (GDP):
• Population;
• Per capita income;
• Price of electricity;
• Expected temperatures for the relevant period;
• Energy efficiency, etc.

A study we conducted earlier found that the correlative dependencies between each of these factors and nationwide electricity consumption are not fully determined. We used statistical methods, applying regression analysis to create models of these dependencies and a corresponding error estimate. Regression models on the „electricity consumption“ factor, as well as each of the many factors that determine it, have proved to have relatively poor prognostic potential. For example, an error of 907,3004926411 thousand tons o.e. (thousand tonnes of oil equivalent) in the study of the relationship „Energy intensity – Electricity consumption“ is unacceptable. This article presents a study conducted using extrapolation procedures that produce results of higher predictive accuracy. It should be noted here that in the scientific literature there are various models for predicting power consumption. Perhaps the most complex model for Bulgaria is the one created in 2017 by a team of scientists from the Bulgarian Academy of Sciences. The model was custom designed by Bulgarian Energy Holding EAD to develop a national energy strategy with a focus on electricity. It includes the following components:

\[ C_t = C_t^{H} + C_t^{A} + C_t^{I} + C_t^{S} + C_t^{T}, \]

where \( C_t \) is the total electricity consumption for the period \( t \) found as the sum of electricity consumption in different sectors: \( C_t^{H} \) is the electricity consumption of households, \( C_t^{A} \) in the Agriculture, forestry and fisheries sector, \( C_t^{I} \) in the Industry sector, \( C_t^{S} \) in the Services sector, \( C_t^{T} \) in the Transportation sector [4]. Critics point out various weaknesses in the model. The most discussed deficiencies are the incorrect accounting of the dynamics of the Bulgarian population, which has been declining at a rapid pace in recent years, anticipating moderate GDP growth in the coming years and disregarding the highly likely scenario involving an economic crisis, etc. In addition, according to experts, Bulgaria has enormous potential for increasing its energy efficiency which would also have an impact on electricity consumption. The specialized literature includes descriptions of various models and approaches for predicting electricity consumption in different countries. Kankal and team have developed a model for predicting electricity consumption in Turkey. It is based on socio-economic and demographic variables (gross domestic product-GDP, population, import and export amounts, and employment). The authors use an artificial neural network and regression analyses [5]. Kumar and Jain use a different approach. They have created Time series models (Grey-Markov, Grey Model with rolling mechanism and singular spectrum analysis) to forecast the consumption of conventional energy in India. They have used the Grey-Markov model to forecast the crude-petroleum consumption, Grey-Model with rolling mechanism to predict the coal, electricity (in utilities) consumption, and SSA to predict natural gas consumption [6]. The grey multivariable model for predicting industrial energy consumption in China is described in [7]. It is an improved grey model with convolution integral GMC(1, n). A nonlinear optimization model is developed to find the optimal parameters. The proposed model exhibits a higher accuracy than GMC(1, n), SARMA and GM(1, 1). Ogcu and team solve the issue of forecasting electricity consumption with neural networks and support vector regression [8]. Artificial intelligence methods such as support vector machine (SVM) and artificial neural networks (ANN) are also largely used in forecasting power consumption [9].

3 NUMERICAL EXTRAPOLATION OF THE RELATION BETWEEN ELECTRICITY CONSUMPTION AND THE FACTORS GDP, ENERGY INTENSITY, POPULATION, ANNUAL INCOME, ELECTRICITY PRICES FOR INDUSTRIES AND HOUSEHOLDS

Knowing which factors influence electricity consumption more strongly, helps to create more efficient and complex prediction models. This study uses publicly available annual data on the factors investigated as follows: Gross domestic product, determined by the income approach (2000-2017): Population (2003-2017); Average annual per capita income (2004-2017); Price of electricity consisting of two components – price for the industry and price for households (2007-2017); and Energy intensity (2004-2017). The data is taken from the INFOSTAT Information system which provides statistical information on the demographic and socio-economic status and development of Bulgaria [10]. INFOSTAT is maintained by the National Statistical Institute of Bulgaria [11].

Let

\[ X_i = \{ x_{i1}, x_{i2}, ..., x_{in} \}, i = 1, 2, ..., 6 \]

be the set of input data from a sample for a specific factor \( X_i \). Then the main nodes on which we will build the prediction models are denoted by:

\[ (x_{ij}, y_{ij}), i = 1, 2, ..., 6, j = 1, 2, ..., n, \]

where \( y_{ij} \) is the value of electricity consumption in the year of measurement of \( x_{ij} \).

Then, if \( x_{i(n+1)} \) is the value of the factor \( X_i \) for which we will forecast consumption and expect that \( x_{i(n+1)} \) is within the sample, i.e.

\[ x_{i(n+1)} \in (\min x_i, \max x_i), \]

it is appropriate to use interpolation procedures.

In most cases, however, the values of consumption determinants follow an increasing or decreasing trend. So, the consumption during the forecast period requires us to use as an argument the value of the factor beyond the sample boundary. For example, when studying the dependence of electricity consumption on GDP, it should be noted that Bulgaria’s GDP, with only slight fluctuations, has been steadily increasing in recent years (fig. 1), according to the World Bank statistics [12]. Moreover, \( x_{i(n+1)} \) does not belong to the specified interval defined by the sample of GDP values. In this case it is most appropriate to use extrapolation methods.
For the other factors, similar trends are observed in the long run – for an upward or a downward trend. To extrapolate the dependence of consumption on the factors mentioned above, MatLab environment was used. The most effective extrapolation model for each factor was found through the following example code:

```matlab
real_y=9737.9; % the actual measured value of consumption methods = {'linear', 'nearest', 'next', 'previous', 'pchip', 'spline'};
% methods - extrapolation methods
n=length(methods); % n - number of extrapolation methods in array methods
for i=1:1:n
    % obtaining an estimated value through extrapolation
    forecast(1,i)=interp1(x,y,test,interp(methods(i)),'extrap');
    % error calculation
    error(i)=abs(real_y-forecast(1,i));
end;
```

The obtained results show that in numerical integration, the absolute errors are within a few dozen gigawatts. These results are significantly better than the results of our experiments to model the same dependencies through correlation analysis. Similar calculations have been made for 4-year periods of power consumption forecasting. To select the most efficient extrapolation methods, error estimates calculated with a MatLab script are used. In the presented example, 6 methods (linear, nearest, next, previous, pchip, and spline) were used, with parameter vectors extracted from the measured data. The most effective extrapolation methods were used in the example, with parameter vectors obtained by numerical integration.

<table>
<thead>
<tr>
<th>Extrapolation method</th>
<th>Predicted result</th>
<th>Actual value</th>
<th>Absolute error</th>
<th>Factor related to energy consumption</th>
<th>Year to be forecasted</th>
<th>Extrapolation method</th>
<th>Predicted result</th>
<th>Actual value</th>
<th>Absolute error</th>
<th>Factor related to energy consumption</th>
<th>Year to be forecasted</th>
</tr>
</thead>
<tbody>
<tr>
<td>linear extrapolation</td>
<td>9705,7777777777</td>
<td>9737.9</td>
<td>32.12222222222260</td>
<td>Population - Electricity consumption</td>
<td>(2003-2016)</td>
<td>linear extrapolation</td>
<td>9645,153490611458</td>
<td>9737.9</td>
<td>38.67003891050555</td>
<td>Electricity price for the households – Electricity consumption (in the households) 2007-2016</td>
<td>2017</td>
</tr>
<tr>
<td>linear extrapolation</td>
<td>2357,3703891051</td>
<td>2318.7</td>
<td>457.3199999999952</td>
<td>linear extrapolation</td>
<td>2600.6</td>
<td>linear extrapolation</td>
<td>1861.380000000005</td>
<td>2318.7</td>
<td>457.3199999999952</td>
<td>Electricity price for the households – Electricity consumption (in the households) 2007-2016</td>
<td>2017</td>
</tr>
</tbody>
</table>

The results of the calculations made are presented in table 1. All consumption values and absolute errors are measured in thousands of tons of oil equivalent (toe).

### TABLE 1. Extrapolation of the Dependence of Electricity Consumption on Individual Factors Over a One-Year Forecast Period

<table>
<thead>
<tr>
<th>Factor related to energy consumption</th>
<th>GDP – Electricity consumption (2000-2016)</th>
<th>Year to be forecasted</th>
<th>Extrapolation method</th>
<th>Predicted result</th>
<th>Actual value</th>
<th>Absolute error</th>
<th>Factor related to energy consumption</th>
<th>Year to be forecasted</th>
<th>Extrapolation method</th>
<th>Predicted result</th>
<th>Actual value</th>
<th>Absolute error</th>
<th>Factor related to energy consumption</th>
<th>Year to be forecasted</th>
</tr>
</thead>
</table>

The results of the calculations made are presented in table 2.

### TABLE 2. Evaluation of the Error in Extrapolation of the Dependence of Electricity Consumption on the Individual Factors for a 4-Year Forecasting Period

<table>
<thead>
<tr>
<th>Forecasting period</th>
<th>Extrapolation method</th>
<th>Mean absolute error (GWh)</th>
<th>Mean absolute percentage error (%)</th>
<th>Mean squared error (GWh)</th>
<th>Factor related to energy consumption</th>
</tr>
</thead>
</table>

For the factors of energy intensity and electricity consumption, the most accurate method of extrapolation for each factor was found through the minimization of the absolute errors. The obtained results show that in numerical integration, the absolute errors are within a few dozen gigawatts. These results are significantly better than the results of our experiments to model the same dependencies through correlation analysis. Similar calculations have been made for 4-year periods of power consumption forecasting. To select the most efficient extrapolation methods, error estimates calculated with a MatLab script are used. In the presented example, 6 methods (linear, nearest, next, previous, pchip, and spline) were used, with parameter vectors extracted from the measured data. The most effective extrapolation methods were used in the example, with parameter vectors obtained by numerical integration.

![GDP of Bulgaria for the period 1997-2017](image)
Forecasting period  |  2014 – 2017  
Extrapolation method  |  cubic/pcchip extrapolation  
Mean absolute error (GWh)  |  682.4686899503004  
Mean absolute percentage error (%)  |  1.846271431040587  
Mean squared error (GWh)  |  585921.1477165715  
Factor related to energy consumption  |  Population – Electricity consumption (2003-2013)  
Forecasting period  |  2014 – 2017  
Extrapolation method  |  linear extrapolation  
Mean absolute error (GWh)  |  2078.345102250802  
Mean absolute percentage error (%)  |  5.61524622749457  
Mean squared error (GWh)  |  52.237.143.431514022  
Factor related to energy consumption  |  Annual income – Electricity consumption (in the households) (2004-2013)  
Forecasting period  |  2014 – 2017  
Extrapolation method  |  linear extrapolation  
Mean absolute error (GWh)  |  108.7222875226039  
Mean absolute percentage error (%)  |  1.234343563070950  
Mean squared error (GWh)  |  17304.41860196234  
Factor related to energy consumption  |  The price of electricity for the industry - Electricity consumption (in the Industry sector) (2007-2013)  
Forecasting period  |  2014 – 2017  
Extrapolation method  |  cubic/pcchip extrapolation  
Mean absolute error (GWh)  |  73.976732459482490  
Mean absolute percentage error (%)  |  0.697945408687615  
Mean squared error (GWh)  |  7496.98976340804  
Factor related to energy consumption  |  The price of electricity for the households – Electricity consumption (in the households) (2007-2013)  
Forecasting period  |  2014 – 2017  
Extrapolation method  |  cubic/pcchip extrapolation  
Mean absolute error (GWh)  |  1003.054090909111  
Mean absolute percentage error (%)  |  11.351638528091435  
Mean squared error (GWh)  |  1253.753.120885177  

An analysis of the results of one-year and four-year forecasts show that extrapolation methods work well for a forecasting period not too distant from the sample, but when data is forecast for a longer period, errors become larger.

4 NUMERICAL EXTRAPOLATION OF THE CONNECTION BETWEEN ELECTRICITY CONSUMPTION AND PAIRED FACTORS

To investigate the dependences of electricity consumption on groups of paired factors, similar studies were conducted using two-dimensional extrapolation:

\[ P = F_{ij}(X_i, X_j), \ i \neq j, \]

where: \( P \) is the power consumption considered as a function \( F_{ij} \) of the variables \( X_i \) and \( X_j \) which model some of the factors studied – GDP, Energy Intensity, Population, Annual Income, Electricity price for Industry, and Electricity price for households.

For example, the relationship between the consumption \( P \) with the pair of factors – average annual income \( X_1 \) and population \( X_2 \), is investigated with different types of extrapolation. A minimal prediction error was obtained with the thin-plate spline method (fig. 2).

**Fig. 2 View of the surface extrapolating the relationship of the pair of factors „average annual income” and “population” with the power consumption**

For each of the studied dependences of the electricity consumption on the pair of factors, different extrapolation methods have been experimented and an error estimate has been made. The method with the least absolute error has been chosen as the most effective one. Table 3 summarizes the most effective extrapolation methods, as well as the actual and estimated values of electricity consumption and the absolute errors measured in thousands of tons of oil equivalent.

**TABLE 3. EXTRAPOLATION OF THE DEPENDENCE OF ELECTRICITY CONSUMPTION ON TWO FACTORS (WITH A ONE-YEAR FORECAST)**

<table>
<thead>
<tr>
<th>Group of factors related to energy consumption (GDP, Intensity of the economy) – Electricity Consumption 2000-2016</th>
<th>Group of factors related to energy consumption (GDP, Population) - Electricity Consumption 2004-2016</th>
</tr>
</thead>
<tbody>
<tr>
<td>Year to be forecasted</td>
<td>2017</td>
</tr>
<tr>
<td>Extrapolation method</td>
<td>( P = 0.07914X_1 + (1.858e + 04)X_2 – 5150 )</td>
</tr>
<tr>
<td>Forecasted value</td>
<td>9737.9</td>
</tr>
<tr>
<td>Measured value</td>
<td>9737.9</td>
</tr>
<tr>
<td>Absolute error</td>
<td>3.2276</td>
</tr>
<tr>
<td>Year to be forecasted</td>
<td>2017</td>
</tr>
<tr>
<td>Extrapolation method</td>
<td>( P = 14.65X_1 + 0.6217X_2 – (1.559e – 05)x_1^2 – (1.703e – 06)x_2x_1 – (3.435e – 08)x_1^2 – (2.778e + 06) )</td>
</tr>
<tr>
<td>Forecasted value</td>
<td>9926.2</td>
</tr>
<tr>
<td>Measured value</td>
<td>9737.9</td>
</tr>
<tr>
<td>Absolute error</td>
<td>88.3</td>
</tr>
</tbody>
</table>

**TABLE 3. EXTRAPOLATION OF THE DEPENDENCE OF ELECTRICITY CONSUMPTION ON TWO FACTORS (WITH A ONE-YEAR FORECAST)**

<table>
<thead>
<tr>
<th>Group of factors related to energy consumption (GDP, Intensity of the economy) – Electricity Consumption 2000-2016</th>
<th>Group of factors related to energy consumption (GDP, Population) - Electricity Consumption 2004-2016</th>
</tr>
</thead>
<tbody>
<tr>
<td>Year to be forecasted</td>
<td>2017</td>
</tr>
<tr>
<td>Extrapolation method</td>
<td>( P = (4.587e + 05)x_1 – 0.2366x_2 – 0.05977x_1x_2 + (1.779e – 08)x_1^2 + (7.778e + 05) )</td>
</tr>
<tr>
<td>Forecasted value</td>
<td>9705.313782180077</td>
</tr>
<tr>
<td>Measured value</td>
<td>9737.9</td>
</tr>
<tr>
<td>Absolute error</td>
<td>32.5862178199222</td>
</tr>
<tr>
<td>Year to be forecasted</td>
<td>2017</td>
</tr>
<tr>
<td>Extrapolation method</td>
<td>Thin-plate spline</td>
</tr>
<tr>
<td>Forecasted value</td>
<td>2298.8</td>
</tr>
<tr>
<td>Measured value</td>
<td>2318.7</td>
</tr>
<tr>
<td>Absolute error</td>
<td>19.8684</td>
</tr>
<tr>
<td>Group of factors related to energy consumption (Price of electricity for the households, Annual income) - Electricity Consumption in the households 2007-2016</td>
<td>Group of factors related to energy consumption (Annual income, Population) - Electricity Consumption in the households 2004-2016</td>
</tr>
<tr>
<td>Year to be forecasted</td>
<td>2017</td>
</tr>
<tr>
<td>Extrapolation method</td>
<td>Lowess (Linear)</td>
</tr>
<tr>
<td>Forecasted value</td>
<td>2337.4</td>
</tr>
</tbody>
</table>
The continuous increase in electricity consumption has put the hand of economic systems from toys to sophisticated machines – economic systems – family, company, municipality, state, etc. – are getting more and more dependent on electricity. Any device – from toys to sophisticated machines – needs electricity to function. Their mechanic analogues and hand-built devices are generally ineffective or uninteresting. The continuous increase in electricity consumption has put the task of efficient use and reduced consumption in the foreground. Reducing the electricity costs in one system has different dimensions. One possibility is to increase energy efficiency, which can be achieved through buildings renovation, use of power-saving appliances and power transmission elements, and many more. Another important aspect typical of large electricity consumers and power producers is the reduction of energy surplus produced, which explicitly or implicitly leads to higher electricity bills for households, higher prices of goods and services produced, etc. Predicting the power consumption of a system allows for its efficient planning and management. The article presents the results of a study of the dependences between electricity consumption and socio-economic and demographic factors – GDP, Energy Intensity, Population, Annual Income, Electricity Prices for the Industries and Households in Bulgaria. One-factor and two-factor extrapolation models have been created. They have helped investigate the relations between power consumption and individual factors, as well as between power consumption and some pairs of factors. For a one-year prediction, the most successful models are found to be generated using the factors GDP, Energy Intensity, Average Annual Income, and the pair of factors (GDP & Intensity of the economy). The presented results are part of a larger project aimed at creating a complex model for forecasting electricity consumption in the Republic of Bulgaria.

ACKNOWLEDGMENT

The work is funded by the SP19-FMI-004 project at the Research Fund of the University of Plovdiv "Paisii Hilendarski".

### REFERENCES


