# Climate Impact and Feasibility Analysis of Global Green GDP Based on Improved Weight Method and K-Means Method

Linyang Tian [1, #], Yunqiang Sun [2, *, #], Jiaji Chi [2]

[1] LUT School of Engineering Science, Lappeenranta-Lahti University of Technology, Lahti, Finland, 15240

[2] LUT School of Energy Systems, Lappeenranta-Lahti University of Technology, Lappeenranta, Finland, 53850

* Corresponding Author Email: Yunqiang.Sun@student.lut.fi

[#]These authors contributed equally

**Abstract.** Gross domestic product (GDP) is a widely used economic indicator, but it does not consider economy related to natural resources. It may not accurately reflect a country's economic well-being. In this study, the potential of using "green GDP" (GGDP) and the GGDP index as primary indicators of economic health with a focus on climate implications is explored. The article proposes a new GGDP model and index that incorporate three key factors including environment, economy, and society with four quantifiable indicators within each factor. The weights for the indicators are calculated based on improved weight method containing entropy weight method and CRITIC method. GGDP model places emphasis on climate mitigation and employs grey correlation analysis and univariate linear regression. Furthermore, the GGDPGI model has been developed to assess the feasibility of replacing GDP with GGDP as the primary measure of economic growth. Our results suggest that using GGDP and the GGDP index as primary indicators is not only feasible but also meaningful for evaluating a country's economic growth.

**Keywords:** Entropy weight method, CRITIC Method, K-means method, GGDP, GGDP index.

## 1. Introduction

### 1.1. Background

Gross Domestic Product (GDP) is the final product of production activities of all the resident units in a country (or region) in a certain period of time. It is one of the most well-known and commonly used measures of the health of a nation's economy. At the moment, a lot of countries boost their GDP through the exploitation of oil, natural gas and coal resources. Some countries of having rich forests or mineral resource boost their GDP by cutting trees or exploiting mineral resources without abstinence[1]. The phenomenon leads to the loss of biodiversity, air and water pollution, land desertification and other negative environmental impacts. This can not make countries and regions develop sustainably[2].

However, there is a new measurement of the health of a nation's economy called Green GDP (GGDP). GGDP[3] is an extension and improvement of GDP, which can take environmental and sustainability perspectives and factors into consideration. GGDP may be a better way to measure the health of a nation's economy[1]. Therefore it is important to establish an evaluation model to decide which method is better in order to make a nation's economy boost in a healthier way.

### 1.2. Our Work

Our study aims to explore the feasibility of using global GGDP as a primary indicator for measuring national economic health. To achieve this, we conducted the following steps:

Firstly, we prepared the models and collected data to select 30 countries and indicators.

Secondly, we used the improved weight method to calculate weights.

Thirdly, we established a GGDP model based on the results and measurable indicators such as global temperature and forest area. Then we analyzed the impact of GGDP on climate mitigation using the grey analysis method and univariate linear regression.

Next, we used the K-means clustering method to group the 30 selected countries into three categories based on their GGDP index. We then utilized the grey prediction model to fit and estimate the GGDP index variation from 1995 to 2025 in selected European countries. Additionally, we predicted the potential impact of GGDP on the economy and examined the correlation between the GGDP index and the sum of GDP and the average PPP for each country category.

Then, we analyzed the feasibility of the models and performed sensitivity analysis to demonstrate their stability.

### 1.3. Arrangement

**Section I** involves the preparation of the models, we make assumptions and define the notations in the chapter. **Section II** is methodology, explaining the method of establishing models. **Section III**, we arrange experiment as the third part of our paper to interpret the process of selecting indicators and countries. **Section IV** displays GGDP model and GGDPGI model.

## 2. Methodology

### 2.1. Weights

#### 2.1.1 Entropy Weight Method

We used the EWM to determine index weights objectively based on the information contained in each indicator's value[4]. A smaller entropy indicates greater variation in the index value and more information, resulting in a higher weight in the comprehensive evaluation[5].

It is necessary to first convert them into relative values through standardization, as the units of measurement for various indicators are not standardized.

Let:

$$x_{ij}=|x_{ij}| \tag{1}$$

Then we solve the homogenization problem of different quality index values. we use different algorithms for the data in the high and low indicators.

For positive indicators:

$$x'_{ij}=\left[\frac{x_{ij}-\min(x_{1j},x_{2j},\cdots,x_{nj})}{\max(x_{1j},\cdots,x_{nj})-\min(x_{1j},\cdots,x_{nj})}\right] \tag{2}$$

For negative indicator:

$$x'_{ij}=\left[\frac{\max(x_{1j},x_{2j},\cdots,x_{nj})-x_{ij}}{\max(x_{1j},\cdots,x_{nj})-\min(x_{1j},\cdots,x_{nj})}\right] \tag{3}$$

Calculate the proportion of characteristics of each evaluation object under item $j$. Set the standardized data of the $j^{th}$observation of the $i^{th}$ evaluation object, and calculate the feature weight of the $i^{th}$evaluation object under the $j^{th}$ indicator:

$$p_{ij}=\frac{x_{ij}}{\sum_{i=1}^{n} x_{ij}},\ (i=1,2,\cdots,n,) \tag{4}$$

Calculate the weights and information entropy of each indicator, the information entropy of data is defined to be $E_j = -\ln\frac{1}{n}\sum_{i=1}^{n} p_{ij}\ln p_{ij}$ $(E_j \geq 0)$

$$w_j=\frac{d_j}{\sum_{j=1}^{m} d_j},\ j=1,...,m \tag{5}$$

$d_j$ equals to 1-$E_j$ .

### 2.1.2 Improved Weight Method

The new perspective of CRITIC(Criteria Importance Through Inter criteria Correlation) weighting method[6-7] is to use two indicators, the comparative strength and the conflict indicator. The contrast strength is expressed using standard deviation, and if the data standard deviation is larger, the fluctuation is larger and the weight will be higher[7].

The CRITIC method calculates the weight of indicators around two aspects: contrast and contradiction.

Information Capacity. Assume indicator i, information capacity $C_j$, σ means the standard deviation and $f_j$ magnitude of the contradiction between indicator j and the rest of the indicators.

$$C_j=\sigma_j \cdot f_j \tag{6}$$

Calculate Weights and Index

$$w_j'=\frac{C_j}{\sum_{j=1}^{n} C_j} \tag{7}$$

Differences between EWM and CRITIC are stability and flexibility. CRITIC is sensitive to data instability, while EWM is relatively robust and stable. However, CRITIC can account for nonlinear relationships and interactions between criteria. We calculate the average from these two models to obtain a more accurate value [7].

$$W_j=\frac{w_j+w_j'}{2} \tag{8}$$

### 2.2. GGDP Index and GGDP

We calculate the comprehensive index of each country.

$$s_i=\sum_{j=1}^{m} W_j \cdot p_{ij} \tag{9}$$

$x_{ij}$ represents the value of the corresponding indicator. $m$ means the amount of countries, which equals to 30. Then, we get the value of GGDP of 30 countries.

$$GGDP=s_i \times GDP \tag{10}$$

### 2.3. Grey Correlation Analysis

We used univariate linear regression model and grey correlation analysis to prove the measurable impact of the global green GDP model. Gray correlation analysis[8] reflects the degree of correlation between curves.We use $Z$-score normalization to standardize them. The specific formula is:

$$x' = \frac{x - \mu}{\delta} \tag{11}$$

$\mu$ means the original data, $\delta$ is Standard deviation of the original data. The grey correlation analysis formula is used to calculate the correlation between each factor of each year and the total amount of global GGDP.

$$\varepsilon_i(k) = \frac{min_i min_k |x_0(k)-x_i(k)| + \rho * max_i max_k |x_0(k)-x_i(k)|}{|x_0(k)-x_i(k)| + \rho * max_i max_k |x_0(k)-x_i(k)|} \tag{12}$$

We assume that the comprehensive grey correlation degree(CGCD) is $y(x_0, x_i)$, then

$$y(x_0, x_i) = \frac{1}{n} \sum_{k=1}^{n} y(x_0(k), x_i(k)) \tag{13}$$

### 2.4. Univariate Linear Regression

We perform a one-dimensional linear regression analysis[9] on forest cover area as an example to further explore the influence of GGDP and climate in depth.

We can accomplish the following task first assume such a univariate linear equation if you have *n* independent observations $(x_1,y_1)(x_2,y_2),\cdots,(x_n,y_n)$

$$\begin{cases} y_i = \beta_0+\beta x_1+\varepsilon_i,\ i=1,2,...,n \\ E\varepsilon_i=0, D\varepsilon_i=\sigma^2 \text{ and } \varepsilon_1\varepsilon_2,...,\varepsilon_n \end{cases} \tag{14}$$

We can solve from the least squares method:

$$\begin{cases} \hat{\beta}_0 = \bar{y}-\hat{\beta}_1\bar{x} \\ \hat{\beta}_1 = \dfrac{\overline{xy}-\bar{x}\bar{y}}{\overline{x^2}-\bar{x}^2} \end{cases} \qquad \hat{\beta}_1 = \frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\sum_{i=1}^{n}(x_i-\bar{x})^2} \tag{15}$$

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n}x_i,\ \bar{y} = \frac{1}{n}\sum_{i=1}^{n}y_i,\ \bar{x}^2 = \frac{1}{n}\sum_{i=1}^{n}x^2{}_i,\ \overline{xy} = \frac{1}{n}\sum_{i=1}^{n}\overline{xy}_i \tag{16}$$

For regression equation Significance testing of $Y = \beta_0 + \beta_1 x$ boils down to the hypothesis

$$H_0 : \beta_1 = 0; H_1 : \beta_1 \neq 0 \tag{17}$$

### 2.5. K-means Method

After we obtain GDP index, we can divide it into countries with low, medium, and high GDP index using the K-means method [10].

Assigning object points to the cluster closest to the cluster center requires a nearest-neighbor measurement strategy. We use the Euclidean distance considering the characteristics of the model, which can be expressed as follows:

$$d(x,y)=\sqrt{(x_1-y_1)^2+(x_2-y_2)^2+\cdots+(x_n-y_n)^2}=\sqrt{\sum_{i=1}^{n}(x_i-y_i)^2} \tag{18}$$

#### 2.5.1 The calculation of the new clusters

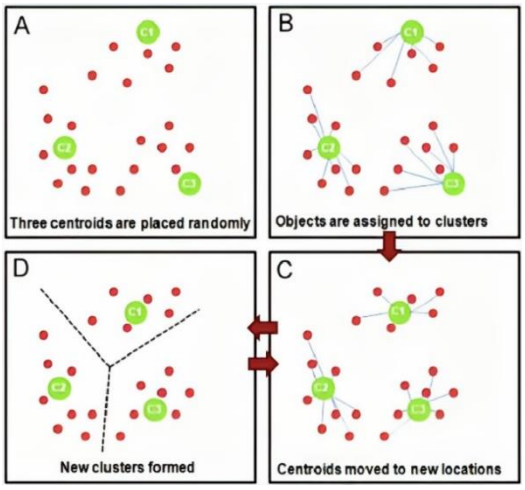



**Figure 1.** Process of K-means algorithm

The distance to the point with the minimum average distance within the cluster is calculated to determine the centre for each of the *k* clusters generated after classification. When the centre of each cluster no longer changes, the *K*-means algorithm can be stopped. The process can be seen in Figure 1.

### 2.6. Grey Prediction Model

We selected gray prediction model GM(1,1)[11,12] as our analysis method to measure the change of the selected countries in GGDP Index.

$$d^1(n)=d^0(1)+d^0(2)+\cdots+d^0(n) \tag{19}$$

$$d^1=(d^1(1),d^1(2),\cdots,d^1(n)) \tag{20}$$

where $d^0(n)$ are original data, and $d^1$ are obtained by accumulating $d^0(n)$

$$z^0(n)=ad^0(n)+(1-a)d^0(n-1) \tag{21}$$

The sequence obtained in this way is called a Proximate Sequence, and the parameter a is also called the Proximate Coefficient.

After the processing, we have the initialized data to establish GM(1,1) Model, which means the the first-order differential equation.

$$\gamma(k)=\frac{d^0(k-1)}{d^0(k)} \quad k=2,3\cdots,n \tag{22}$$

We can obtain the first-order differential equation.

$m(k)+az^1(k)=b$

Where a is called the development coefficient, c is called the background value, and b is called the gray action in gray system theory.Whitening Model is shown in formula (23)

$$\frac{dd^1(t)}{dt}+ad^1(t)=b \tag{23}$$

Ratio-to-difference can be calculated as well.

$$C=\frac{S_2}{S_1} \tag{24}$$

Where $C$ is Ratio-to-difference and $S_1$ is variance of sequence $d^1(k)$, $S_2$ is the variance of residual of sequence $d^1(k)$. Model Accuracy Class is rather precise if $C$ is smaller than 0.35.

## 3. Experiment

### 3.1. Factors and Indicators

We create the GGDP model which analyzes multiple green indicators to obtain the global average GGDP index to measure a country's economic development[13]. The model comprises three primary indicators and twelve secondary indicators, with each primary indicator including four secondary indicators. These indicators consist of five positive and seven negative indicators, reflecting that a country's economic health is determined not only by green sustainable development but also by indicators related to resource consumption and environmental damage. GGDP factors and indicators displays the selected indicators in Figure 2.

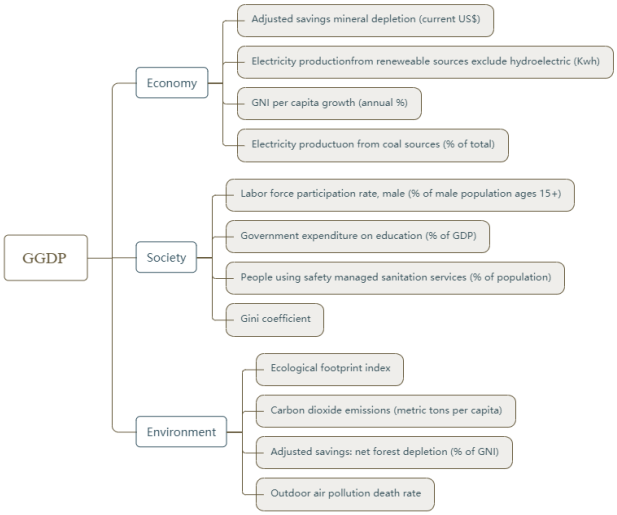



**Figure 2.** Selected factors and indicators

**Economy( Eco)** We selected consumption of mineral resources (Eco1) and electricity production from coal (Eco4) as secondary indicators to measure the impact of non- renewable resource consumption on the green economy[14]. Fossil fuels have been the primary source of energy, but their limited supply leads to increased energy prices, scarce resources, and increased greenhouse gas emissions. Renewable energy generation (Eco2) can promote economic and social development while conserving resources and protecting the environment[15]. The relationship between GNI growth rate(Eco3) and the green economy is interactive. Achieving economic growth in a green way can increase sustainability and achieve long-term economic and social development goals.

**Society(Sco)** We divide the factor into four secondary indicators. The employment rate (Soc1) is an indicator of an economy's ability to create job opportunities. Investment in education (Sco2) is an investment in human resources, and the degree of investment determines the degree of income it generates. The GINI coefficient[16] (Soc4) measures the degree of income or wealth inequality in a society, with a range of values from 0 to 1. The level of healthcare is closely related to a country's economic development, and the proportion of the population using healthcare services with safe management (Sco3) assesses whether healthcare institutions have taken sufficient measures to ensure the safety and quality of healthcare services.

**Environment(Env)** Environment is chosen as another important indicator to measure GGDP[17]. The Ecological Footprint Index(Env1) reflects the impact of human activities on ecosystems and assesses the impact of human activities on Earth's resources and the sustainability of ecosystems. Per capita carbon dioxide emission(Env2) is used to measure a country's resource and environmental development and protection. Considering the importance of forest resources in global climate change and economic development, we made net forest depletion as the third indicator(Env3). In addition, outdoor air pollution mortality rate is selected as a secondary indicator(Env4) to measure the impact of a country's industrial and economic development on people's health. The purpose of all these indicators is to measure the impact of economic development on the environment and human health, and to promote the sustainable development of the economy.

### 3.2. Country selection

We select thirty countries as samples to establish the model. The results are shown in Table 1. The sign ‘*’ beside country names are developed countries .

**Table 1.** 30 countries selection

| | |
|---|---|
| Asia | China Japan* India Saudi Arabia Philippines Pakistan Vietnam Israel* Singapore* |
| Europe | Russia United Kingdom* Turkey France* Germany* Romania Poland* |
| North America | United States* Canada* Mexico |
| South America | Brazil Argentina Colombia Peru |
| Africa | South Africa Egypt Sudan Nigeria Morocco Kenya |
| Oceania | Australia* New Zealand* |

### 3.3. Data collection

We obtain relevant indicator data from the *World Bank* and *Our World in Data*. However, we use the latest available data for these factors from the time period of 2019-2021 as an approximation for 2020 data due to incomplete data on some countries' factors in 2021. Additionally, we evaluate 30 representative countries based on continent and development level, using data from 1995 to 2020, with each five-year interval serving as a time node for recording information.

## 4. Result

### 4.1. Results of Weight

We get the following results of weights of each indicators and factors in Figure 3, Figure 4, Figure 5 and Figure 6.

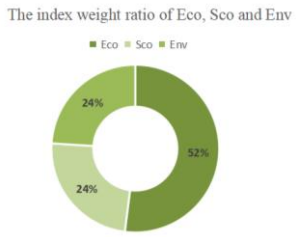



**Figure 3.** Weights of factors

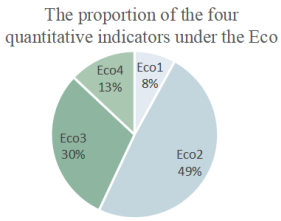



**Figure 4.** Weights of Eco

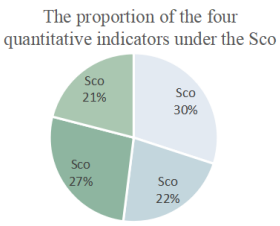



**Figure 5.** Weights of Sco

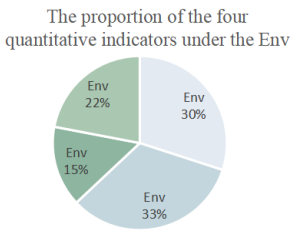



**Figure 6.** Weights of Env

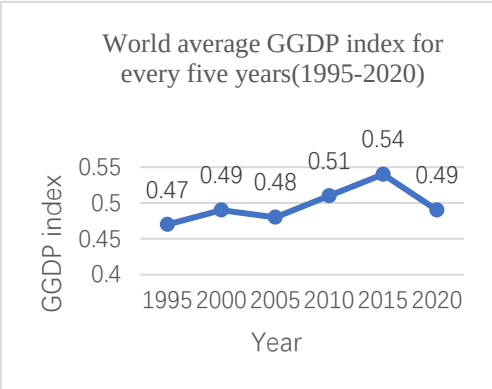



**Figure 7.** GGDP index variation

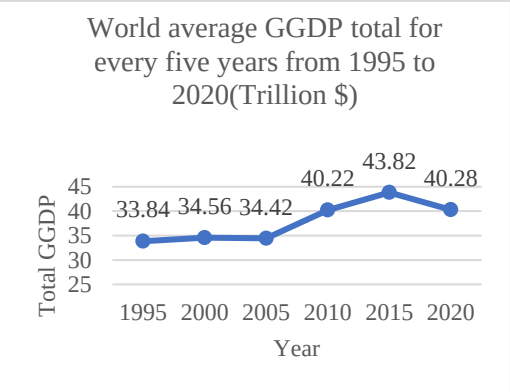



**Figure 8.** Average GGDP variation

### 4.2. GGDP Model

We obtain the results of GGDP index and total GGDP in Figure 7, Figure 8.

We obtained correlation coefficients between total green GDP and forest area, greenhouse gas emissions as well as global average temperature data respectively for each five-year period from 1995-2020.

Then, we get grey correlation degree(CGCD) in Table 2.

**Table 2.** grey correlation degree

| | Forest area | Greenhouse gas emissions | Global mean temperature |
|---|---|---|---|
| CGCD | 0.707 | 0.681 | 0.795 |

We can draw some preliminary conclusions as follows. There are high correlations between GGDP and global average temperature, forest cover, and greenhouse gas emissions with correlations of 0.795, 0.707 and 0.681 respectively. This indicating that there are strong correlations and positive relationships between GGDP and global average temperature. The correlation between GGDP and global average temperature is strong and positive.

After that we use the method of univariate linear Regression and get the results of coefficients and regression line. The results are shown in Table 3.

**Table 3.** Coefficients of univariate linear regression equation

| $\hat{\beta}_0$ | $\hat{\beta}_1$ | $\hat{\beta}_0$confidence interval | $\hat{\beta}_1$ confidence interval | $r^2$ | $F$ | $p$ |
|---|---|---|---|---|---|---|
| 1.1373 | 0.0474 | (-2.9003,0.6257) | (0.0010,0.0937) | 0.6684 | 8.0618 | 0.0469 |

$$y=1.1373x+0.0474 \tag{25}$$

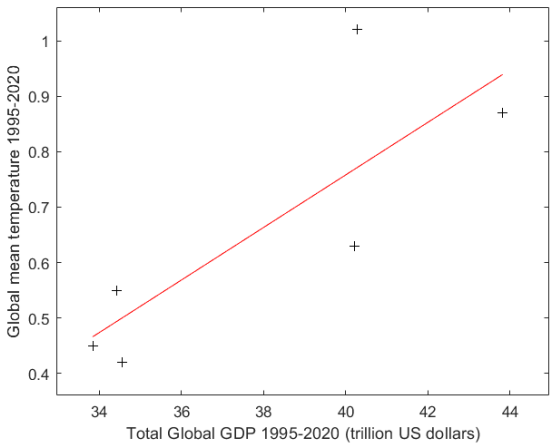



**Figure 9.** Univariate linear regression line

The equation (25) and Figure 9 indicate that the growth of GGDP will have some positive impact on forest cover area, i.e., GGDP growth may promote the increase of forest cover area. An increase in forest cover can help mitigate climate change by absorbing $CO_2$. Therefore, the growth of GGDP may have some positive impact on climate change mitigation.

### 4.3. GGDPGI Model

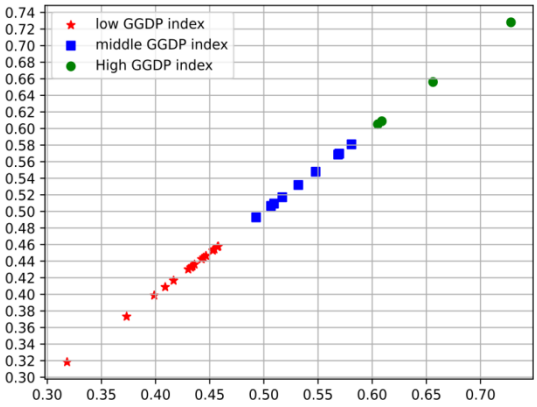



**Figure 10.** Results of K-means

As depicted in Figure 10, we have categorized the 30 countries into three groups based on their level of GGDP index. The green circle, blue square, and red pentagon respectively represent countries with high, medium, and low GGDP index. We have conducted a brief statistical analysis to better understand the distribution of GGDP and to gain insight into the current global equity situation. The statistical data for all categories is presented in Table 4. Furthermore, we have selected six European countries - including the high GGDP index countries of the United Kingdom and Germany, the medium GGDP index country of France, and the low GGDP index countries of Romania, Poland, and Russia to estimate the variation of GGDP index from 1995 to 2025. The results are shown in Figure 11.

**Table 4.** Categories' information based on K-means

| Index level | High | Middle | Low |
|---|---|---|---|
| Number of countries | 4 | 8 | 18 |
| Maximum GGDP Index | 7.28098 | 0.580918 | 0.492887 |
| Minimum GGDP Index | 0.605189 | 0.506583 | 0.318340 |
| Average GGDP Index | 0.649572 | 0.54148625 | 0.42949461 |

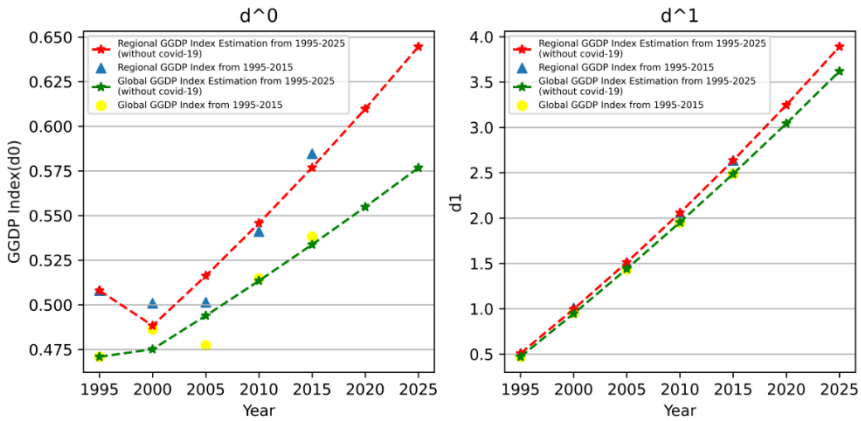



**Figure 11.** Variation and estimation of GGDP index in selected countries

We conducted a data analysis to explore the viability of using GGDP as a substitute for GDP. We employed the *K*-Means method of cluster analysis to classify countries into three groups based on

their GGDP values and utilized the Grey Prediction model to predict the economic conditions of each group. We selected two indicators, the total GDP and the average purchasing power parity (PPP), to measure the economic status of the various country groups. If GGDP were to replace GDP, the impact on countries with different GGDP values would differ. Our data analysis provides insights into how different countries may be affected based on their GGDP values. We performed a level ratio test of the sequence, and all level ratios fell within the admissible coverage interval, indicating the accuracy of our modeling with GM(1,1). All data were sourced from the public data of the World Bank. Figure 12 illustrates the forecast data of the total GDP and purchasing power of the countries in 2025, 2030, and 2035, with each indicator representing the country with the GGDP index from low to high, respectively, from left to right.

**Table 5.** The predicted value of the sum of countries' GDP and Purchasing Power Parity

| Year | The sum of countries' GDP (trillions of dollars) | | | Purchasing Power Parity (trillion International dollars) | | |
|---|---|---|---|---|---|---|
| Index level | Low | Middle | High | Low | Middle | High |
| 2025 | 34.29 | 40.10 | 46.97 | 10.03 | 12.27 | 15.08 |
| 2030 | 34.46 | 43.81 | 55.82 | 6.42 | 8.29 | 10.80 |
| 2035 | 17.93 | 24.27 | 32.66 | 2.36 | 3.16 | 4.31 |

Based on Table 5, we can see that the sum of green GDP and the average purchasing power parity (PPP) for countries in each of the three categories with low, medium, and high GGDP values are continually increasing. This suggests a positive trend for the future global economic development, despite the initial state of the green economy. We observe that in some cases, the sum of DFP for countries with medium GGDP values exceeded that of countries with high GGDP values after 2025.

## 5. Conclusion

In this paper, the article concentrates on a novel method for calculating Green Gross Domestic Product (GGDP) and develops a model to assess its global climate impact. Our approach synthesizes the advantages and limitations of previously proposed GGDP models, offering a more comprehensive perspective on the subject.

The improved weight method is used to effectively weigh the various factors contributing to GGDP, resulting in a more accurate and robust calculation. Furthermore, the article analyzes the measurable impact of GGDP on global climate using grey relational analysis and univariate linear regression. The article demonstrates a strong correlation between GGDP and global climate and the positive influence of GGDP on climate change. Then, the K-means method is selected to categorize 30 typical countries into three groups based on their GGDP index, considering the three factors of environment, society, and economy. Each factor contains four secondary indicators to measure the impact of GGDP on the world.

Our study presents a novel method for calculating GGDP and a model for evaluating its global climate impact. The correlation between GGDP and global climate, as well as the positive influence of GGDP on climate change and the global economy are found in the article. Our results can inform policy decisions and promote sustainable development.

## References

[1] Talberth, J.; Bohara, A.K. 2006. Economic openness and green GDP, Ecological Economics 58(4): 743-758.

[2] Stjepanović S, Tomić D, Škare M. Green GDP: An analyses for developing and developed countries [J]. 2019.

[3] Vaghefi N, Siwar C, Aziz S. Green GDP and sustainable development in Malaysia[J]. Current World Environment, 2015, 10(1):1.

[4] K.S. Rodgers, M. H. Kido, P. L. Jokiel, T. Edmonds, and E. K. Brown, "Use of integrated landscape indicators to evaluate the health of linked watersheds and coral reef environments in the Hawaiian Islands," Environmental Management, vol. 50, no. 1, pp. 21–30, 2012.9

[5] Sahoo M M, Patra K C, Swain J B, et al. Evaluation of water quality with application of Bayes' rule and entropy weight method[J]. European Journal of Environmental and Civil Engineering, 2017, 21(6): 730-752.

[6] M. P. Amiri, "Project selection for oil-fields development by using the AHP and fuzzy TOPSIS methods," Expert Systems with Applications, vol. 37, no. 9, pp. 6218–6224, 2010.

[7] Alinezhad A, Khalili J, Alinezhad A, et al. CRITIC method[J]. New Methods and Applications in Multiple Attribute Decision Making (MADM), 2019: 199-203.

[8] Tsai M S, Hsu F Y. Application of grey correlation analysis in evolutionary programming for distribution system feeder reconfiguration [J]. IEEE Transactions on Power Systems, 2009, 25(2): 1126-1133.

[9] Jiang W. The data analysis of Shanghai air quality index based on linear regression analysis [C]//Journal of Physics: Conference Series. IOP Publishing, 2021, 1813(1): 012031.

[10] Likas A, Vlassis N, Verbeek J J. The global k-means clustering algorithm[J]. Pattern recognition, 2003, 36(2): 451-461.

[11] Tien T L. A research on the grey prediction model GM (1, n) [J]. Applied mathematics and computation, 2012, 218(9): 4903-4916.

[12] Zheng-Xin Wang, Qin Li, Ling-Ling Pei, A seasonal GM(1,1) model for forecasting the electricity consumption of the primary economic sectors, Energy, vol 154, 2018, Pages 522-534

[13] Borowy, I. & Harris, B. (2023). Introduction: Health and development. In I. Borowy & B. Harris (Ed.), Health and Development (pp. 1-30). Berlin, Boston: De Gruyter Oldenbourg.

[14] Nguyen H X, Yamamoto R. Modification of ecological footprint evaluation method to include non-renewable resource consumption using thermodynamic approach [J]. Resources, Conservation and Recycling, 2007, 51(4): 870-884.

[15] Asratie. (2022). Determinants of electricity production from renewable source excluding hydroelectricity in selected East African countries: Panel ARDL approach. Cogent Economics & Finance, 10(1). https://doi.org/10.1080/23322039.2022.2080897

[16] Chen C N, Tsaur T W, Rhai T S. The Gini coefficient and negative income [J]. Oxford Economic Papers, 1982, 34(3): 473-478.

[17] Spash C. The development of environmental thinking in economics [J]. Environmental Values, 1999, 8(4): 413-435.