USING IMPROVED GREY FORECASTING MODEL TO ESTIMATE THE ELECTRICITY CONSUMPTION DEMAND IN VIETNAM

Phan Van Thanh*
Scientific management & International Affairs, Quang Binh University 
312 Ly Thuong Kiet Street, Dong Hoi city, Quang Binh Province

Abstract: On the basis of the grey prediction models, this study uses the previous data (from 1980 to 2014) from the website of the World Bank and applies two algorithm models to forecast the electricity consumption in Vietnam. The simulation results show that Fourier Residual Modified GM (1, 1) (abbreviated as FRMGM (1, 1)) is an effective model with an average accuracy of prediction at 99.13%. Therefore, the FRMGM (1, 1) model is strongly suggested for forecasting the electricity consumption demand in Vietnam.

Keywords: electricity consumption demand, GM (1, 1); FRMGM (1, 1), Vietnam

1 Introduction

Worldwide energy consumption is rising fast because of the increase in human population, continuous pressures for better living standards, emphasis on large-scale industrialization in developing countries, and the need to sustain positive economic growth rates. Given this fact, a sound forecasting technique is essential for accurate investment planning of energy production/generation and distribution. In Vietnam, the electricity consumption forecasting plays a significant role in strategic planning of an electric utility company since many activities need to be planned in advance such as location and construction of new substations, creating new transmission and distribution networks, improving existing systems, and/or construction of new power generation plants.

The grey system theory, established in the 1980s by Deng [1], is a quantitative method dealing with grey systems that are characterized by both partially known and partially unknown information [2–5]. As a vital part of the grey system theory, grey forecasting models with their advantages in dealing with uncertain information by using as few as four data points [6, 7]. These models have been successfully applied in various fields such as tourism [8, 9], energy [10, 11], financial and economic [12–14], and IC industry [15].

Because the grey forecasting model has two major advantages, namely fewer samples requirement and simple computation, this study tries to apply two models to estimate the

* Corresponding: thanhkem2710@gmail.com
Submitted: July 27, 2017; Revised: March 09, 2018; Accepted: April 27, 2018
Vietnamese electricity consumption. The results will be helpful for the top managers in formulating policies as well as orienting development in the power sector. The remaining of this paper is organized as follows. In section 2, the concept of the grey theory is presented and the fundamental function of traditional grey forecasting model “(GM (1, 1)” and modified GM (1, 1) by Fourier series are shown. On the basis of the fundamental function of GM (1, 1) and FRMGM (1, 1), the empirical results are discussed in section 3. Finally, section 4 concludes this paper.

2 The concept of grey prediction

2.1 GM (1, 1) algorithm

GM (1, 1) is the basic model of grey forecasting modelling, a first-order differential model with one input variable which has been successfully applied in many fields of research. It is obtained as in the following procedure.

Step 1: Let raw matrix \(X^{(0)}\) stand for the non-negative original historical time series data

\[
X^{(0)} = \{x^{(0)}(t_i)\}, \quad i = 1, 2, ..., n
\]  

where \(x^{(0)}(t_i)\) is the value at time \(t_i\), and \(n\) is the total number of modelling data points.

Step 2: Construct \(X^{(1)}\) by one time accumulated generating operation (1-AGO), which is

\[
X^{(1)} = \{x^{(1)}(t_i)\}, \quad i = 1, 2, ..., n
\]  

where

\[
x^{(1)}(t_k) = \sum_{i=1}^{k} x^{(0)}(t_i), \quad k = 1, 2, ..., n
\]  

Step 3: \(X^{(1)}\) is a monotonic increasing sequence which is modelled by the first-order linear differential equation

\[
\frac{dX^{(1)}}{dt} + aX^{(1)} = b
\]  

where parameter “\(a\)” is called the developing coefficient and “\(b\)” is named the grey input.

Step 4: In order to estimate parameter “\(a\)” and “\(b\)”, Eq. (4) is approximated as:

\[
\frac{\Delta X^{(1)}(t_k)}{dt_k} + aX^{(1)}(t_k) = b
\]  

where

\[
\Delta X^{(1)}(t_k) = x^{(1)}(t_k) - x^{(1)}(t_{k-1}) = x^{(0)}(t_k)
\]
\[ \Delta t_k = t_k - t_{k-1} \] (7)

If the sampling time interval is units, then let \( \Delta t_k = 1 \), using

\[ z^{(1)}(t_k) = px^{(1)}(t_k) + (1 - p)x^{(1)}(t_{k-1}), \ k = 2, 3, ..., n \] (8)

to replace \( X^{(1)}(t_k) \) in Eq. (1.5), we obtain

\[ x^{(0)}(t_k) + az^{(1)}(t_k) = b, \ k = 2, 3, ..., n \] (9)

where \( z^{(1)}(t_k) \) in Eq. (8) is the termed background value, and \( p \) is the production coefficient of the background value in the range of \((0, 1)\), which is traditionally set to 0.5.

Step 5: From Eq. (9), the value of parameter “\( a \)” and “\( b \)” can be estimated using the least-square method. That is

\[
\begin{bmatrix}
a \\
b
\end{bmatrix} = (B^T B)^{-1} B^T Y_n
\] (10)

where

\[
B = \begin{bmatrix}
-z^{(1)}(t_2) & 1 \\
-z^{(1)}(t_3) & 1 \\
\vdots & \vdots \\
-z^{(1)}(t_n) & 1
\end{bmatrix}
\] (11)

and

\[
Y_n = \begin{bmatrix}
x^{(0)}(t_2), x^{(0)}(t_3), ..., x^{(0)}(t_n)
\end{bmatrix}^T
\] (12)

Step 6: The solution of Eq. (4) can be obtained after parameter “\( a \)” and “\( b \)” has been estimated. That is

\[
\hat{x}^{(1)}(t_k) = \left( x^{(0)}(t_k) - \frac{b}{a} \right) e^{-a(t_k-t_1)} + \frac{b}{a}, k = 1, 2, 3, ...
\] (13)

Step 7: Applying an inverse accumulated generating operation (IAGO) to \( \hat{x}^{(1)}(t_k) \), the predicted datum of \( x^{(0)}(t_k) \) can be estimated as:

\[
\begin{cases}
\hat{x}^{(0)}(t_1) = x^{(0)}(t_1) \\
\hat{x}^{(0)}(t_k) = \hat{x}^{(0)}(t_{k-1}), \ k = 2, 3, ... \\
\hat{x}^{(0)}(t_k) = \hat{x}^{(0)}(t_k) - x^{(1)}(t_{k-1}) \text{, } k = 2, 3, ...
\end{cases}
\] (14)
2.2 Fourier Residual Modified GM (1, 1) algorithm

The overall procedure to obtain the modified model is as follows:

Let \( x \) be the original series of \( m \) entries and \( v \) is the predicted series (obtained from GM (1, 1)). On the basis of the predicted series \( v \), a residual series named \( \varepsilon \) is defined as

\[
\varepsilon = \{ \varepsilon(k) \}, \ k = 2, 3, \ldots m
\]  

where

\[
\varepsilon(k) = x(k) - v(k), \ k = 2, 3, \ldots m
\]

According to the definition of the Fourier series, the residual sequence of GM (1, 1) can be approximately expressed as

\[
\hat{\varepsilon}(k) = \frac{1}{2} a_{(0)} + \sum_{i=1}^{Z} \left[ a_i \cos\left(\frac{2\pi i}{m-1} (k)\right) + b_i \sin\left(\frac{2\pi i}{m-1} (k)\right) \right], \ k = 1, 2, 3, \ldots, m
\]

where \( Z = \left(\frac{m-1}{2}\right) - 1 \) is called the minimum deployment frequency of the Fourier series [13] and only takes an integer number; therefore, the residual series is rewritten as:

\[
\varepsilon = PC
\]

where

\[
P = \begin{bmatrix}
\frac{1}{2} \cos\left(\frac{2\pi 1}{m-1} \times 2\right) & \sin\left(\frac{2\pi 1}{m-1} \times 2\right) & & & \cos\left(\frac{2\pi Z}{m-1} \times 2\right) & \sin\left(\frac{2\pi Z}{m-1} \times 2\right) \\
\frac{1}{2} \cos\left(\frac{2\pi 1}{m-1} \times 3\right) & \sin\left(\frac{2\pi 1}{m-1} \times 3\right) & & & \cos\left(\frac{2\pi Z}{m-1} \times 3\right) & \sin\left(\frac{2\pi Z}{m-1} \times 3\right) \\
& \ldots & & \ldots & \ldots & \ldots \\
\frac{1}{2} \cos\left(\frac{2\pi 1}{m-1} \times m\right) & \sin\left(\frac{2\pi 1}{m-1} \times m\right) & \ldots & \cos\left(\frac{2\pi Z}{m-1} \times m\right) & \sin\left(\frac{2\pi Z}{m-1} \times m\right) \\
& & & & & \\
\end{bmatrix}
\]

and

\[
C = [a_0, a_1, b_1, a_2, b_2, \ldots, a_Z, b_Z]
\]

Parameters \( a_0, a_1, b_1, a_2, b_2, \ldots, a_Z, b_Z \) are obtained by using the ordinary least-square method (OLS) which results in the equation

\[
C = (P^T P)^{-1} P^T [\varepsilon]'
\]

Once the parameters are calculated, the modified residual series is then achieved on the basis of the following expression:

\[
\hat{\varepsilon}(k) = \frac{1}{2} a_{(0)} + \sum_{i=1}^{Z} \left[ a_i \cos\left(\frac{2\pi i}{m-1} (k)\right) + b_i \sin\left(\frac{2\pi i}{m-1} (k)\right) \right]
\]
From the predicted series \( V \) and \( \hat{e} \), the Fourier modified series \( \hat{v} \) of series \( v \) is determined by:

\[
\hat{v} = \{\hat{v}_1, \hat{v}_2, \hat{v}_3, ..., \hat{v}_k, ..., \hat{v}_n\}
\]  \hspace{1cm} (24)

where

\[
\hat{v} = \begin{cases} 
\hat{v}_1 = v_1 \\
\hat{v}_k = v_k + \hat{e}_k & (k = 2, 3, ..., m)
\end{cases}
\]  \hspace{1cm} (25)

2.3 Valuation performances

Regarding the evaluation performance of the volatility model for forecasting, there are some common approaches, including the root of mean square error, mean absolute error, and mean absolute percentage error (MAPE).

This study uses MAPE [16] to identify the grey prediction models with good performance; small MAPE is taken to indicate good forecasting performance. MAPE is defined as follows:

\[
MAPE = \frac{1}{n} \sum_{k=2}^{n} \frac{x^{(0)}(k) - \hat{x}^{(0)}(k)}{x^{(0)}(k)} \times 100\% 
\]  \hspace{1cm} (26)

Where \( x^{(0)}(k) \) indicates the actual value in time period \( k \); \( \hat{x}^{(0)}(k) \) indicates the forecast value in time period \( k \).

And the grade of MAPE is divided into four levels [17]. More detailed shown in Table 1.

<table>
<thead>
<tr>
<th>MAPE</th>
<th>( \leq 10% )</th>
<th>10–20%</th>
<th>20–50%</th>
<th>&gt;50%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grade levels</td>
<td>Excellent</td>
<td>Good</td>
<td>Qualified</td>
<td>Unqualified</td>
</tr>
</tbody>
</table>

In addition, indicator \( \rho \) is used to evaluate the predicted accuracy of the forecast model, higher \( \rho \) is taken to indicate a good forecasting performance. This is defined as follows:

\[
\rho = 100 - MAPE
\]  \hspace{1cm} (27)

3 Data and empirical results

The data on electricity consumption from 1980 to 2014 in Vietnam were obtained from the World Bank on December 12\textsuperscript{th} 2016 [18]. From the historical data, parameters \( a \) and \( b \) are –0.1323 and 0.7783 calculated through the algorithms of the GM (1, 1) model (section 2.1). The fundamental of GM (1, 1) for electricity consumption is found as follows:
\[ \hat{x}^{(1)}(k) = \left[ 3.29 + \frac{0.7783}{0.1323} \right] e^{-0.1323(k-1)} - \frac{0.7783}{0.1323} \]

The residual series attained form GM (1, 1) is then modified with the Fourier series, which results in the modified model FRMGM (1, 1) according to the algorithms stated in section 2.2. The evaluation index of GM (1, 1) and FRMGM (1, 1) is summarized in Table 2.

**Table 2.** Forecasted value of the electricity consumption (Units: kWh per capita)

<table>
<thead>
<tr>
<th>Years</th>
<th>Actual value</th>
<th>Forecasted value by GM(1,1)</th>
<th>% error</th>
<th>Forecasted value by FRMGM(1,1)</th>
<th>% error</th>
</tr>
</thead>
<tbody>
<tr>
<td>1980</td>
<td>54.562</td>
<td>54.562</td>
<td>0.00</td>
<td>54.562</td>
<td>0.00</td>
</tr>
<tr>
<td>1981</td>
<td>55.115</td>
<td>39.942</td>
<td>27.53</td>
<td>53.756</td>
<td>2.47</td>
</tr>
<tr>
<td>1982</td>
<td>57.859</td>
<td>44.639</td>
<td>22.85</td>
<td>59.218</td>
<td>2.35</td>
</tr>
<tr>
<td>1983</td>
<td>57.541</td>
<td>49.889</td>
<td>13.30</td>
<td>56.182</td>
<td>2.36</td>
</tr>
<tr>
<td>1984</td>
<td>67.011</td>
<td>55.756</td>
<td>16.80</td>
<td>68.370</td>
<td>2.03</td>
</tr>
<tr>
<td>1985</td>
<td>70.344</td>
<td>62.314</td>
<td>11.42</td>
<td>68.985</td>
<td>1.93</td>
</tr>
<tr>
<td>1986</td>
<td>73.412</td>
<td>69.642</td>
<td>5.14</td>
<td>74.771</td>
<td>1.85</td>
</tr>
<tr>
<td>1987</td>
<td>80.146</td>
<td>77.833</td>
<td>2.89</td>
<td>78.787</td>
<td>1.70</td>
</tr>
<tr>
<td>1988</td>
<td>87.176</td>
<td>86.987</td>
<td>0.22</td>
<td>88.535</td>
<td>1.56</td>
</tr>
<tr>
<td>1989</td>
<td>94.158</td>
<td>97.217</td>
<td>3.25</td>
<td>92.799</td>
<td>1.44</td>
</tr>
<tr>
<td>1990</td>
<td>98.096</td>
<td>108.650</td>
<td>10.76</td>
<td>99.455</td>
<td>1.39</td>
</tr>
<tr>
<td>1991</td>
<td>102.004</td>
<td>121.429</td>
<td>19.04</td>
<td>100.645</td>
<td>1.33</td>
</tr>
<tr>
<td>1992</td>
<td>105.464</td>
<td>135.710</td>
<td>28.68</td>
<td>106.823</td>
<td>1.29</td>
</tr>
<tr>
<td>1993</td>
<td>116.334</td>
<td>151.670</td>
<td>30.37</td>
<td>114.975</td>
<td>1.17</td>
</tr>
<tr>
<td>1994</td>
<td>134.473</td>
<td>169.508</td>
<td>26.05</td>
<td>135.832</td>
<td>1.01</td>
</tr>
<tr>
<td>1995</td>
<td>159.302</td>
<td>189.444</td>
<td>18.92</td>
<td>157.943</td>
<td>0.85</td>
</tr>
<tr>
<td>1996</td>
<td>186.996</td>
<td>211.724</td>
<td>13.22</td>
<td>188.355</td>
<td>0.73</td>
</tr>
<tr>
<td>1997</td>
<td>210.613</td>
<td>236.624</td>
<td>12.35</td>
<td>209.253</td>
<td>0.65</td>
</tr>
<tr>
<td>1998</td>
<td>241.173</td>
<td>264.453</td>
<td>9.65</td>
<td>242.532</td>
<td>0.56</td>
</tr>
<tr>
<td>1999</td>
<td>261.369</td>
<td>295.555</td>
<td>13.08</td>
<td>260.010</td>
<td>0.52</td>
</tr>
<tr>
<td>2000</td>
<td>295.037</td>
<td>330.315</td>
<td>11.96</td>
<td>296.396</td>
<td>0.46</td>
</tr>
<tr>
<td>2001</td>
<td>335.345</td>
<td>369.163</td>
<td>10.08</td>
<td>333.986</td>
<td>0.41</td>
</tr>
<tr>
<td>2002</td>
<td>387.037</td>
<td>412.580</td>
<td>6.60</td>
<td>388.396</td>
<td>0.35</td>
</tr>
<tr>
<td>2003</td>
<td>443.074</td>
<td>461.103</td>
<td>4.07</td>
<td>441.715</td>
<td>0.31</td>
</tr>
<tr>
<td>2004</td>
<td>493.624</td>
<td>515.332</td>
<td>4.40</td>
<td>494.983</td>
<td>0.28</td>
</tr>
<tr>
<td>2005</td>
<td>579.922</td>
<td>575.940</td>
<td>0.69</td>
<td>578.563</td>
<td>0.23</td>
</tr>
</tbody>
</table>
Table 2 clearly shows that the FRMGM (1, 1) model is better than the GM (1, 1) model in this case with the MAPE and $\rho$ is 0.87% and 99.13%, respectively. Therefore, this study suggests that the FRMGM (1, 1) model should be used for the estimation of the electricity consumption demand in the future. The forecasted value in 2018 to 2020 is shown in Table 3.

Table 3. Forecasted value by FRMGM (1, 1)

<table>
<thead>
<tr>
<th>Year</th>
<th>Forecasted value by FRMGM (1, 1)</th>
<th>Electricity consumption (kWh per capita)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2018</td>
<td>2456.85</td>
<td></td>
</tr>
<tr>
<td>2019</td>
<td>2738.37</td>
<td></td>
</tr>
<tr>
<td>2020</td>
<td>3058.10</td>
<td></td>
</tr>
</tbody>
</table>

Table 3 shows that the forecasting values in 2019 and 2020 will be over 2738 and 3058 kWh per capita, respectively. This figure indicates that the demand for electricity consumption in Vietnam will grow significantly in the future. This is a reference for the managers in the power sector to make a good decision in planning and development.

4 Conclusion

The electricity sector is an important industry in the socio-economic development. In Vietnam, the rapid development of the economy leads to the increasing demand for electric consumption. Through simulation, this study found that FRMGM (1, 1) is the fitting model in order to forecast the electricity consumption demand in Vietnam with an accuracy of 99.13%. On the basis of this result, this study strongly suggests that FRMGM (1, 1) is an effective tool to
estimate the electricity consumption demand in the future. Further, the results of this study can be a good reference for the policymakers to make a good decision in the planning and development of the power sector. Due to the data limitations, this study just compares the forecasted and actual values during the period from 1980 to 2014. Future research could also utilize different models of grey forecasting models such as Grey Verhulst model, the GM (2, 1) model to compare with the proposed model in the current study.

References


