Abstract:

Gross Domestic Product (GDP) is the total pecuniary or mart value of all final commodity and services that are produced within country’s borders in a given time. We choose GDP to predict in Iraq since 2000 to 2018. The state and governments rely on GDP to help shape policy or decide how much public spending is affordable. Combining grey regression is a modern statistical technique of modeling, using this type of model is related to its highly accuracy therefore, in this study we used combined grey regression model to predict the gross domestic production of Iraq because it gives less Mse than grey or regression models alone which is equal to 1.3165 and estimated parameter of grey C1, regression C2 are equals to (2.9 and 0.205) respectively and the intercept C3 is equal to 1.8569 the outcome showed that the new model could attain preferable predicting result contrasted with different predicting methods.

Keywords: Gross Domestic Product, Grey Model, Linear Regression Model

(predictions the GDP of Iraq by using Grey –Linear Regression Combined Model)

Sham Azad Rahim [1], Shahla Othman Salih [2], Ayad Othman Hamdin [3], Hindreen Abdullah Taher

1,4 Department of Statistics and Computer, College of Commerce, University of Sulaimania, Sulaimania, Iraq

2,3 Department of Statistics and Informatics, College of Administration & Economic, University of Sulaimania, Sulaimania, Iraq

Email: sham.rahim@univsul.edu.iq, Shahla.salih@univsul.edu.iq, Ayad.hamdin@univsul.edu.iq, Hindreen.taher@univsul.edu.iq

1

الملخص:

إن إجمالي إنتاج المحلي (GDP) هو إجمالي القيمة المالية أو القيمة الإجمالية لجميع السلع والخدمات التي يتم إنتاجها داخل البلد معين في وقت معين. لقد اختارنا إجمالي إنتاج المحلي في العراق لعام 2000 إلى 2018 لتثبيتها في السنوات القادمة. يجب على الحكومة أن تعتمد على إجمالي الناتج المحلي بشكل مباشر للمساعدة في تشكيل السياسة وتحديد مدى تحمله على مقدار الإنفاق العام. إن النموذج بين نموذج الانحدار ونموذج الرمادي هو أسلوب إحصائي جديد للنمذجة، واستخدام هذا النوع من النماذج يعود إلى الدقة العالمية لدى الأساليب، لذلك في هذا البحث نستخدم نموذج الانحدار الرمادي المدمج للتنبؤ بالإجمالي الإنتاج المحلي لدولة العراق لأن هذا النموذج يعطي نتائج أكثر دقة وأثر متوسط مربعات الخطأ للموديل BSAWSY = 1.3165 وهو أقل من متوسط مربعات الخطأ للنموذج الرمادي C2 = 17.0932 و النموذج للنماذج الرمادية C1 (2.9 و 205) على التوالي، والمعدل المقدر للنموذج الرمادي C2 = 15.0063 و النموذج للنماذج الرمادية C1 = 17.0932. أظهرت النتائج أن النموذج المدمج ملائم لتمثيل البيانات في الدراسة.

DOI: http://dx.doi.org/10.25098/4.2.25
1. Introduction

1.1 Gross Domestic Product (GDP)

GDP is one of the major measures of eventual merchandise and serving that produced by a country at a specific time. To capture economic activities GDP is the most substantial index; it limited measure of people’s material living standard so, it is not a good indicator of society’s prosperity.

Calculating GDP usually is done depend on countries’ currencies themselves. However, for comparison among countries GDP values must be turn in to a mutual currency. The current exchange rate is oftentimes used as a conversion but these can lead to false comparison of the true volumes of final goods and services in GDP. A preferable process is to use purchasing power parities (PPPs). PPPs are currency transformer that control for variations in price levels of products between countries and so let to international comparison of the volumes of GDP and the size of economies. (OECD, 2009).

1.2 Definition of GDP

The Bureau of Economic Analysis (2019) defined GDP “as the value of the commodities and services created by the country’s economy less the value of commodities and services applied up in production. The sum of subjective consuming expenses, overall personal local investment, pure export of commodities and services, and government consumption expenses and overall exploitation is also considered as a GDP definition”.

Mankiw (2001) stated that the measurement of the flow of coinage in economy is GDP. Beside this, he mentioned that GDP could be computed in two ways. Firstly, a total income that comes from the production of bread that equal to the sum of wages and profit. Secondly, the overall expenditure on buying bread. Moreover, The Economic Times (2019) defined GDP “as the final value of the goods and services generated within the geographic boundaries of a country during a particular time, normally a year”.

DOI: http://dx.doi.org/10.25098/4.2.25
1.3 Nominal and Real GDP

One of the ways to measure the average dollar value of produced goods at any specific of time is GDP. However, it is not a proper method to gauge the differences in the average quantities of final goods and services as it fluctuates from year to year for two reasons. First, it may increase because a country manufactures more goods and services; we call this increase growth. Second, it may rise because the average cost of produced goods and services could increase, even though the countries might produce the same quantities of goods produced, this called growing inflation.

To separate the increase in GDP that comes from growth and the increase that comes from inflation we calculate the value of GDP in every year using a mutual set of prices. These prices are the ones that takeover in one year – called the standard year. When GDP is used to measure current prices it is called nominal GDP and when it used to measure standard year prices it is called real GDP. Increases in living level are measured by changes in real GDP per person. (Farmer, 1997)

Referring to Garratt (2013), nominal GDP is the same as actual GDP, as far as it will be affected by the change in price and quantity of products, this is in one had. On the other hand, real GDP has a constant price estimate and it affected only by the change in output.

1.4 Component of GDP

O’Sullivan, Sheffrin and Perez, (2012) and Arnold (2008), stated that economists divide GDP into four broad categories, each corresponding to different types of purchases represented in GDP:

1. Consumption expenditures: It is spending on durable and nondurable goods.
2. Private investment expenditures: It is the total of all purchases produced capital goods.
4. pure exports: pure purchases by the external strip (local exports minus local imports)

\[ GDP = C + I + G + (X - M) \]

1.5 Why GDP matters

Fagan (2019), discussed that, GDP used to commerce, economics, government officials depended on statistics to estimate the economy’s growth to build informed decisions.

Where rule is about rates of exchange, taxation and trading policies, Policymakers will look to GDP when.

The speed of economy growing affects business situations, investment resolution, and whether workers can find jobs or not.

Government’s state and local depended on GDP and equal statistics to use in shape policy or choose how much public spending is available, dependent on statistics the Economists study GDP and to help inform their research.
2.1 Simple linear regression

Linear regression is a statistical method that used to summarize and investigate relationship among quantitative variables. One of them is independent variable or explanatory variable that denoted by X, and the other is dependent or response variable that denoted by Y. Simple linear regression study one response variable and one explanatory variable, while multiple linear regression study one response variable and two or more explanatory variables. Regression analyses are used in many filed, the major use of regression is to analyze the effect of a quantitative variable on another quantitative variable, and furthermore it used to predict future values.

2.2 Linear regression model[7]

Regression model is a mathematical function that one of the two variables can be calculated in terms of the other variable. The simplest case for this function is when we have one independent variable that is associated with the dependent variable in a direct line relationship as follows:

\[ \hat{y} = b_0 + b_1x + e_i \]  

where:

\( \hat{y} \): Response variable

\( b_0 \): The intercept

\( b_1 \): The slope

\( x \): Explanatory variable

\( e_i \): Residual

To be able to apply the regression model and to get a perfect result there are five assumptions that have to be achieved:

1. The existence of a linear relationship.
2. Normality distributed.
3. Lack of multicollinearity.
4. Lack of autocorrelation.
5. Homoscedasticity.

Any breach of these assumptions leads to the insufficiency results of OLS method, and therefore these violations have to be addressed or we must use other methods of estimation.

The linear regression model might not be applicable for all data. Non-linearity may be existing, and then transformations of the data might be sufficient to make the linear regression model convenient. The logarithmic transformation might be useful, but there is a preferable technique that can be used to solve the non-linearity problem which is the combined grey linear regression model.
2.3 Grey model\(^{[1][3][8]}\)

The grey system theory proposed by Chinese professor Deng Julong is a theory that trades with uncertainty cases with the characteristics of less data and incomplete knowledge. 

During years of exploration and growth of grey system it used as a newly emerging scientific correction for its theoretical template framework consisting of systems analysis, modeling, predicting, estimation, and techniques of optimization.

2.4 Combined grey linear regression models\(^{[2][9]}\)

The impairment of the original linear regression model where no exponential growth is taking into account can be amended by using grey models. Also grey models can improve the weakness of GM (1,1) model that involved non-sufficient linear factor. For studying series with both linear propensity and exponential development propensity combined grey linear models is the most appropriate models. For such sequences, characterize modeling process can be used as follows:

Let \(X^{(o)} = \{X^{(o)}(1), X^{(o)}(2), \ldots, X^{(o)}(n)\}\) be a range of data. Its first order generation range \(X^{(1)} = \{X^{(1)}(1), X^{(1)}(2), \ldots, X^{(1)}(n)\}\) . for GM (1,1) model, we can get:

\[\hat{X}^{(1)}(t + 1) = \left(X^{(o)}(1) - \frac{b}{a}\right) \exp(-at) + \frac{b}{a} \quad (2.2)\]

Its form can be written as follows:

\[\hat{X}^{(1)}(t + 1) = C_1 \exp(\nu t) + C_2 \quad (2.3)\]

Then the sum of the linear regression equation \((y=ax+b)\), and the exponential equation \((y=a^\nu \exp(x))\) is used to get the sequence of \(X^{(1)}(t)\):

\[\hat{X}^{(1)}(t + 1) = C_1 \exp(\nu t) + C_2(t) + C_3 \quad (2.4)\]

Where \(\nu, C_1, C_2, C_3\) are the parameters that should be estimated.

For getting the parameters values, we undertake the sequences:

\[Z(t) = \hat{X}^{(1)}(t + 1) - \hat{X}^{(1)}(t)\]

\[= C_1 \exp[\nu(t + 1)] + C_2(t + 1) + C_3 - C_1 \exp(\nu t) - C_2 t - C_3\]

\[= C_1 \exp(\nu t) [[\exp(\nu)] - 1] + C_2 \quad (2.5)\]

\((t= 1,2, 3, \ldots, n-1)\)

Let

\[Y_m(t) = Z(t + m) - Z(t)\]

\[= C_1 \exp[\nu(t + m)][\exp(\nu - 1)] + C_2 - C_1 \exp(\nu t)[\exp(\nu - 1) - 1] - C_2\]

\[= C_1 \exp(\nu t)[\exp(\nu m) - 1][\exp(\nu) - 1] \quad (2.6)\]
\[ Y_m(t + 1) = C_1 \exp[v(t + 1)][\exp(v_m) - 1][\exp(v) - 1] \quad (2.7) \]

The ratio of the previous two equations is

\[ Y_m(t + 1)/Y_m(t) = \exp(v) \quad (2.8) \]

this equation is solved to obtain \( v \):

\[ v = \ln[Y_m(t + 1)/Y_m(t)] \quad (2.9) \]

By substituting the value of \( \bar{X}^{(1)} \) in equation (2.5) by the equation (2.6), we get the estimated values \( \bar{V} \) for \( v \). From various values of \( m \) we get another values of \( \bar{V} \) and then we use the average of these values \( \bar{V} \) to get the approximate value of \( v \).

For equation (2.6) if \( m=1 \) then:

\[ Y_1(t) = Z(t + 1) - Z(t), \quad t=1, 2, \ldots, n-2 \]

\[ \bar{V}_1(t) = \ln[Y_1(t + 1)/Y_1(t)], \quad t=1, 2, \ldots, n-3 \]

For \( m=2 \)

\[ Y_2(t) = Z(t + 2) - Z(t), \quad t=1, 2, \ldots, n-3 \]

\[ \bar{V}_2(t) = \ln[Y_2(t + 1)/Y_2(t)], \quad t=1, 2, \ldots, n-4 \]

For \( m=n-3 \)

\[ Y_{n-3}(t) = Z(t + n - 3) - Z(t), \quad t=1, 2 \]

\[ \bar{V}_{n-3}(t) = \ln[Y_{n-3}(t + 1)/Y_{n-3}(t)], \quad t=1 \]

We calculated many \( \bar{V} \) values \( (n-3)+(n-4)+\ldots+2+1=(n-2)(n-3)/2 \) and get the average of \( \bar{V} \) to estimate the amount of \( v \).

\[ \bar{V} = \frac{\sum_{m=1}^{n-3} \sum_{t=1}^{n-2-m} \bar{V}_m(t)}{(n-2)(n-3)/2} \quad (2.10) \]

Let \( L(t)= \exp(\bar{V}t) \), then equation(2.4) rewired as:

\[ \bar{X}^{(1)}(t) = C_1 L(t) + C_2 t + C_3 \quad (2.11) \]

Using the methods of least squares estimate, we obtain the values of \( C_1, C_2, C_3 \).

\[ X^{(1)} = \begin{bmatrix} x^{(1)}(1) \\ x^{(1)}(2) \\ \vdots \\ x^{(1)}(n) \end{bmatrix}, \quad \bar{C} = \begin{bmatrix} C_1 \\ C_2 \\ C_3 \end{bmatrix}, \quad A = \begin{bmatrix} L(1) & 1 & 1 \\ L(2) & 2 & 1 \\ \vdots & \vdots & \vdots \\ L(n) & n & 1 \end{bmatrix} \]

Then we obtain

\[ X^{(1)} = AC \quad (2.12) \]

DOI: http://dx.doi.org/10.25098/4.2.25
Then we get:
\[ \hat{C} = (A^TA)^{-1}A^TX^{(1)} \]  
\( (2.13) \)

Now, we can achieve the predicted values of the cumulating produced sequence are by:
\[ \hat{X}^{(1)}(t+1) = \hat{C}_1e^{(vt)} + \hat{C}_2(t) + \hat{C}_3 \]  
\( (2.14) \)
\[ \hat{X}^{(0)}(t) = \hat{X}^{(1)}(t+1) - \hat{X}^{(1)}(t) , \ (t = 1,2,\ldots,n-1) \]  
\( (2.15) \)

If \( C_1 = 0 \) then the accumulation stands for a linear regression model, if \( C_2 = 0 \), then the accumulation stands for a GM (1,1) model.

2.5 Posterior Deviation Ratio test (c) \(^{[5,6]} \)

The Posterior deviation ratio test (c) of the grey model is also used as a criterion to judge the predicting capability of the grey model, and (c) is defined as follows:

1) Calculating the average of the original sequence:
\[ \bar{x}^{(0)} = \frac{1}{n}\sum_{i=1}^{m}x^{(0)}(t) \]  
\( (2.16) \)

2) Calculating the original sequence \( (x^{(0)}) \) mean square error:
\[ s_1 = \left( \frac{1}{m} \sum_{i=1}^{m} [x^{(0)}(t) - \bar{x}^{(0)}]^2 \right)^{1/2} \]  
\( (2.17) \)

3) Calculating residual mean:
\[ \bar{e} = \frac{1}{m} \sum_{t=1}^{m} e^{(0)}(t) \]  
\( (2.18) \)

4) Calculating residual mean square error:
\[ s_2 = \left( \frac{1}{m} \sum_{i=1}^{m} [e^{(0)}(t) - \bar{e}]^2 \right)^{1/2} \]  
\( (2.19) \)

5) Calculation variance ratio (c):
\[ c = \frac{s_2}{s_1} \]  
\( (2.20) \)

The corresponding precision rank of different levels of posterior deviation ratio test is organized in table:

Table (2-1) shows the levels of posterior deviation ratio test (c)

<table>
<thead>
<tr>
<th>Precision rank</th>
<th>posterior deviation ratio test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Highly accurate</td>
<td>( c \leq 0.35 )</td>
</tr>
<tr>
<td>Good</td>
<td>( c \leq 0.50 )</td>
</tr>
<tr>
<td>Reasonable</td>
<td>( c \leq 0.65 )</td>
</tr>
<tr>
<td>Inaccurate</td>
<td>( c \geq 0.65 )</td>
</tr>
</tbody>
</table>
3 Applications

3.1 Data of the study

The data of this study has been taken from the world bank site, which is about the yearly gross domestic product in billions Iraqi dinars of Iraq country since the period 2000 till 2018.

3.2 Estimating the parameters of combined model

The parameter V is estimated depending on equations (2.7 and 2.8) which is equaled 0.1009 and the parameter C_1, C_2 and C_3 estimated through equations (2.13) which are equals to (2.9, 0.205 and 1.8569) respectively, the estimated GDP have been computed through equations (2.14), the actual and predicted values of GDP is shown in below table:

Table (3-1): The actual, predicted and error values of combined model

<table>
<thead>
<tr>
<th>Years</th>
<th>GDP</th>
<th>GDP-hat</th>
<th>Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>2000</td>
<td>5.02137</td>
<td>3.74025</td>
<td>1.28112</td>
</tr>
<tr>
<td>2001</td>
<td>4.131457</td>
<td>3.24475</td>
<td>0.886707</td>
</tr>
<tr>
<td>2002</td>
<td>4.102293</td>
<td>2.71761</td>
<td>1.384683</td>
</tr>
<tr>
<td>2003</td>
<td>2.958579</td>
<td>1.71248</td>
<td>1.246099</td>
</tr>
<tr>
<td>2004</td>
<td>5.323536</td>
<td>3.9641</td>
<td>1.359436</td>
</tr>
<tr>
<td>2005</td>
<td>7.35336</td>
<td>5.68561</td>
<td>1.66775</td>
</tr>
<tr>
<td>2006</td>
<td>9.558795</td>
<td>8.29374</td>
<td>1.265055</td>
</tr>
<tr>
<td>2007</td>
<td>11.14558</td>
<td>9.80614</td>
<td>1.33944</td>
</tr>
<tr>
<td>2008</td>
<td>15.70261</td>
<td>14.09421</td>
<td>1.6084</td>
</tr>
<tr>
<td>2009</td>
<td>13.06432</td>
<td>12.072</td>
<td>0.99232</td>
</tr>
<tr>
<td>2010</td>
<td>16.20646</td>
<td>17.54693</td>
<td>-1.34047</td>
</tr>
<tr>
<td>2011</td>
<td>21.73271</td>
<td>22.23851</td>
<td>-0.5058</td>
</tr>
<tr>
<td>2012</td>
<td>25.42255</td>
<td>26.41973</td>
<td>-0.99718</td>
</tr>
<tr>
<td>2013</td>
<td>27.36</td>
<td>28.3831</td>
<td>-1.0231</td>
</tr>
<tr>
<td>2014</td>
<td>27.36</td>
<td>28.68453</td>
<td>-1.32453</td>
</tr>
<tr>
<td>2015</td>
<td>20.72</td>
<td>21.1548</td>
<td>-0.4348</td>
</tr>
<tr>
<td>2016</td>
<td>20.12</td>
<td>21.4268</td>
<td>-1.3068</td>
</tr>
<tr>
<td>2017</td>
<td>22.87</td>
<td>23.936</td>
<td>-1.066</td>
</tr>
<tr>
<td>2018</td>
<td>26.72</td>
<td>28.05813</td>
<td>-1.33813</td>
</tr>
</tbody>
</table>

DOI: [http://dx.doi.org/10.25098/4.2.25](http://dx.doi.org/10.25098/4.2.25)
From the above table it is obvious that the GDP of Iraq country raised linearly and fall as a sin wave damped, the GDP in 2013 and 2014 was recorded the highest amount which is equal to 27.36 for both.

Table (3-2): represents the computed mean square error of the three candidate models

<table>
<thead>
<tr>
<th>Models</th>
<th>GM(1,1)</th>
<th>Simple Regression</th>
<th>Combined</th>
</tr>
</thead>
<tbody>
<tr>
<td>MSE</td>
<td>15.0063</td>
<td>17.0932</td>
<td>1.3165</td>
</tr>
</tbody>
</table>

Up to the above table it is clear that the MSE of the combined model is more less than the MSE of the two others models which is equal to 1.3165

![Figure (1): The scatter plot of the actual and predicted values of GDP](image)

3.3 Testing the model

To test the hypothesis that the model is adequate we used posterior deviation ratio test depending on equation (2.20) the ratio value of (c) computed and it is equal to 0.0473 it implies that the combined model is highly accurate.
4. Conclusions

1. Using this type of model gives more reliable estimated parameters than GM (1,1) and simple linear regression models.
2. The GDP of Iraq in 2003 has recorded the lowest levels which is 2.96 billion because of the American war on Iraq.
3. The GDP of Iraq increased sharply after 2005 but it is not a sign to say that it brings happiness to Iraqi citizens if it does not distribute equally on them.
4. We cannot say that this GDP is the pure GDP of Iraq because there is black market which is not controlled by the government.
5. Some of the products are crashed down or not sell, these products value are not recorded in GDP.
6. According to the 986 decision of republic of Iraq to exchange a part of oil product with food which is called ration card, here also this part of oil is not entered in GDP.

References

2. FAGAN. Doreen. (2019). What is GDP, and why is it important? [online]. Received at: https://www.stlouisfed.org/open-vault/2019/march/what-is-gdp-why-important

DOI: http://dx.doi.org/10.25098/4.2.25