

Analysis of Green GDP Revolution Neural Network Prediction Model

Xinyi Li*

School of Statistics, Shanxi University of Finance and Economics, Taiyuan, China, 030031

* Corresponding Author Email: lxy_20031030@163.com

Abstract. In recent decades, GDP, the core indicator of the national economy, has played a key role in measuring the state of the economy and the level of development. However, GDP falls short as a measure of economic health, failing to fully account for the interrelationships between resources, the environment and the economy. In order to more comprehensively reflect the balance between the economy and the environment, we have introduced green GDP as an important indicator to measure the degree of social and economic health. This paper discusses in depth the differences in the definition, understanding and calculation methods of GGDP. An in-depth comparison of GDP and GGDP is made based on the grey forecast model to analyze their differences in measuring socio-economic health. At the same time, in order to measure social and economic health status more accurately, this paper classifies green GDP based on feedforward neural network model theory and clustering algorithm, builds a neural network prediction model, analyzes the correlation degree between different factors and GGDP, and analyzes the influence degree of different factors in resources and environment on social and economic status and development level. Through reasonable summary and analysis of the research results, considering the state of social and economic health, this study provides useful reference suggestions for the future government to implement economic measures and policies, and has positive significance for promoting the innovation and sustainable development of economic indicators.

Keywords: Green GDP, Neural Network Prediction Model, Clustering Algorithm.

1. Introduction

In recent decades, GDP (the final result of production activities carried out by all resident units in a certain period of time in a country or region), as the core indicator of the national economy, has been an important indicator to measure the economic status and development level of a country or region. However, GDP still has many shortcomings: the interrelationship between resources, environment and economy is not fully reflected. The more resources are consumed and the more serious the environmental degradation, the higher the GDP, which promotes economic growth to a certain extent. Economic activities such as resource restoration and environmental pollution control have been carried out, and GDP has once again increased, which is reflected in economic growth^[1]. From the perspective of natural and sustainable economic development, this dual effect has increased GDP to a certain extent and caused a certain degree of ineffective economic growth. Therefore, we need to change the way GDP is calculated to use the GGDP form - in order to urge humanity to contribute to the conservation of the planet^[2].

At present, the definition and understanding of green economy and the analysis and research methods of GGDP^[3] calculation is different at home and abroad. The concept of "green economy" was first put forward in the Green Economy Blueprint published by British environmental economist Pearce in 1989^[4]. Environmental economists believe that economic development must be able to withstand the natural environment and human beings themselves, and will not cause social division and ecological crisis due to blind pursuit of production growth, and will not cause sustainable economic development due to the depletion of natural resources. The green economy also specifically refers to social capital organizations (SOCs)^[5], which refer to local communities, business groups, trade unions, national legal and political organizations, and international environmental treaties (such as the Law of the Sea and the Montreal Convention). These social organizations, they argue, are more than the sum of individuals. Regardless of the level of the organization, it derives its individual habits,

norms, emotions, traditions, procedures, memories, and cultures to foster different levels of efficiency, dynamism, motivation, and creativity dedicated to the creation of human well-being. Li Fanglin^[6] mentioned the calculation method of production method and expenditure method of Green GDP in his article, and He Yumei^[7] mentioned the construction of Green GDP2.0 accounting index system and the cost accounting of various natural resources consumption in her article. However, although their research methods are constantly proposing and constructing new calculation methods of Green GDP, I will choose appropriate calculation methods based on them and thoroughly compare the differences between GDP and Green GDP in measuring social and economic health by means of gray prediction models. At the same time, I made a reasonable summary and analysis of the results of GDP and Green GDP to consider the social and economic health status. The research on this problem can provide reference suggestions for the future government departments to implement economic measures and policies.

2. The basic fundamental of neural network

2.1. The structure of neural network

The most basic and commonly used model in grey system modeling is GM (1,1) model, which is a model established by summing up the original sequence $X^{(0)}$ to generate $X^{(1)}$ ^[8]. Its grey differential equation is as follows^[9]:

$$X^{(0)}(k) + aZ^{(1)}(K) = b \quad (1)$$

Where, a and b are undetermined parameters,

$$Z^{(1)}(k) = 0.5X^{(1)}(k) + 0.5X^{(1)}(k - 1) \quad (2)$$

The corresponding whitening equation is:

$$\frac{dx^{(1)}(t)}{dt} + ax^{(1)}(t) = b \quad (3)$$

The solution is as follows:

$$X^{(1)}(t) = b/a + Cexp(-at) \quad (4)$$

When $t=0$, $x^{(1)}(0) = \frac{b}{a} + C$, solve $C = x^{(1)}(0) - \frac{b}{a}$

Thus, the time response function of the original equation is:

$$X^{(1)}(t) = [x^{(1)}(0) - b/a]exp(-at) + b/a \quad (5)$$

The time response function regards all the data as a set of special data that conforms to the change law of a certain function and uses these limited discrete data to fit a function or differential equation to form a change law of the data. According to this rule, the development trend of the data can be predicted, which shows the advantage of the minority data modeling in the gray system. Suppose the parameters have been determined, the time response function (5) is taken as the time response function of GNNM (1,1)^[10], denoted as:

$$y(t) = [x^{(1)}(0) - b/a]exp(-at) + b/a \quad (6)$$

If we multiply both sides of equation (6) by $\frac{1}{1 + exp(-ak)}$ we get

$$x^{(1)}(k)1/(1 + exp(-ak)) = [x^{(1)}(0) - b/a](exp(-ak))/(1 + exp(-ak)) + b/a \times 1/(exp(-ak)) \quad (7)$$

Then there is,

$$\begin{aligned} x^{(1)}(k) &= [x^{(1)}(0) - b/a](exp(-ak))/(1 + exp(-ak)) + b/a \times 1/(1 + exp(-ak))[1 + exp(-ak)] \\ &= [x^{(1)}(0) - b/a][1 - 1/(1 + exp(-ak))] + b/a \times 1/(1 + exp(-ak))[1 + exp(-ak)] \quad (8) \end{aligned}$$

$$= \{[x^{((1))}(0) - b/a] - x^{((1))}(0) \times 1/(1 + \exp(-ak)) + 2 \times b/a \times 1/(1 + \exp(-ak))\}$$

Equation (8) is mapped to the neural network to obtain the GNNM (1,1) network structure, as shown in Figure 1.

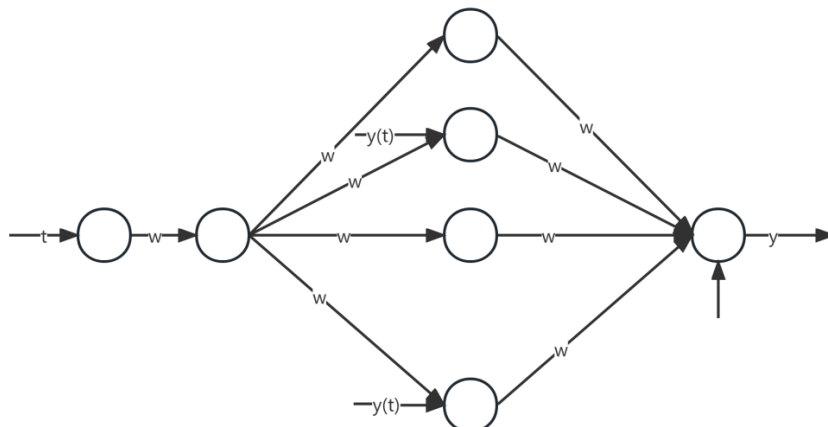


Figure 1. Neural network structure

2.2. Basic assumption and symbols

- (1) Assuming that the level of economic development is relatively stable and there will be no explosive growth or decline in a short time;
 - (2) It is assumed that, except for the indicators studied, other indicators have little impact on the system and can be ignored;
 - (3) Suppose we assume that the same policy has the same impact in different regions.
 - (4) We do not consider the impact of special circumstances unless mentioned in the article.
- The specific description of each data is shown in Table.1.

Table 1. Symbols

Symbols	Definition
$F_i (i = 1, 2, \dots, m)$	The score of each component
$r(X_i, X_j)$	Correlation coefficient between indicators
$\alpha_i (i = 1, 2, \dots, m)$	The contribution rate of each component
X	Input vector
$Cov(X_i, X_j)$	Correlation coefficient between indicators
r	Network prediction error
ρ	Pearson correlation coefficient
τ	Spearman correlation coefficient
	Kendall coefficient

3. Results

3.1. The establishment of simulation model 1

First, we will do the data processing.

GDP: Gross domestic Product (GDP) is the core index in the system of national economic accounting. It refers to the total value of the products and services newly produced by all the citizens of a country (or region) in a certain period. It refers to the final results of the production and operation

activities of all the resident units of a country (or region) in a certain period. The added value created by foreign-funded enterprises in China is included in GDP_[11].

Resource consumption cost: refers to the value of resources consumed in economic activities. Since economic development depends on fossil energy and water resources, we choose fossil energy consumption and water consumption to measure the cost of resource consumption. In addition, scholars also take cultivated land resources as the accounting of resource consumption cost, but considering the annual data of cultivated land occupation changes of various departments, this paper does not include it_[12].

Environmental improvement cost: refers to the positive economic benefits brought by the use of waste or the improvement of environmental conditions, including the output value of waste comprehensive utilization and the ecological benefits of the garden. However, due to data limitations, we exclude environmental improvement costs and use the following formula to calculate China's green GDP. Shen et al. (2018) argued that environmental improvement costs had little impact on the estimated results, so ignoring environmental improvement costs had no significant impact on the results_[13].

The cost of environmental pollution: also known as the cost of environmental degradation, refers to the value of environmental pollution loss and the cost of environmental protection. Currently, most relevant studies use the cost of carbon emissions as a proxy for environmental costs, ignoring neither the cost of sulfur dioxide emissions nor the damage caused by water and solid pollution. Choose an air pollution account represented by carbon and sulfur dioxide emissions, a water pollution account represented by wastewater emissions, and a solid pollution account represented by solid waste emissions and solid waste storage to measure environmental pollution costs_[14].

GGDP: Subtract the consumption of natural capital from the gross domestic product to get the environmentally adjusted gross domestic product_[15].

Secondly, we give the construction system of Green GDP. Physical quantities of environmental pollutants are calculated according to industry data, and then converted into real values expressed in currency. Green GDP is obtained by eliminating environmental pollution and resource consumption costs of traditional GDP.

Finally, we get a model of Green GDP.

$$W_{carbon} = D_{carbon} \times P_{carbon} \quad (9)$$

W_{carbon} is the environmental pollution cost of carbon emission, D_{carbon} is the total carbon emission, and P_{carbon} is the conversion factor.

$$W_{SWD} = D_{SWD} \times P_{SWD} \quad (10)$$

$$W_{SWS} = D_{SWS} \times P_{SWS} \quad (11)$$

Where, W_{SO_2} is the cost of environmental pollution emitted by SO_2 , P_{SO_2} is the shadow price of SO_2 , and D_{SO_2} is the emission of SO_2 during fuel combustion and production. By definition, the shadow price of pollutants can be expressed as the change of environmental output caused by the change of each unit of pollutant at a certain output level, so the shadow price of SO_2 can measure the impact of SO_2 emissions on output.

$$W_{WP} = D_{WP} \times P_{WP} \quad (12)$$

W_{wp} is the environmental cost of wastewater discharge, and D_{wp} is the cost of wastewater treatment.

$$W_{SWD} = D_{SWD} \times P_{SWD} \quad (13)$$

$$W_{SWS} = D_{SWS} \times P_{SWS} \quad (14)$$

W_{SWD} represents the environmental pollution cost of solid waste discharge, solid waste discharge on D_{SWD} , treatment cost per unit of solid waste discharge on P_{SWD} , environmental pollution cost per unit of solid waste storage on W_{SWS} , storage amount of solid waste on D_{SWS} , and treatment cost per

unit of solid waste storage on P_{sws} .

$$W_F = D_F \times P_F \quad (15)$$

W_F represents the cost of fossil energy consumption F the shadow price of fossil energy consumption in an economy.

$$W_W = D_W \times P_W \quad (16)$$

W_W represents the cost of water resource consumption is distributed by transpiration.

So, the Green GDP result is:

Green GDP = Traditional GDP - (Environmental pollution cost + Resource depletion cost) + Environmental improvement cost

Some methods of assigning monetary values exhibit a degree of arbitrariness, and the availability and reliability of data remains a challenge for many countries, particularly developing countries, as the required data can be subject to incomplete coverage, measurement errors and biases. Therefore, Amount can be converted into cost according to the current price, such as the price of water resources and mineral resources, or into electric energy, which can be calculated as follows^[16]:

$$GreenGDP = GDP - (KtCO_2 \times P_{CDM}) - (Twaste \times 74kWh \times Pelect) - \left(\frac{GNI}{100} \times \%NRD \right) \quad (17)$$

At the stage of trying to replace GDP with GGDP, the indicators of Green GDP are clustered and analyzed, and their environmental resources are divided into nine categories^[17]: Land resources(main indicators have cultivated land resource consumption reduction of cultivated land quality index, index) ten land desertification and water erosion; Grassland resources (main indicators are grassland area of grassland area of cost reduction, the affected area, grassland quality index, grassland degradation);Water (main indicators available freshwater resources, consumption reduction, declining water quality, Flood disaster);Forest resources (the main indicators include forest area, forest cutting amount and cutting density, forest affected area);Wild animal and plant resources (the main indicators include loss of wild animal and plant species, loss of population, increase of rare species, Wildlife habitat consumption reduction);Mineral and energy resources (the main indicators include the amount of mineral and energy savings, the amount of exploitation and the total amount of pollution produced);Marine resources (the main indicators include the quality of offshore water bodies, the degree of water pollution and the decline of offshore biological resources);Renewable resources, It refers to the natural resources that can be restored or even exceed the original scale and level through labor input (the main indicators include water resources recycling, forest resources increased by afforestation, grassland increased by grass planting, cultivated land increased by county desertification control (grassland), species, wetland and fishery resource reserves increased by conservation measures);environmental resources (greenhouse gas emissions). Production and consumption of ozone-depleting substances, concentration of suspended particulate matter (TSP), production of "three industrial wastes") Secondly, the Kendall grade correlation coefficient was used for correlation analysis. In addition, the CART regression tree is used to build the model^[18]. Based on the principle of minimum mean square error, the factors such as humidity, soil evaporation, and precipitation were divided. The specific steps are as follows:

- **Step1:** For any partition factor A , the corresponding partition points
- **Step2:** Parent node splits into two children (hypothetical)
- **Step3:** The data sets of the two nodes are D_1 and D_2
- **Step4:** To minimize the mean square error of D_1 and D_2
- **Step5:** At the same time, the factor and factor value corresponding to the minimum sum of mean square error of D_1 and D_2 are used as the dividing point
- **Step6:**

$$\min_{A,s} [\min_{c_1} \sum_{x_i \in D_1(A,s)} (y_i - c_1)^2 + \min_{c_2} \sum_{x_i \in D_2(A,s)} (y_i - c_2)^2] \quad (18)$$

c_1 is the mean of the first node, c_2 is the mean of the second node

- **Step7:** The autoregressive moving average model ARIMA

$$y_t = \mu + \phi_t + \sum_{i=1}^q \theta_i \phi_{t-i} + \sum_{i=1}^p \gamma_i y_{t-i} \quad (19)$$

3.2. Model 1 results and analysis

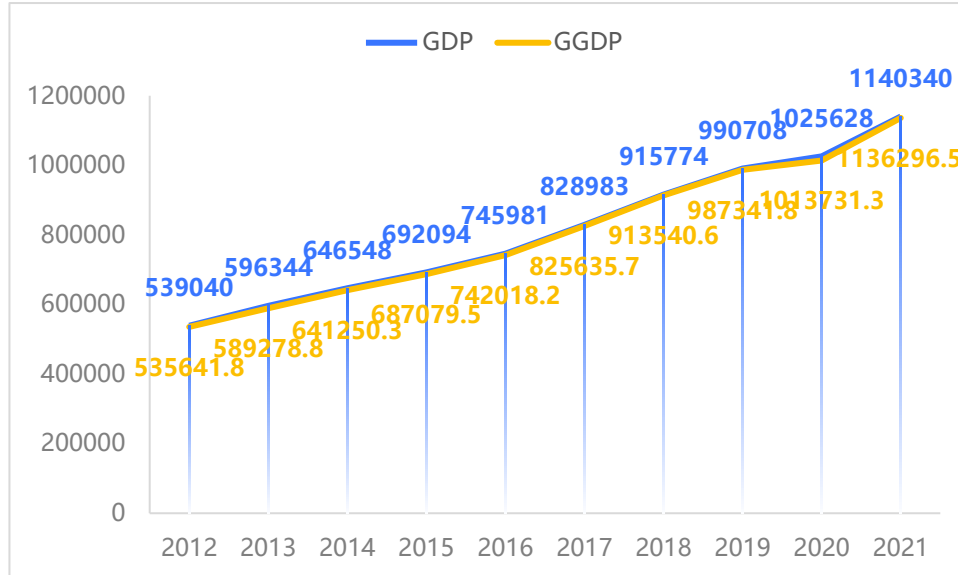


Figure 2. Comparison of China's GDP and GGDP data

As shown in Figure 2, through the comparison of GDP and GGDP data, it can be seen that the value of GDP is always higher than that of GGDP, and the data change trend of GGDP and GDP is almost identical. Replacing GDP with GGDP may meet resistance. Models suggest that, globally, the shift is worth it.

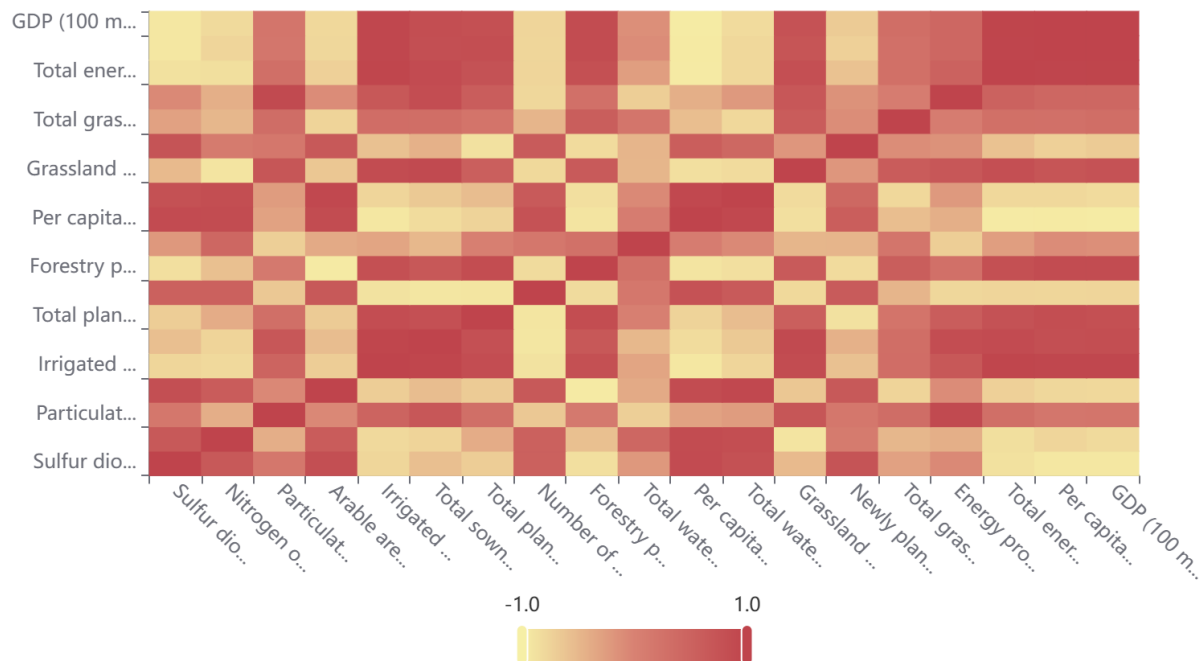


Figure 3. Correlation coefficient heat map

As shown in Figure 3, there is a strong correlation between the various environmental factors affecting GGDP, and they are closely related to GDP. In particular, the relationship between energy

resources, grassland resources and irrigation resources are very high.

Table 2. Intra-group correlation coefficient result table

Intra-group correlation coefficient analysis							
	Intra-group correlation	95%confidence interval		Use the F-test with a true value of 0			
		Lower limit	Upper limit	value	df1	df2	P
Individual measurements ICC (1, 1)	0.962	0.93	0.984	254.821	16	153	0.000***
Average measurements ICC (1, k)	0.996	0.993	0.998	254.821	16	153	0.000***
Note: ***, **, and * represent the significance levels of 1%, 5%, and 10%, respectively							

As shown in Table.2, the results of the ICC analysis, including confidence intervals, F-test results, etc. The results of the intra-group correlation coefficient for a single measurement showed that the significance P value was 0.000***, which showed significance at the level, rejecting the null hypothesis, indicating that the consistency of reliability was credible. And the correlation coefficient is 0.962, indicating that the reliability of the data is very strong.

The results of the intra-group correlation coefficient for the average measurement showed that the significance P value was 0.000***, which showed significance at the level, rejecting the null hypothesis, indicating that the consistency of reliability was credible. And the correlation coefficient is 0.996, indicating that the reliability of the data is very strong.

Table 3. results

Method	MAE	RMSE	R ²
APIMA	1.447	1.764	0.662
Prophet	1.532	1.843	0.392
Arima+Prophet+random forest	1.031	1.226	0.825

In this paper, APIMA, Prophet, and random forest are also used to test the prediction results. The smaller MAE is, the smaller RMSE is and the larger R² is, which indicates that the accuracy of model prediction is higher. As shown in Table.3, the results are credible.

3.3. The establishment of simulation model 2

Based on the above results, this paper explores the impact of GGDP when it is converted into GDP. In this paper, casualties and average annual wages due to disasters caused by environmental degradation are used because it is difficult to determine the associated morbidity and associated working hours for human health loss. In this part, it is difficult to obtain the environmental maintenance costs due to the accelerated depreciation rate of fixed assets due to environmental degradation. The impact of economic activities on the quality of the environment is reflected in the discharge of various wastes. This paper argues that the associated pollution control costs are the loss of environmental capacity (quality) due to environmental pollution. The feedforward neural network used in this paper is a simple gradient descent method that aims to minimize the error of the network's computational output. The training process is to continuously adjust the weights and thresholds to reduce the network error to a predetermined minimum or stop at a predetermined training step^[19].

- **Step1:** The sample set was sampled using the iterative sampling method
- **Step2:** Randomly generate k training sets
- **Step3:** A decision tree corresponding to it is further generated
- **Step4:** When the its decision tree is trained, the variance of the distance between all the

predicted points and the actual points in the decision tree is calculated

- **Step5:** The weight of the its decision tree is:

$$w_r(i) = \frac{1/\delta^2(i)}{\sum_{j=1}^T 1/\delta^2(j)} \quad (20)$$

- **Step6:** Thus, the predicted value of the algorithm can be obtained:

$$\hat{u}(x) = \sum_{i=1}^k w_r(i)Z_i \quad (21)$$

3.4. Model 2 results and analysis

We calculate the value of natural resource depletion through user costs. The calculation of the loss value of depletion of natural resources caused by economic activities should include the consumption of coal, oil and natural gas and the resource losses caused by forestry cutting, water resources consumption, fishery resources consumption and Marine resources consumption. However, due to the limitation of data, this paper only calculates the loss of coal, oil and natural gas and other mineral resources, forestry resources, water resources and fishery resources. In this way, the green GDP calculated by us will be higher than the real value of green GDP, but this does not affect the original intention of this paper to accelerate the environmental resources accounting and establish a green GDP accounting system to ensure sustainable development^[20].

As a large ecosystem, natural resources have a wide range of ecological functions, so it is difficult for us to divide them into a comprehensive single ecological function. Furthermore, due to the limitations of current statistical data and the relative lack of knowledge of the author, only a part of the single ecological functions, namely forest ecological functions and cultivated land ecological functions, are calculated in this paper. Garden and grassland ecological functions.

4. Conclusions

This paper uses the neural network model and the grey prediction model to compare the differences between GDP and green GDP in measuring socio-economic health, establishes the calculation method of green GDP, and proves that green GDP has certain advantages over GDP in measuring socio-economic health through comparative results and test analysis, and the transformation of replacing GDP with GGDP is worthwhile on a global scale.

References

- [1] Zhang Qi. Research on the relationship between GDP and green GDP, GEP and natural resource value[J]. China Statistics,2023, (01):11-14.
- [2] Liu Qihang, Ouyang Kang. Rethinking the performance evaluation of green GDP under the vision of "dual carbon"[J]. Theory Monthly,2023, (10):91-97.
- [3] Duan Lin.Green GDP accounting and sustainable development evaluation of Sichuan Province[J]. Foreign Investment in China,2023, (16):42-44.
- [4] Yi Wenfang. The evolution of China's development Concept and the choice of green development Times [J]. Environment,2023, (07):75-77.
- [5] Tang Xuebin, Shu Ping, Zhang Shoulei. Spatial econometric analysis of social capital and regional economic growth: Based on the perspective of Chamber of Commerce [J/OL]. Journal of Guangxi Normal University (Philosophy and Social Sciences Edition),2023, (12):1-18.
- [6] Li Fanglin, Zhang Haibing. A comparative study of current GDP and green GDP [J]. Shanghai Statistics,2003, (06):19-22.
- [7] He Yumei, Wu Shasha. Method based on resource value loss of green GDP accounting system to build [J]. Journal of statistics and decision, 2017, (17): 5-10.

- [8] Zhao Junhai, Hua Linwei, Wang Yu. Study on axial compressive capacity of concrete-filled steel tube with hollow sandwich based on BP neural network of particle swarm optimization [J/OL]. Progress in building steel structures, 2023, (12):1-9.
- [9] Yen-Tseng Hsu, Chuy-Shin Cheng, Chowan-Chia Wu. Grey Dynamic Modeling and Prediction Control of Macroeconomic System[J]. Journal of Grey System .1997, (31).
- [10] Kong Xinhai, Chen Jiajia, Zhao Yong. Grey GM (1,1, k, k~2) model background value and time response function optimization [J]. Operations Research and Management, 2012, (07):109-113.
- [11] Si Puxu. GDP forecast of Shandong Province based on variable selection and ARIMAX model [D]. Shandong University,2020:000983.
- [12] Qiu Jinping. Application of resource consumption accounting in cost budget management [J]. Journal of Finance and Accounting,2020, (12):1002-8072.
- [13] Li Y, Yang P, Wang H. Collecting coal-fired power environmental tax to promote wind power development and environmental improvement[J]. Acta Scientific Malaysia,2018, (1).
- [14] Zhou Xinjun, Yang Xiaofang. Research on environmental cost management and control under the background of dual-carbon target [J]. Shandong Textile Economy,2023, (10):27-31.
- [15] Wang Yuchen. Study on the relationship between informal system and green GDP growth in China [D]. Xizang University,2023:2023.000014.
- [16] Gao Xirong, Zhao Xinyue, Chai Chengxin. Green GDP accounting based on resource and environment value compensation: A case study of Chongqing [J]. Ecological Economy, 2019,39(05):148-153.
- [17] Technology -Green Technology; Study Results from Guilin University of Electronic Technology Broaden Understanding of Green Technology (Measuring the Efficiency of China's Carbon Market: A Comparison Between Efficient and Fractal Market Hypotheses) [J]. Journal of Technology,2020.
- [18] Wang Xiyue. Prediction and application of stock market based on multi-factor model and CART classification regression tree [D]. Zhengzhou University,2022: 2021.002873.
- [19] Chen Yuhang, Wang Yuhong, Nan Lu et al. Short-term load forecasting based on self-organizing map-Feedforward neural network and Prophetic hybrid model [J/OL]. Modern Electric Power,2023:1-9.
- [20] Yang Qijie. Path selection of Promoting Green Development with Green GDP [J]. Journal of Hebei Youth Management Institute,2022,34(05): 13-126