A Restudy of the Impact of Climate on Brazil Based on National Vulnerability Model

Bao Wang*
School of North China Electric Power University, Baoding, China

*Corresponding author e-mail: 3360675177@qq.com

Abstract. Based on research of national vulnerability, Grey prediction and regression analysis are applied to predict the impact of environment change of Brazilian fragility. According to the result, the impact of climate change on Brazilian agricultural production is relatively larger. We define the breakthrough point as that climate change severely undermines national economic stability and predict that Brazil will reach the breakthrough point in 2060 without state intervention measures.

Keyword. Brazil, climate, national vulnerability, Grey prediction, regression analysis.

1. Introduction
Based on big data analysis, most environmental problems are caused by humans. To satisfy their own selfish desires, people are willing to sacrifice the environment at the expense of pesticides, deforestations and so on. As a result, environmental problems are getting worse. For unstable state, they are usually faced with plenty of challenge such as resource shortage, energy pinch, refuges explosion and many other problems. They seek access to resources, energy and everything they need as much as possible, while ignoring the sustainable development, ignoring the harmonious development of man and nature. They rarely accelerate the environmental policies or environmental protection measures, and even if they do, they have no capacity for long-term implementation. Ultimately, long-term accumulations of environmental problems lead to irregular climate changes.

Consequently, based on the national vulnerability model, we use Grey prediction and regression analysis to analyze the impact of climate vulnerability in Brazil, simultaneously, we have predicted the impact of climate on Brazil in the next 30 years.

2. Assumptions
Climate change mainly affects the following: droughts, floods and extreme temperatures, average precipitation depth, nitric oxide emissions, other greenhouse gas emissions, and carbon dioxide emissions.

We believe that climate change will only bring about changes in the following aspects: rainfall, temperature, sea level rise, and natural disaster frequency directly.

3. The GM Model
GM model is a prediction method for predicting grey system which refers to a system we partly know information and uncertain relationship between factors. It can identify the dissimilarity degree of
development trend between the system factors, namely, the method always operates by analyzing correlation information, and achieving the raw data to find the law of the system change, generating strong regularity of data sequences, and then establishing the corresponding differential equation model, to predict the future development trend.

First of all, we make necessary inspection and processing for the known data in order to ensure the feasibility of the GM (1,1) modeling method. We select \( x^{(0)}(1), x^{(0)}(2), \ldots, x^{(0)}(n) \) as the original data and then calculate the class ratio of the sequence.

\[
\lambda(k) = \frac{x^{(0)}(k-1)}{x^{(0)}(k)}, \quad k = 2, 3, \ldots, n
\]

If all the class ratios are falling within the volume range \( X = \left[ e^{-2}, e^{2} \right] \), the sequence \( X^{(0)} \) is used to establish the GM (1, 1) model and could make gray prediction. Otherwise, we should transform the data properly, such as translation transformation.

\[
y^{(0)}(k) = x^{(0)}(k) + c, \quad k = 1, 2, \ldots, n
\]

Where ‘c’ makes the level ratio of the data column falling within the allowable cover.

Secondly, we produce our own GM (1,1) model. Let us say that \( x^{(0)} = (x^{0}(1), x^{0}(2), \ldots, x^{0}(n)) \) satisfies the above requirements and establishes the GM model by analyzing the data column.

\[
x^{0}(k) + az^{(1)}k = b
\]

The estimate of a and b is obtained by regression analysis, and the corresponding albino model is.

\[
\frac{dx^{(1)}(t)}{dt} + ax^{(1)}(t) = b
\]

And then, we get the predicted value as:

\[
x^{(1)}(k+1) = x^{(1)}(k+1) - x^{(1)}(k), \quad k = 1, 2, \ldots, n - 1
\]

Accordingly, the predicted value is obtained:

\[
x^{(0)}(k+1) = x^{(1)}(k+1) - x^{(3)}(k), \quad k = 1, 2, \ldots, n - 1
\]

Thirdly, we check out the predictions. On one hand, we can make residual test by calculating the relative residuals.

\[
\varepsilon(k) = \frac{x^{(0)}(k) - x^{(0)}(k)}{x^{(0)}(k)}, \quad k = 1, 2, \ldots, n
\]

If the formula is true for all \( |\varepsilon(k) < 0.1| \), then the higher requirement is considered; otherwise, if the formula is valid for all \( \varepsilon(k) < 0.2 \), it is considered to meet the general requirements.

On the other hand, we can apply the class ratio deviation test, which is calculated according to the following formula.

\[
\rho(k) = 1 - \frac{1 - 0.5a}{1 + 0.5a} \lambda(k)
\]

If the formula is true for all \( |\rho(k) < 0.1| \), then the higher requirement is considered; otherwise, if the formula is valid for all \( \rho(k) < 0.2 \), it is considered to meet the general requirements.

4. The Multivariable Linear Regression Model

Multiple regression analysis prediction method is a method to predict the prediction model by analyzing the correlation between two or more independent variables and one dependent variable. When there is a linear relationship between the independent variable and the dependent variable, it is called multivariate linear regression analysis.
We set \( y \) as the dependent variable, \( x_1, x_2, x_3, \ldots, x_k \) as the independent variable, and when the independent variable is linear with the dependent variable, the multiple linear regression model is:

\[ y = b_0 + b_1 x_1 + b_2 x_2 + \cdots + b_k x_k + e \]

Where, \( b_0 \) is a constant term, \( b_1, b_2, b_3, \cdots, b_k \) is the regression coefficient, \( b_1 \) is the fixed time of \( x_1, x_2, x_3, \cdots, x_k \), and the effect of \( x_1 \) per unit on \( y \), that is, the partial regression coefficient of \( x_1 \) on \( y \); In the same way that \( b_2 \) is \( x_2 \) and \( x_k \) is fixed, \( x_2 \) increases the effect of each unit on \( y \), that is, the partial regression coefficient of \( x_2 \) to \( y \), and so on.

And then, we perform parameter estimation of multiple linear regression.

It solves the parameter with least square method under the premise that the squared error sum \( \sum e^2 \) is minimum. Taking the two-linear regression model as an example, the standard equations of regression parameters are obtained.

\[
\begin{align*}
\sum y &= nb_0 + b_1 \sum x_1 + b_2 \sum x_2 \\
\sum x_1y &= b_0 \sum x_1 + b_1 \sum x_1^2 + b_2 \sum x_1x_2 \\
\sum x_2y &= b_0 \sum x_2 + b_1 \sum x_1x_2 + b_2 \sum x_2^2
\end{align*}
\]

You can find the value of \( b_1, b_2 \) by solving this equation, and it can also be obtained by the following matrix.

\[
b = (xx)^{-1} \cdot (xy)
\]

\[
\begin{bmatrix}
b_0 \\
b_1 \\
b_2
\end{bmatrix} = 
\begin{bmatrix}
\sum x_1 \\
\sum x_2 \\
\sum x_1x_2
\end{bmatrix}^{-1} \cdot 
\begin{bmatrix}
\sum y \\
\sum x_1y \\
\sum x_2y
\end{bmatrix}
\]

After that, we conduct a multivariate linear regression model test to determine the model’s availability.

We will find multiple linear regression in the \( R^2 \) also have multiple R-squared figures, it is a total change in the dependent variable, the regression equation the explanation of the changes the ratio of sum of squares (regression), \( R^2 \), the greater the return site each of sample sites, the more the degree of fitting all the relationship between the dependent and independent variables and the more closely. The calculation formula is:

\[
R^2 = \frac{\sum(y - \bar{y})^2}{\sum(y - \bar{y})^2} = 1 - \frac{\sum(y - \bar{y})^2}{\sum(y - \bar{y})^2}
\]

Where,

\[
\sum (y - \bar{y})^2 = \sum y^2 - (b_0 \sum y + b_1 \sum x_1 y + b_2 \sum x_2 y + \cdots + b_k \sum x_k y)
\]

\[
\sum (y - \bar{y})^2 = \sum y^2 - \frac{1}{n} (\sum y)^2
\]

We also carried out to estimate standard errors, that is, according to the actual value of the dependent variable \( y \) and the regression equation and the estimated standard errors between the estimated standard error is smaller, the regression equation fitting degree of the process.

\[
S^2 = \frac{\sum(y - \bar{y})^2}{n-k-1}, \quad v_k = \frac{s^2}{y}
\]

Where, \( k \) is the number of independent variables in the multivariate linear regression equation.

Moreover, we text the significance of the regression equation, which means to test the significance of the whole regression equation, or to evaluate whether all the independent variables whether it is
closely related to the dependent variable. F-test can be used frequently. The calculation formula of F-statistic is as follows:

\[ F = \frac{\sum (\hat{y} - \bar{y})^2 / k}{\sum (y - \hat{y})^2 / (n - k - 1)} = \frac{R^2 / k}{1 - R^2 / (n - k - 1)} \]

According to the given significant level \( \alpha \), the degree of freedom \((k,n-k-1)\) check F-distribution table, and obtain the corresponding critical value \( F_a \). If \( F > F_a \), the regression equation has significant meaning and the regression effect is significant. \( F < F_a \), the regression equation is not significant and the regression effect is not significant.

Besides, we test the significance level of regression coefficient by T-test. T-test is a method to test whether the regression coefficients in the regression model are significant, so that only those factors that have significant influence on the dependent variables are retained in the model. The statistics \( t_i \) are calculated at the time of inspection. According to the given significant level and degrees of freedom \( n - k - 1 \) in T-distribution table, which refers threshold \( t_a \) or \( t_{a/2} \), if \( t > t_a \) or \( t_{a/2} \), the regression coefficients \( b_i \) and 0 has significant levels, on the other hand, there was no significant difference with 0. The calculation formula of statistic \( t \) is as follows:

\[ t_i = \frac{b_i}{S_y \sqrt{C_{ij}}} = \frac{b_i}{S_{bi}} \]

Where, \( C_{ij} \) is the \( j \)-th element on the main diagonal of the inverse matrix of the regression coefficient matrix \((x'x)^{-1}\) in the multivariate linear regression equation. Take binary linear regression for example, the following formula can be used to calculate:

\[ C_{11} = \frac{S_{22}}{S_{11}S_{22} - S_{12}^2} \]
\[ C_{22} = \frac{S_{11}}{S_{11}S_{22} - S_{12}^2} \]

Where,

\[ S_{11} = \sum (x_1 - \bar{x}_1)^2 = \sum x_1^2 - \frac{1}{n} (\sum x_1)^2 \]
\[ S_{22} = \sum (x_2 - \bar{x}_2)^2 = \sum x_2^2 - \frac{1}{n} (\sum x_2)^2 \]
\[ S_{12} = \sum (x_1 - \bar{x}_1)(x_2 - \bar{x}_2) = S_{21} \]
\[ \sum x_1 x_2 - \frac{1}{n} (\sum x_1)(\sum x_2) \]

A regression coefficient cannot pass the T-test might be caused by the coefficients corresponding to the independent variable on the dependent variable of flat is not significant, at this time, we should establish a more simple regression model or change the independent variable. It is also possible that collinearities among the independent variables should be managed to reduce the total linear effect. We can calculate determination coefficient \( r_i^2 \) separately between two independent variables. If \( r^2 > R^2 \) or close to \( R^2 \), we should try to reduce the effects of multiple linear. It is also possible to calculate the condition number, which \( k = \tau_1 / \tau_p \) (\( \tau_1 \) is the maximum eigenvalue, and the \( \tau_p \) is the minimum eigenvalue), of the eigenvalues of the correlation coefficient matrix the independent variable. If \( k < 100 \), there is no multi-focus co-linearity. If \( 100 \leq k \leq 1000 \), there is a strong multiple collinear between the independent variables. If \( k > 1000 \), there is a serious multicolor in the independent variable.

When the regression model is established according to the dynamic data, the error term \( e \) is also a time sequence, if the error sequences among the items are independent of each other, there is no correlation between the error sequence. If there is a close relationship between error series, the established regression model could not represent the real changes of expression between independent
variable and dependent variable. D.W test is a self-correlation test of error sequence. The test method is the same as the linear regression.

Finally, we establish the optimal regression equation.

5. The Prediction Results and Analysis

Figure 1. Climate summary

Figure 2. Average high and low temperature.
Figure 3. Cloud cover categories

Figure 4. Humidity comfort levels.

Figure 5. Average monthly rainfall.

We selected the mean of data change in the last 20 years to analyze:
Table 1. Summary of data for the past 20 years.

|     | EW   | MPD  | PCE   | POE   | PDE   |
|-----|------|------|-------|-------|-------|
| IOS | 0.4355 | 0.1761 | 0.0858 | 0.7263 | 0.188 |
| ITS | 0.4355 | 0.6116 | 0.6974 | 1.4237 | 1.6117 |
| MP  | beta\_1=0.9502 | beta\_2=0.3335 |
| AVC | 0.4355 | 0.3118 | 0.2963 | 0.2815 | 0.2675 |
| RESIDUAL |       |       | 0.2669 |       |       |

Through the establishment of the above model, we study the impact of climate on Brazil. First of all, we through the Grey prediction model to predict the change of various other factors caused by climate change in Brazil in the next forty years, then, we use multiple linear regression to predict Brazil's fragility, predict the future five years. According to our results, we found that the negative impact of climate change on Brazil is obvious, the impact of climate change on Brazil's agricultural production is relatively larger. Crop yields and production decline and rising production costs will directly affect the balance between supply and demand of grain in Brazil, on the one hand, will also weaken the competitive advantage in the international agricultural products trade in Brazil, adding to the pressure of the international balance of payments, and even affect the stability of the macroeconomic situation in 2060 (Figure 6).

![Crop yield](image)

Figure 6. The prediction of crop yield.

6. Conclusion
In this paper, we apply the regression analysis method to study the accuracy of the model in the past 20 years to calibrate the accuracy of the model. On this basis, we use the gray prediction method to predict the impact of environmental changes on agriculture in Brazil over the next 30 years. Climate change has a negative impact on crops. Without state intervention, the decline in crop yields and rising production costs will directly affect the balance of food supply and demand in Brazil, and will also reduce the competitive advantage of its agricultural products and affect economic development and stability. Therefore, the Brazilian government needs to increase support for environmental protection to a certain extent and reduce the impact of environmental changes on it.

References
[1] Han Ye. Analysis on the impact of climate change on water circulation and water resources [J]. Private science and technology, 2017(09):194.
[2] Tong Jiang, Xiucang Li, Duqing Qing, Jiashuan Yuan, Erda Lin. The main conclusions and new cognition of climate change 2014: impact, adaptation and vulnerability [J]. Research progress on climate change, 2014,10(03):157-166.
[3] Xiao Kang. Multiple symbiosis: innovation in global governance of sino-us climate cooperation [J]. Social science abstracts, 2016(11):50-52.