Research on the Climate Impact of Adopting GGDP Based on K-means Clustering and Linear Fitting Model

Pengyu He[1a], Qingyu Zhou[2b], Xuan Fang[3c], Tao Liu[4d,*]

{1530763672@qq.com[a], 1609187326@qq.com[b], 1599713033@qq.com[c], liutao@neuq.edu.cn[d,*]}

Sydney Smart Technology College, Northeastern University, Qinhuangdao 066004, China[1,2,3], School of Mathematics and Statistics, Northeastern University at Qinhuangdao, Qinhuangdao 066004, China[4,*]

Abstract. When GDP is used to measure the development prospect of a country, a country may choose production modes that consume resources or pollute the environment to improve GDP, which can be replaced by the calculation of Green GDP. A model is established based on the mitigation of global climate pollution by Green GDP. GGDP per capita of ten countries will be calculated according to the formula, and the GGDP per capita will be analyzed by K-means clustering. In addition, the Silhouette coefficient score will be used as the evaluation standard of clustering effect to obtain two different development stages of countries divided by GGDP. Then GGDP of the United States and China are predicted.

Keywords: GDP; GGDP; K-means clustering; Environment protection; Economy

1 Introduction

In recent years, countries have been paying increasing attention to global environmental issues [1-4]. Problems like greenhouse gas emission [5,6] and loss of forest area [7] have drawn much concern from public. Traditional GDP calculation cannot judge the ecological loss along with the economic development. Thus, the concept of “Green GDP” or “GGDP” has been put forward [8,9]. Not only does it include the calculation of GDP, but it also takes serval environment-related factors into consideration, which gives it the ability of measuring environmental resource loss and indicating the country’s overall performance on environment protection. In this essay the predication of the future value and effect on the global environment will be given, as well the influence after adopting GGDP as the main judging standard.

2 General Assumptions and Notations

The following assumptions are made to simplify the question, see Table 1. A country’s environmental pollution costing can be represented by CO2 emission and waste produced. Adopting GGDP won’t influence environment and economy development in a near future. Other factors (population, national territorial area, etc.) remain predictable.

MSEA 2023, May 26-28, Nanjing, People's Republic of China

DOI 10.4108/eai.26-5-2023.2334462

Table 1. Symbol and description

Symbol	Description
x	The independent variable of the regression equation (GDP and GGDP)
y	The dependent variable of the regression equation (amount of greenhouse gas emission)
θ_0	Linear equation constant terms
θ_1	Coefficient of independent variable of linear equation
$J(\theta_0, \theta_1)$	The cost function of the regression equation
i	The number of data used for fitting
p	The significant correlation value between two variables
m	The total number of data used for fitting
α	Learning rate of regression process

3 Model I: A Method for Calculating GGDP

GDP (Gross Domestic product) refers to the market value of all the final products (goods and services) produced by all economic activities in a country (or region) within a certain period of time (a quarter or a year) [10]. Countries often rely on earth resource and usually tend to gather all available resource, carrying out concentrated development to increase GDP in a short term. Therefore, the Green GDP will be considered with environmental pollution costing and natural resource depletion.

Among these methods [11-13] that can measure GGDP, The method suitable for calculating is the one based on [14], which combines the former experience and the techniques provided by other mature methods like SEEA. The equation is listed as following:

$$Green\ GDP = GDP - (KtCO_2 * PCDM) - (Twaste * 74kWh * Pelect) - (\frac{GNI}{100} * \%NRD) \quad (1)$$

It should be noted that GDP (Gross domestic product) in this equation does not deduct the fabricated assets depreciation or expending of natural resources. Kt CO2 refers to the weight of CO2 emission and $(KtCO_2 * PCDM)$ refers to the total value of CO2 emission, where PCDM represents the average market price of carbon per Kt. Twaste represents the weight (ton) of waste produced and the section $(Twaste * 74kWh * Pelect)$ refers to the total value of waste produced in one year, which turns the value of the waste into the form of the value of electricity. 74kWh is the amount of electricity produced by the waste per ton. GNI represents gross national income and NRD represents natural resource depletion. The last section of the equation refers to the natural resource depletion (including forest, energy and mineral) while the first two sections refer to the environmental pollution costing. This method also provides feasible quantization ways to evaluate the environmental and ecological consumption, which means it's accessible and can get the result more easily.

4 Model II: Global Climate Mitigation Impact Projections

4.1 Data of GGDP in Developed and Developing Countries

Since the world was originally divided into developed countries and developing countries, and the climate environment and natural resources of different continents differ greatly, in order to make the clustering more comprehensive and scientific, different types of countries are selected from different continents to participate in the clustering. Five developing countries (China, Brazil, Russia, Malaysia and Egypt) and five developed countries (USA, UK, Japan, Australia, France) are selected. GGDP per capita of the above ten countries in different years is as Table 2.

Table 2. GGDP per Capita of the Countries

	China	Brazil	Russia	Malaysia	Egypt
2016	6027.557611	7246.403948	5896.635888	5958.675562	2292.058404
2017	6709.769065	8430.133741	7552.555618	6315.454202	1212.461284
2018	7633.496147	7627.915948	6955.942531	6800.015669	1141.859668
2019	7467.926492	7191.425612	6709.827097	6531.587603	1599.742645
	America	England	Japan	Australia	France
2016	44458.54546	27464.30602	24854.14702	22006.22317	31678.07123
2017	45794.09064	27302.35012	22749.53267	23527.27221	32710.64907
2018	48152.15395	12934.48345	-4248.982725	-24154.23226	23545.58356
2019	49722.6367	-1381.179114	-25858.11482	-69758.37804	12649.17578

4.2 The Establishment of GGDP Impact Clustering Model

The above variables (GGDP per capita) will be used as development characteristics for K-means clustering of ten countries. In selecting the number of categories, the Silhouette score will be used as the evaluation standard for clustering effect [15]. When the number of categories ranges from 2 to 4, the corresponding clustering indexes of ten countries are respectively (0.465691, 0.4790881, 0.4807911). The indexes reach the maximum when the number of categories is 2. Therefore, two categories are obtained, and representative countries will be selected to predict the impact of GGDP.

4.3 Qualitative Analysis of Climate Change

The change of greenhouse gas emissions can well reflect the effect of replacing GDP with GGDP on mitigating climate change. The greenhouse gas emission data of the United States from 2001 to 2020 under the influence of GDP and GGDP as development indicators were intercepted as Fig. 1 to analyze the impact(In which GDP is in trillions, greenhouse gas emission is in hundred million tons). It can be found that with the changes of GDP and GGDP, the overall trend of greenhouse gas values in the United States both declines linearly, but the reduction rate of decline is significantly increased. It can be interpreted that the change has promoted the development plan of the United States to a more sustainable direction and has positive significance for the mitigation of climate pollution.

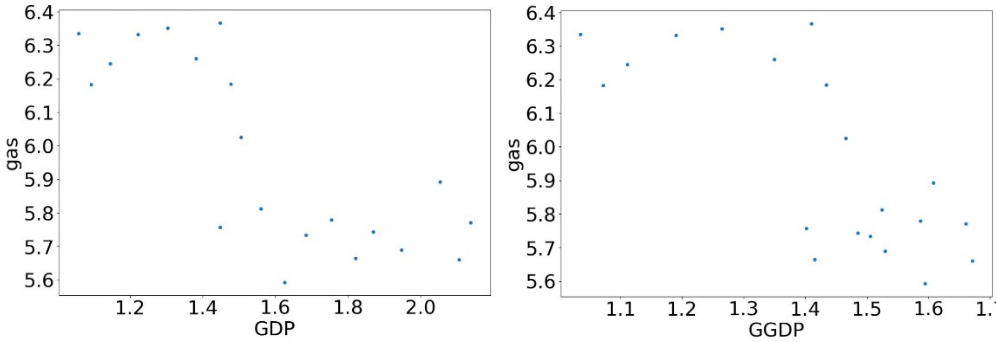

Fig. 1. US greenhouse gas emissions trend along with GDP and GGDP

4.4 The Impact of GGDP on Climate

Liear regression model is used to discuss the impact of GDP growth on greenhouse gas emissions. Set GDP value as x(trillion) and the amount of greenhouse gas emissions as y(hundred million tons). Then let $h(x) = \theta_0 + \theta_1 * x$. The cost function is:

$$J(\theta_0, \theta_1) = \frac{1}{2m} \sum_{i=1}^{m} (h_\theta(x^{(i)}) - y^{(i)})^2 \tag{2}$$

Use result matrix to get the result of $h_\theta(x^{(i)}) - y^{(i)}$:

$$\begin{pmatrix} x^{(1)}{}_1 & x^{(1)}{}_2 \\ \vdots & \vdots \\ x^{(m)}{}_1 & x^{(m)}{}_2 \end{pmatrix} * \begin{pmatrix} \theta_0 \\ \theta_1 \end{pmatrix} - \begin{pmatrix} y^{(1)}{}_1 \\ \vdots \\ y^{(m)}{}_1 \end{pmatrix} \tag{3}$$

Then get the change value of θ_0 and θ_1 :

$$\theta_j := \theta_j - \alpha \frac{\partial}{\partial \theta_j} J(\theta_0, \theta_1) \tag{4}$$

In which α is the learning rate. In order to make the final fitting result better, the value of α (0.01~1) is selected and set to be 0.03 while the number of times of gradient descent is 10000. Since the cost function is too high at the beginning, only the last 9985 times of the function is intercepted. Then it can show the outcome of linear function obtained from US GDP/GGDP data and greenhouse gas emission data from 2001 to 2020 and the effect when GDP/GGDP increases as Fig. 2. It can be concluded that the rate of decline of greenhouse gas emissions is significantly increased, and climate pollution in the United States can be alleviated which means developed countries would pay more attention to climate pollution prevention after GDP is replaced by GGDP.

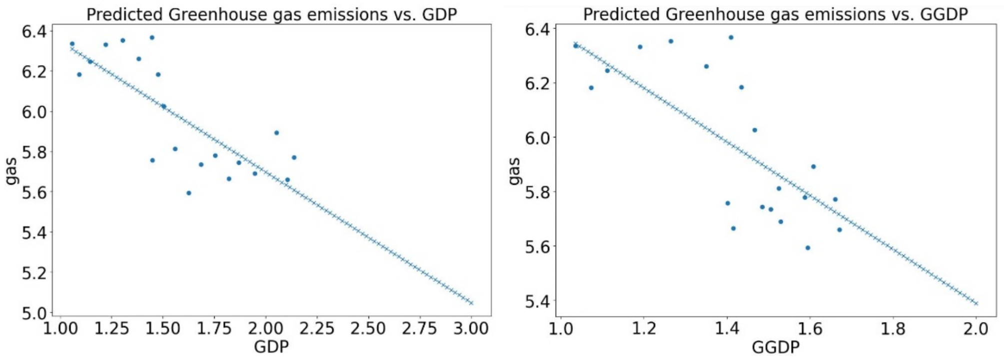

Fig. 2. US GDP/GGDP and greenhouse gas emissions (slope: -0.65/-0.99)

For developing countries, China will be the representative of the discussion. The outcome of linear function obtained from China's GDP/GGDP data and greenhouse gas emission data from 2001 to 2020 and the effect when GDP/GGDP increases as Fig. 3. It is found that in the early stage the value of GGDP is quite small, but it has increased by 59.85 times in the past 20 years, while the GDP has only increased by 10 times in the past 20 years. At the same time, despite slowing down, the increase rate of greenhouse gas emissions accelerates. Because China and the United States are in different stages of development, the United States has preliminarily shifted to the tertiary industry and high-tech industry, while China is just transforming to that. GDP is an important factor in the calculation of GGDP, so it means that industries with high emissions but higher output will also increase the GGDP. Thus, it can be seen that developed countries can adopt GGDP for development direction restrictions, but developing countries still need to have more discussions.

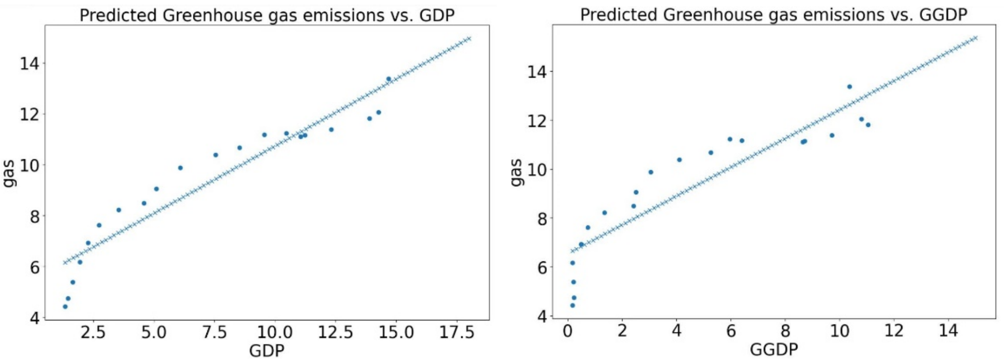

Fig. 3. China GDP/GGDP and greenhouse gas emissions (slope: 0.53/0.59)

It is estimated that after the adoption of GGDP, the reduction rate of greenhouse gas emissions in developed countries will be 1.53 times of the original value, while the rate in developing countries will only be 1.12 times. The GDP per capita of developed countries is about 10 times that of developing countries, and the production activities of developed countries occupy more than 50% part of the global production activities. Therefore, it can be inferred that global climate pollution can be effectively alleviated after the adoption of GGDP.

5 Conclusion

The study chose a most effective way to calculate the GGDP and ten countries around the world are divided by GGDP per capita using K-means clustering while two most representative countries America and China are chosen for deeper analysis using linear fitting method. It can be seen that after adopting GGDP the amount of greenhouse gas emission declines faster, which indicates that adopting GGDP can effectively ameliorate environmental problem. Based on the analysis, the further actions of adapting GGDP are recommanded. It would be considered healthy for both the whole nation's development and whole planet's ecosystem.

References

[1] O'Brien, K.: Global environmental change II: From adaptation to deliberate transformation. Progress in Human Geography. Vol. 36, pp. 667-676 (2012)

[2] Liu, T.: Porosity reconstruction based on Biot elastic model of porous media by homotopy perturbation method. Chaos, Solitons & Fractals. Vol. 158, pp. 112007 (2022)

[3] Moser, K.A., Baron, J.S., Brahney, J., et al.: Mountain lakes: Eyes on global environmental change. Global and Planetary Change. Vol. 178, pp. 77-95 (2019)

[4] Liu, T.: Parameter estimation with the multigrid-homotopy method for a nonlinear diffusion equation. Journal of Computational and Applied Mathematics. Vol. 413, pp. 114393 (2022)

[5] Meinshausen, M., Meinshausen, N., Hare, W., et al.: Greenhouse-gas emission targets for limiting global warming to 2 C. Nature. Vol. 458, pp. 1158-1162 (2009)

[6] Gillingham, K., Stock, J.H.: The cost of reducing greenhouse gas emissions. Journal of Economic Perspectives. Vol. 32, pp. 53-72 (2018)

[7] McAlpine, C.A., Johnson, A., Salazar, A., et al.: Forest loss and Borneo's climate. Environmental Research Letters. Vol. 13, pp. 044009 (2018)

[8] Talberth, J., Bohara, A.K.: Economic openness and green GDP. Ecological Economics. Vol. 58, pp. 743-758 (2006)

[9] Boyd, J.: Nonmarket benefits of nature: What should be counted in green GDP?. Ecological Economics. Vol. 61, pp. 716-723 (2007)

[10] Xu, L., Yu, B., Yue, W.: A method of green GDP accounting based on eco-service and a case study of Wuyishan, China. Procedia Environmental Sciences. Vol. 2, pp. 1865-1872 (2010)

[11] Liu, T.: A wavelet multiscale-homotopy method for the parameter identification problem of partial differential equations. Computers & Mathematics with Applications. Vol. 71, pp. 1519-1523 (2016)

[12] Liu, T., Yu, J., Zheng, Y., et al.: A nonlinear multigrid method for the parameter identification problem of partial differential equations with constraints. Mathematics. Vol. 10, pp. 2938 (2022)

[13] Liu, T., Xia, K., Zheng, Y., et al.: A homotopy method for the constrained inverse problem in the multiphase porous media flow. Processes. Vol. 10, pp. 1143 (2022)

[14] Stjepanović, S., Tomić, D., Škare, M.: Green GDP: An analyses for developing and developed countries. E+M. Ekonomie a Management. Vol. 22, pp. 4-17 (2019)

[15] Hamerly, G., Elkan, C.: Learning the k in k-means. Advances in Neural Information Processing Systems. Vol. 16, pp. 1-8 (2003)