Hindawi
Complexity
Volume 2021, Article ID 6659302, 14 pages
https://doi.org/10.1155/2021/6659302

Research Article

Use of BP Neural Networks to Determine China's Regional CO_2 Emission Quota

Yawei Qi, Wenxiang Peng, Ran Yan, and Guangping Rao

School of Information Management, Jiangxi University of Finance and Economics, Nanchang 330032, China

Correspondence should be addressed to Yawei Qi; qiyawei@jxufe.edu.cn

Received 26 November 2020; Revised 4 December 2020; Accepted 14 December 2020; Published 6 January 2021

Academic Editor: Abd E.I.-Baset Hassanien

China declared a long-term commitment at the United Nations General Assembly (UNGA) in 2020 to reduce CO_2 emissions. This announcement has been described by Reuters as "the most important climate change commitment in years." The allocation of China's provincial CO_2 emission quotas (hereafter referred to as quotas) is crucial for building a unified national carbon market, which is an important policy tool necessary to achieve carbon emissions reduction. In the present research, we used historical quota data of China's carbon emission trading policy pilot areas from 2014 to 2017 to identify alternative features of corporate CO_2 emissions and build a backpropagation neural network model (BP) to train the benchmark model. Later, we used the model to calculate the quotas for other regions, provided they implement the carbon emission trading policy. Finally, we added up the quotas to obtain the total national quota. Additionally, considering the perspective of carbon emission terminal, a new characteristic system of quota allocation was proposed in order to retrain BP including the following three aspects: enterprise production, household consumption, and regional environment. The results of the benchmark model and the new models were compared. This feature system not only builds a reasonable quota-related indicator framework but also perfectly matches China's existing "bottom-up" total control quota approach. Compared with the previous literature, the present report proposes a quota allocation feature system closer to China's policy and trains BP to obtain reasonable feature weights. The model is very important for the establishment of a unified national carbon emission trading market and the determination of regional quotas in China.

1. Introduction

Because of the 2020 COVID-19 epidemic, human beings have become more aware of their relationship with nature and of the importance of sustaining a harmonious coexistence of man and nature. In a time of significant crises, including the COVID-19 epidemic and climate change, the international community agrees that only through the development and implementation of green and low-carbon technologies, society can achieve high-quality economic recovery [1–3]. On Sept. 22, during the General Debate of the 75th Session of the UNGA, Chinese President Xi Jinping declared that China aims to reach CO_2 emissions peak before 2030 and achieve carbon neutrality before 2060. According to this, China plans to restore its economy by promoting low-carbon technologies and lifestyle. Reducing CO_2 emissions has become an important goal of China's 14th Five-Year Plan.

During the "13th Five-Year Plan" period, the Chinese government learned from the successful experience of the European Union, Japan, and other economies in reducing CO_2 emissions and began to explore the use of market-based methods: CO_2 emission trading systems [4, 5]. From 2014 to 2019, the central government implemented a pilot CO_2 emission trading policy in 7 provinces and cities, and each regional government formulated relevant trading standards and rules. The government in the pilot regions implemented the "allocation + trading" quota management rules for emission-control enterprises, that is, the emission-control enterprises received free quotas issued by the government at the beginning of the performance period. These quotas are determined after an enterprise's self-inspection and CO_2 emissions report is issued and a third-party verification is performed. If an enterprise's CO_2 emission is exceeded/remained during the period, it can be bought/sold in the carbon trading market. After six years, China's pilot carbon market has increased and has become the world's second largest carbon market in terms of quota trading volume. Preliminary statistics have shown that a total of 2,837 emission-control agencies, 1082 nonemission-control agencies, and 11,169 individuals have participated in the pilot market. The cumulative transaction volume of the 7 pilot market quotas has reached 406 million tons, and the cumulative transaction volume is about 9.28 billion yuan. By the end of 2019, China's carbon intensity was reduced by about 48.1% as compared with data from 2005. In addition, nonfossil energy accounted for 15.3% of the primary energy consumption, which means that China has achieved its 2020 emission reduction target ahead of schedule. Thus, China is currently applying market mechanisms to control and reduce greenhouse gas emissions and to promote the green and low-carbon transformation of economic development. Moreover, the implementation of the CO_2 emissions trading market represents not only an important institutional innovation for China but also an important policy tool to implement international agreements for emissions reduction. Given the significant emissions reduction results due to the implementation of the CO_2 emission trading market policies in the pilot regions, the Chinese government has initiated the creation of a unified national CO_2 emission trading market to help all regions in the country to reduce carbon emissions. It is expected that with these new rules, the CO_2 emission peak target by 2030 will be achieved as soon as possible. Therefore, the design and implementation of a unified national CO_2 emission market is an issue that needs to be studied urgently. The foundation for a proper design and implementation of a CO_2 emission trading market program to achieve the intended emission reductions resides in the correct determination of the national quota and quotas for each region (province and city).

After finding alternative features of corporate carbon emissions, we used quota data on China's carbon emission trading policy pilot areas from 2014 to 2017 and the BP model to calculate the quotas of other regions in the sample interval, provided that such regions implemented the carbon emissions trading policy. We obtained the national total quota. With respect to carbon emission terminal, we divided the quota allocation system into three aspects: enterprise production, household consumption, and regional environment and then retrained the BP to obtain new results which were compared to those of the benchmark model. We found out that (1) from 2014 to 2016, China's total quota displayed a yearly increase and a sudden decrease in 2017. During the initial stage of the national CO_2 emission trading market program, the national quota is expected to maintain a relatively high level. Later, during the following 3 to 4 years, through adjustment and adaptation, quotas in each region are expected to show a downward trend and increase in the change rate. This forces enterprises to either participate in the CO_2 emission trading market or improve their technology to reduce emissions. (2) Considering the feature system built by adding household consumption and regional environment, the training model displays a smaller loss rate, and the test results (other regional quotas) describe the actual situation in a more accurate way. Thus, when building a unified national CO_2 emission trading market and determining quotas for various regions in addition to enterprise production, it is more reasonable to consider a feature system that takes into account household consumption and regional environment. At the same time, this feature system can be used in combination with China's "bottom-up" total control and postadjustment method. It not only allows regional quota decision makers to predict CO_2 emissions in advance through the existing data in the feature system and the trained model before obtaining the final real summary of the CO_2 emissions but also allows enterprises considering CO_2 price when making investment decisions and trying to make profits in the market.

2. Literature Review

Previous research has mainly studied the following two aspects: (a) initial distribution of quotas for CO_2 emissions according to different principles and methods and (b) the allocation efficiency according to the initial allocation of carbon emission quotas.

The publications in the first area reported the study of distribution subjects and distribution methods. CO_2 emission quota allocation can be divided according to the two following perspectives: region and industry. Region includes the initial quota allocation among countries and the initial quota allocation among different provinces within the same country. In this context, the earliest research studied CO_2 emission quota allocation among countries. In 1992, the United Nations Conference on Environment and Development established "common but differentiated responsibilities" as the principle of international environmental cooperation. Later, in 1998, Rose et al. introduced the principle of CO_2 emission rights allocation, which should include equity and efficiency. However, what kind of allocation method involves "equity" and how to balance the importance of "equity" and "efficiency" are controversial topics [6]. Some scholars in developed countries believe that allocating CO_2 emission rights based on population size is in line with the principle of equity [7]. However, other experts in developing countries have proposed that "accumulated emissions per capita" is more related to equity. According to this concept, the allowed CO_2 emissions per capita in developing countries during the development stage should be higher than those in developed countries [8, 9]. Later, more researchers used different features and methods to quantitatively analyze the initial quota [10–12]. Based on population, GDP, and CO_2 emission data for 132 countries, Wang et al. proposed the Gini coefficient optimization model that optimizes the historical CO_2 emissions quota for various countries and is able to project future quotas [13].

Chinese scholars have focused on CO_2 quota allocation among provinces and industries in China [14–19]. Song et al. considered the comprehensive distribution principle of three indicators: hereditary system, egalitarianism, and payment, to create the provincial environmental fixed cost allocation optimization model (FCAM), which will be able to determine the allocation of provincial CO_2 emission rights for 2020, with a more balanced equity and efficiency [20]. From the perspective of equity and efficiency, Yu and Wu used a master-slave hierarchical interactive iterative algorithm based on satisfaction to build a two-level planning model (the upper-level planning model based on equity and the lower-level planning model based on efficiency) to optimize the allocation of CO_2 emission rights between provinces [21]. Qian et al. used Chinese enterprise carbon patent data, and from a consumption and production perspective created the stochastic frontier model to measure regional CO_2 emission efficiency. Later, they used the estimated efficiency value to numerically simulate the regional allocation of CO_2 emission rights [22]. Wu et al. proposed a coupling model of China's multiregional CGE and CO_2 trade (CGE-3MS). The model showed the decision-making and optimization process of trading CO_2 units, and they analyzed the impact of the carbon market on the economy and emission-control industries in China under different initial quotas [23].

Compared with the first piece of literature, the second one focused on whether the existing allocation of quotas between regions and industries is reasonable and effective. In general, allocation efficiency is used to measure whether the allocation is reasonable. Some Chinese scholars used the original DEA and the zero-sum DEA models to measure quota allocation for Chinese provinces. They found out that the results of the zero-sum DEA model were better than those obtained with the original DEA model [24]. In addition, other scholars revised the original DEA model and proposed a new allocation method that included the evaluation of DEA efficiency and historical CO_2 emissions. Later, they used China's emission commitment as the decision-making guidance and selected the maximization of the average efficiency as the final goal [25, 26]. Besides, after proving that the original DEA model was inefficient to determine CO_2 emissions quotas, some scholars used other models to redistribute the quotas of various industries [27–29]. Huang and Zhang used the SBM and RE/CE models to get a more comprehensive efficiency that reveals Chinese energy use and the CO2 emission situation. From the empirical study of 30 regions in China, they found that the southern region of China has the most efficient score, while northeastern China has poor performance. Price factor has a significant influence on energy use and CO_2 emission efficient score of some provinces [30].

In summary, we have identified two points that should be taken into account when creating regional quota allocation systems and building the corresponding models. First, feature selection should be performed with caution. Factors related to the regional CO_2 emission accounting as well as those related to CO_2 emissions of the accounting entities (enterprises participating in the CO_2 emission trading market) should be selected. Second, when using the feature system data in the allocation model, it is necessary to carefully determine the weight of each feature to ensure the scientific, rigorous, and accurate quota allocation.

Therefore, in the present research, we first selected features that are closely related to the accounting entities (emission-control enterprises), such as the number of industrial enterprises exceeding designated size in the region, the energy consumption per unit of industrial added value, and the proportion of coal-fired energy included in the total energy consumption. We also considered features that are related to regional accounting CO_2 emissions, such as per capita energy consumption, per capita carbon emissions, forest coverage, and green areas. Then, we built a feature system that can be used in conjunction with China's "bottom-up" total control and postadjustment method to provide a predictive model. In addition, this model allows enterprises considering CO_2 prices when making investment decisions, thereby stimulating corporate green innovation. In addition, for the first time, we considered the use of the BP in order to determine the weight of each feature based on historical quota data. This is different from the previous literature, which used subjective weight determination methods. In addition, our methods are more in line with China's actual CO_2 emission situation and more accurate during calculations. Finally, the feature system and the corresponding BP neural network model proposed in the present research can be used to calculate the quotas of other regions during the same period and also predict future quotas for the same region.

3. Empirical Analysis

3.1. Data and Feature Statistics

3.1.1. Data. On October 29, 2011, China´s Development and Reform Commission indicated that China should start implementing the pilot carbon emission trading policy. Specifically, Beijing, Tianjin, Shanghai, Chongqing, Guangdong, Hubei, and Shenzhen were labeled as the pilot regions. This allowed each regional government to determine the transaction start time, corresponding transaction subject enterprises, and quotas according to local conditions. Although the central government promulgated the pilot policy in 2011, each pilot region actually started the carbon emission trading system between 2013 and 2014. Therefore, considering the complete CO_2 quota and the availability of additional data, we selected 31 regions (5 cities and 26 provinces) as the research sample for the period 2014–2017. With respect to sample pretreatment, the following aspects were considered: (1) given that Shenzhen belongs to Guangdong Province and that both of them are pilot regions, Shenzhen was analyzed separately; (2) regions with serious missing values in historical feature data were eliminated (Tibet); (3) taking into account that the different dimensions of the data may affect the prediction, all features were normalized using

$$b_{ijt} = \frac{a_{ijt} - \min a_{ijt}}{\max a_{ijt} - \min a_{ijt}}, \tag{1}$$

where a_{ijt} indicates the jth original feature of the i region in the t year, $\min a_{ijt}$ and $\max a_{ijt}$ represent the minimum and maximum values of the original feature j in 31 regions in the t year, respectively, and b_{jit} indicates the j feature of i region in the t year after normalization.

Regional features (except that for Shenzhen) were obtained from the National Bureau of Statistics and China Statistical Yearbook. Features of Shenzhen were obtained from the Shenzhen Statistical Yearbook, and the historical quotas data came from the 2014–2018 Beijing Carbon Market Annual Report.

China and its regions have not released official data on CO_2 emissions. Thus, we used the CO_2 emission calculation method given in the 2006 IPPC National Greenhouse Inventory Guidelines and reported by Qi [31]:

$$\begin{aligned} CO_2 &= \sum_{i=1}^{n} EF_i \times E_i \times 10^{-6} \\ &= \sum_{i=1}^{n} CEF_i \times NCV_i \times OR_i \times \left(\frac{44}{12}\right) \times E_i \times 10^{-6}, \end{aligned} \tag{2}$$

where CO_2 indicates carbon dioxide emissions (t) and EF_i represents the carbon dioxide emissions factor for a specific fuel $(kg - CO_2/kg, m^3)$. We considered 11 types of energy consumption including raw coal, coal, coke, crude oil, fuel oil, gasoline, kerosene, diesel, liquefied petroleum gas, refinery dry gas, and natural gas $(n = 11)$; E_i is the fuel consumption for fuel i (kg, m^3); CEF_i represents the carbon content of the fuel $i\,(tC/TJ)$; NCV_i indicates the average low calorific value of fuel i $(kJ/kg, m^3)$; OR_i denotes the carbon oxidation rate of fuel $i\,(\%)$. The carbon dioxide emission coefficients of various energy sources are shown in Table 1.

Table 1: The CO_2 emission coefficients of various energy sources.

Energy	NCV	CEF	OR	EF
Raw coal	20 908	26.37	0.94	1.900 3
Coal	20908	25.8	0.92	1.8300
Coke	28435	29.5	0.93	2.860 4
Crude oil	41816	20.1	0.98	3.020 2
Fuel oil	41816	21.1	0.98	3.170 5
Gasoline	43070	18.9	0.98	2.925 1
Kerosene	43070	19.5	0.98	3.017 9
Diesel	42652	20.2	0.98	3.095 9
Liquefied petroleum gas	50179	17.2	0.98	3.101 3
Refinery dry gas	46055	18.2	0.98	3.011 9
Natural gas	38931	15.3	0.99	2.162 2

Notes: (1) NCV comes from China "General Principles of Comprehensive Energy Consumption Calculation" (GB/T 2589–2008); (2) CEF and OR come from China's "Provincial Greenhouse Gas Inventory Compilation Guide" (NDRC Office [2011] No. 1041).

3.1.2. Feature Statistics. Considering the calculation order, there are two carbon allocation methods:

(1) The "up-bottom" allocation method is applied from a macro (provincial) perspective, according to the general principle of efficiency and equity. This methodology considers population, provincial economic level (GDP), degree of industrialization (industrial structure), historical factors (accumulated carbon emissions per capita), and natural endowments to determine quotas for the different regions and for enterprises. In addition, the method allows determining the amount of quotas in advance, in such a way that participating entities are able to consider CO_2 prices in decision-making processes. However, due to the lack of data regarding actual emissions of the participating entities, when the quotas correspond to the actual situation and, in consequence, entities are motivated to participate in market transactions emissions reduction, these processes display great uncertainty.

(2) In the case of the "bottom-up" allocation method, the level-by-level summary determines the quotas for each region through the calculation of emissions at the microlevel (emission terminal). Considering carbon dioxide production terminal in human society (mainly consumers and enterprises), the enterprise component considers the number of equipment that emits CO_2, the scale of the enterprise, the scientific research level of enterprise emissions reduction (patents and R&D investment), and so on. The consumer component involves population size, per capita carbon emissions, and so on. Taking into account the emission data of the participating entities, this method is suitable for adjusting quotas afterwards (without specifying quotas in advance). However, it may eventually result in oversupply of quotas due to the excessively high emission limits. This situation may discourage participants to consider CO_2 prices when they are making investment decisions. In consequence, participants are not motivated to reduce emissions.

At the same time, given the problem of excessive CO_2 emissions caused by humans, people are also planting trees (forest carbon sinks) and using photosynthesis to reduce emissions. Therefore, resident life should also be considered when calculating quotas.

In 2020, the central government will formally begin to build a national carbon emission trading market. First, it will issue a national carbon emission trading quota setting and allocation implementation plan for the power industry (2019–2020). The plan involves a "bottom-up" quota determination method, that is, relevant provincial departments will be required to determine the list of key emitters and their actual output. Later, they will identify key emitters' quotas based on the benchmark method (the free quotas in each region for 2019–2020 were preallocated according to the 70% of the power (heat) supply of each key emitter in 2018). Then, after the quotas of all the key emitters in each region are verified, they will be added up to form the total quota of the region, and the regional quotas will be further added to obtain the total quota of the country.

According to China's "bottom-up" quota allocation method, it is most accurate to use the CO_2 emission equipment data of emission-control enterprises in the regions. However, since China's national carbon emission trading market is still in its infancy, relevant data (corporate power supply (heat) units, actual output) of emission-control enterprises in regions have not been released.

Also, "bottom-up" quota allocation methodologies present several limitations, which may result in oversupply of quotas, insufficient demand, low CO_2 price, and inactive market due to excessively high emission limits. Because of this, participants usually do not consider CO_2 prices when they are in the process of investment decision-making and cannot effectively motivate participants to reduce emissions.

Therefore, based on the operating experience of The European Union Emission Trading Scheme (EU ETS) and Regional Greenhouse Gas Initiative (RGGI), we proposed a more reasonable "top-down" allocation method in combination with China's existing "bottom-up" quota allocation method [32]. This method was divided into two steps. The first step involved the development of a feature system that is more suitable for China's provincial carbon quota allocation. This was performed by selecting features that are related to participating entities. For example, (a) the current emission-control enterprises in China are mainly industrial enterprises; thus, the region that contains more industrial enterprises will have more quotas; (b) energy consumption, including coal-fired energy that is used by emission-control enterprises in production activities ranks first in regional energy consumption; thus, energy structure and energy consumption per unit GDP are both factors affecting regional quotas. The second step involves training suitable BP to calculate nonpilot regional quotas based on historical pilot regional quotas and feature system data.

Most ETS are based on total CO_2 emission control. Determination of total CO_2 emissions should not only consider the overall emission target, but also the regional differences (level of economic development, technological differences, and forest carbon sink).

Based on previous reports and considering different factors included in the "top-down" quota control approach adopted by the EU ETS after phase II [33–35], we divided the features into three categories: (a) enterprise production, (b) household consumption, and (c) regional environment.

(a) Household consumption: in 2019, 30% of China's CO_2 emissions corresponded to consumer and man-made emissions. Household consumption, as the main body of society, is one of the main causes of CO_2 emissions. Regions with large populations display frequent economic activities. Therefore, resident consumption factors depend on per capita GDP, energy consumption, carbon emissions, cumulative carbon emissions, and disposable income.

(b) Enterprise production: in 2019, 70% of China's CO_2 emissions were the result of industrial production or generative emissions. Among them, the carbon dioxide emissions related to the power industry accounted for more than 40%. In addition, those from the steel industry, which are part of the manufacturing industry, accounted for about 15%. Industrial enterprises exceeding permitted size are not only the main body of the industry but also the main entities participating in the carbon emissions trading market. Therefore, factors related to CO_2 emissions from enterprises include the number of industrial enterprises exceeding permitted size, full-time equivalent of R&D personnel in industrial enterprises exceeding designated size, industrial structure, and energy structure, energy consumption per GDP unit, energy consumption per unit of industrial added value, and electricity consumption per GDP unit.

(c) Regional environment: in addition to factors related to resident and enterprises, regional environment also affects CO_2 emissions and decomposition. We considered three aspects, regional economic level, technological level, and green resources, mainly including total freight volume, total passenger volume, degree of openness to the outside world, total gas supply, total supply of liquefied petroleum gas, urbanization rate, technological level, green area, and forest carbon sink. Specific explanations of features are given in Table 2.

3.2. Empirical Model

3.2.1. Specifications of the Backpropagation Neural Network Model (BP). Compared with the subjective weight assignment methods (AHP, Expert Evaluation Method, TOPSIS, etc.) used in the formation of the "top-down" model and reported in previous literature, BP involves a multilayer feedforward network and error direction propagation-learning algorithm. Because of its unique adaptability, learning ability, and strong generalization ability, it is widely used in the fields of automatic identification, predictive estimation, engineering, biology, and medicine, among others. For the purpose of the present research, BP can more objectively and accurately quantify the impact of features on quotas and dynamically reflect the nonlinear impact of features on quotas at different stages [36]. Therefore, after training the BP based on pilot regional quotas and feature system data, we calculated the nonpilot regional quotas.

BP is composed of an input layer, a hidden layer, and an output layer. These three basic elements are fully connected during the whole network training. For example, for a neural network model with only one hidden layer, the process of BP neural network is mainly divided into two stages.

The first stage is the forward propagation of the signal, from the input layer to the hidden layer and finally to the output layer. Assuming that the number of samples is A, input layer has m nodes, output layer has n nodes, and hidden layer has p nodes; x_{ai} is the input/output of the input layer, $a = 1, 2, \ldots, A$, $i = 1, 2, \ldots, m$; B_{aj}, b_{aj} is the input/output of the hidden layer, respectively, $j = 1, 2, \ldots, p$; Y_{ak}, and y_{ak} are the input/output of the output layer, respectively, $k = 1, 2, \ldots, n$; y^*_{ak} is the expected label of the output layer (historical quotas); w_{ij} and w_{jk} are the weights from input to hidden and hidden to output, respectively; , θ_j and γ_k are the biases from input to hidden and hidden to output, respectively.

Table 2: Specific explanation of features.

Feature	Explanation
GDP per capita	GDP/population
Energy consumption per capita	Energy consumption/population
Degree of opening to the outside world	Total import and export/GDP
Carbon emissions per capita	Total CO_2 emissions/population
Cumulative carbon emissions per capita	Cumulative CO_2 emissions/Cumulative population
Total passenger volume	The number of passengers actually carried by the means of transport in a certain period of time
Disposable income per capita	Disposable income/population
Industrial structure	Added value of tertiary industry/GDP
Energy consumption per unit of GDP	Energy consumption/GDP
Energy consumption per unit of industrial added value	Energy consumption/industrial added value
Electricity consumption per unit of GDP	Electricity consumption/GDP
Energy structure	Coal-fired energy/energy consumption
Total freight volume	In a certain period of time, the actual weight of the cargo carried by the means of transport
The number of industrial enterprises above designated size	Including independent auditor industrial enterprises and affiliated industrial production units
Full-time equivalent of R&D personnel in industrial enterprises above designated size	The number of full-time staff plus part-time staff is converted to the total number of full-time staff based on workload
Technological level	Number of patents granted/10,000 people
Forest carbon sink	Forest area/total land area
Green area	Including park area, production green area, protective green area, subsidiary green area, and other green area
Total gas supply	Refers to the amount of natural gas supplied by gas units during the reporting period
Total supply of liquefied petroleum gas	Refers to the amount of liquefied petroleum gas supplied by gas units during the reporting period
Urbanization rate	Urban population/population

Input layer to hidden layer: determine input function, $B_{aj} = x_{ai}w_{ij} + \theta_j$, and then transform B_{aj} into b_{aj} through activation function $b_{aj} = f(B_{aj})$.

Hidden layer to output layer: determine input function, $Y_{ak} = b_{aj}w_{jk} + \gamma_k$, and then transform Y_{ak} into y_{ak} through activation function, $y_{ak} = f(Y_{ak})$, where y_{ak} is the final result.

Later, the loss function is determined and loss is calculated according to y_{ak} and y^*_{ak}. When the loss is either smaller than the set range or reaches the upper limit of the number of iterations, the model ends the training; otherwise, it enters a second stage.

The second stage is the backpropagation of the loss. The loss information is returned along the original propagation path through the learning signal. Starting from the last layer, the weight and bias are corrected layer by layer, and finally the loss is within the set range.

Weight and bias update formula:

$$\begin{aligned} W^{l*} &= W^l + \Delta W^l, \\ \theta^{l*} &= \theta^l + \Delta\theta^l, \\ \Delta W^l &= -\eta \frac{\partial \text{Loss}}{\partial W^l}, \\ \Delta\theta^l &= -\eta \frac{\partial \text{Loss}}{\partial \theta^l}, \end{aligned} \tag{3}$$

where W^{l*} and θ^{l*} indicate the updated weight and bias of the l layer, W^l and θ^l indicate the original weight and bias of the l layer, ΔW^l and $\Delta\theta^l$ represent the correction part of weight and bias of the l layer, and η is a fixed value that indicates the learning rate.

The fundamental part of backpropagation is to minimize the loss through the update of weights and biases, using the gradient descent method (actually using the chain partial derivative). The specific derivation process is given below.

The following assumptions are considered:

The activation function from the input layer to the hidden layer is $f(x) = (1/(1+e^{-x}))$ (sigmoid).

The activation function from the input layer to the hidden layer is $f(x) = x$.

Loss function is $\text{Loss} = (1/2) \times \sum_{k=1}^{n} (y^*_{ak} - y_{ak})^2 = (1/2) \times \sum_{k=1}^{n} e^2_{ak}$, $e_{ak} = y^*_{ak} - y_{ak}$.

The weights update process of the hidden-output layer:

$$\begin{aligned} \frac{\partial Loss}{\partial w_{jk}} &= \frac{\partial Loss}{\partial e_{ak}} \times \frac{\partial e_{ak}}{\partial y_{ak}} \times \frac{\partial y_{ak}}{\partial Y_{ak}} \times \frac{\partial Y_{ak}}{\partial w_{jk}} \\ &= e_{ak} \times (-1) \times 1 \times b_{aj} = -b_{aj}e_{ak}, \\ w^*_{jk} &= w_{jk} - \eta \frac{\partial Loss}{\partial w_{jk}} = w_{jk} + \eta b_{aj} e_{ak}. \end{aligned} \tag{4}$$

The weight update process of the input-hidden layer:

$$\frac{\partial Loss}{\partial w_{ij}} = \frac{\partial Loss}{\partial e_{ak}} \times \frac{\partial e_{ak}}{\partial y_{ak}} \times \frac{\partial y_{ak}}{\partial Y_{ak}} \times \frac{\partial Y_{ak}}{\partial b_{aj}} \times \frac{\partial b_{aj}}{\partial B_{ij}} \times \frac{\partial B_{ij}}{\partial w_{ij}}$$
$$= e_{ak} \times (-1) \times w_{jk} \times b_{aj}\left(1 - b_{aj}\right) \times x_{ai}$$
$$= -b_{aj}\left(1 - b_{aj}\right) w_{jk} x_{ai} e_{ak},$$
$$w_{ij}^{*} = w_{ij} - \eta \frac{\partial Loss}{\partial w_{ij}} = w_{ij} + \eta b_{aj}\left(1 - b_{aj}\right) e_{ak} w_{jk} x_{ai}. \tag{5}$$

In the same way, the updated value of the bias can be obtained:

Hidden-output layer: $\gamma_k^* = \gamma_k + \eta e_{ak}$

Input-hidden layer: $\theta_j^* = \theta_j + \eta b_{aj}(1 - b_{aj}) w_{jk} e_{ak}$

3.2.2. Key Parameters. According to the model principle previously mentioned, there are 5 key parameters that determine the learning effect of the BP: activation function, loss function, learning rate, the number of hidden layer nodes, and gradient descent algorithm.

(1) Activation function:

For the BP both, the hidden layer and the output layer need to use an activation function.

For the hidden layer, the activation function is generally a nonlinear function. The reason for this is that, if the activation function is a linear function, the output is a linear combination of the input, which is equivalent to the effect of the no hidden layer (the hidden layer is invalid). The introduction of a nonlinear function as the hidden layer activation function makes the network more powerful, increases its ability to learn complex data, and reflects the nonlinear relationship between input and output. Therefore, we introduced four nonlinear activation functions that are widely used (Tables 3 and 4).

Although ReLU has two problems, it is currently the most commonly used activation function for BP. In addition, it is the default activation function used by most feedforward neural networks.

For the output layer, the choice of its activation function depends on whether the problem is a regression problem or a classification problem. In the event, it is a classification problem, the sigmoid activation function represents a good choice; for regression problems, a linear activation function is more appropriate.

(2) Loss function:

With regards to the problem and the output layer activation function to match different loss functions,

(1) Cross-entropy function: it is suitable for binary classification problems, and the output layer activation function is sigmoid

(2) Log-likelihood cost: it is suitable for multi-classification problems, and the output layer activation function is softmax

(3) Mean square error (MSE): it is suitable for regression problems, and the output layer activation function is a linear function

(3) Learning rate:

The learning rate value is an important part of the BP, which represents the speed of information accumulation in the neural network over time, and its value is between [0, 1]. Under ideal circumstances, we would start with a large learning rate and gradually reduce the speed until the loss value no longer diverges (if the learning rate is set too low, the training progress will be very slow because only very few adjustments to the weight of the network are made. However, if the learning rate is set too high, it may bring undesirable consequences on the loss function (Figure 1)).

(4) The number of hidden layer nodes:

The number of hidden layer nodes has a great influence on the prediction accuracy of BP; if the number of nodes is too small, the network cannot perform a proper learning process, and it will need more times to train. In addition, the training accuracy is also affected. When the number of nodes is too large, the training time increases and the network will result in overfitting. However, there is no conventional formula for determining the number of nodes. Some empirical formulas are given below for reference: $l < n - 1$ $l < \sqrt{(m + n)} + a$, $l = \log_2 n$, $l \geq k \times n/(n + m)$, where n indicates the number of input layer nodes, l indicates the number of hidden layer nodes, m is the number of output layer nodes, a represents a constant between 0–10, and k corresponds to the number of samples.

In fact, the number of hidden layer nodes can be roughly calculated according to the reference formula. Later, trial and error is used to find the optimal number of nodes. Generally speaking, the BP error shows a trend where it first decreases and later increases with the increase of hidden layer nodes.

(5) Gradient descent algorithm:

We introduce six well-known gradient descent algorithms (Tables 5 and 6).

According to this analysis, there are no perfect key parameters that can suit all conditions. The appropriate selection of key parameters depends on the specific problem of study. In the present research, we studied a regression problem. Thus, we chose ReLU and $f(x) = x$ as the activation function and MSE as the loss function. Also, Adam may be appropriate as the gradient descent algorithm; however, the learning rate and the number of hidden layer nodes cannot be determined in advance. In summary, the final determination of all key parameters needs BP

Table 3: Activation functions and figures.

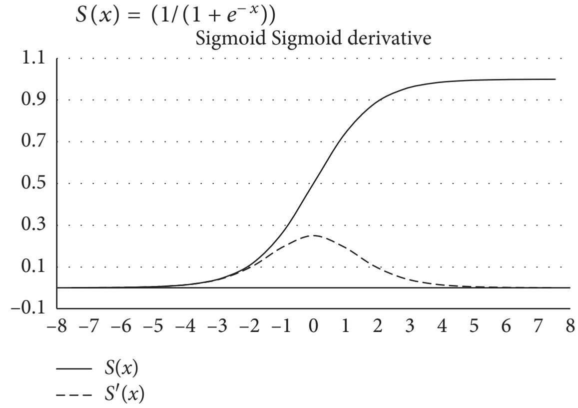
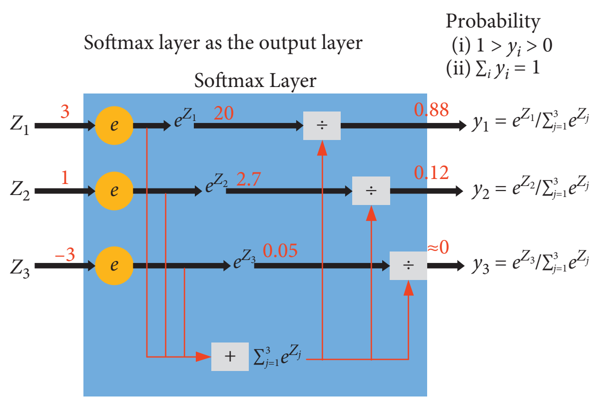
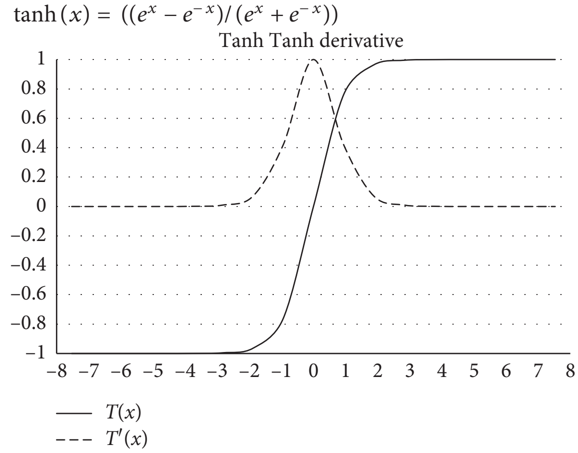
<table>
<tr><th>Name</th><th>Function</th><th>Derivative function</th></tr>
<tr><td>Sigmoid</td><td>$S(x) = (1/(1+e^{-x}))$

</td><td>$S'(x) = S(x)(1 - S(x))$</td></tr>
<tr><td>Softmax</td><td>$\text{Softmax}(Z_i) = (e^{Z_i}/\sum_{c=1}^{C} e^{Z_c}) = p_i$

</td><td>$(\partial \text{softmax}(Z_i)/\partial Z_j) = \begin{cases} p_i(1-p_i), & j = i, \\ -p_j \cdot p_i, & j \neq i. \end{cases}$</td></tr>
<tr><td>tanh</td><td>$\tanh(x) = ((e^{x} - e^{-x})/(e^{x} + e^{-x}))$

</td><td>$T'(x) = 1 - (T(x))^2$</td></tr>
</table>

Table 3: Continued.

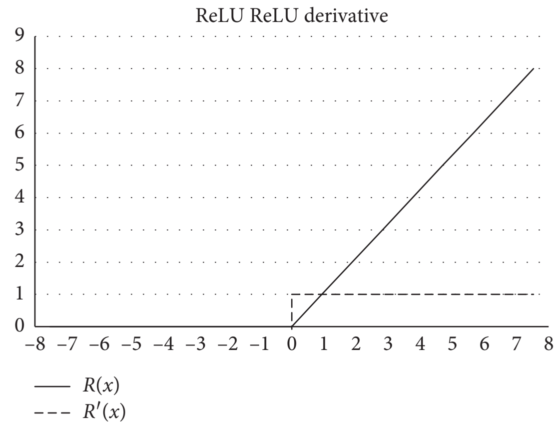
Name	Function	Derivative function
ReLU	$\mathrm{ReLU}(x) = \max(0, x)$ 	$\mathrm{ReLU}'(x) = \begin{cases} 0, & x \le 0, \\ 1, & x > 0. \end{cases}$

Table 4: Advantages and disadvantages of activation functions.

Name	Advantages	Disadvantages
Sigmoid	(1) It can smoothly map the real-number field to [0, 1] (2) Monotonically increasing, continuous derivable, and its derivative form is very simple (3) Suitable for handling binary classification problems	(1) Gradient vanishing, that is, in the process of backpropagation, the derivative will gradually become 0; thus, the parameters cannot be updated and the neural network cannot be optimized (2) Nonzero-centered: the output value of the function is always greater than 0, which will slow down the convergence speed of the model training (3) Exponentiation is relatively time-consuming
Softmax	(1) It maps the output value to (0, 1), and the sum of the mapped output value is 1 (2) It divides the entire hyperspace according to the number of classifications (3) Suitable for multiclassification problems	(1) The operation of Softmax involves the calculation of exponential function; in consequence, an “overflow problem” for computers occurs (2) Not suitable for face recognition tasks
tanh	(1) It can smoothly map the real-number field to [−1, 1] (2) Solve nonzero-centered problem (3) Suitable for handling binary classification problems	(1) Gradient vanishing, that is, in the process of backpropagation, the derivative will gradually become 0; thus, the parameters cannot be updated and the neural network cannot be optimized
ReLU	(1) Solve the gradient vanishing in the positive interval (2) The calculation is simple, no exponential calculation is required, and the activation value can be obtained with only one value (3) The convergence speed is much faster than sigmoid and tanh	(1) Nonzero-centered (2) Dead ReLU problem: it is “vulnerable” during training; when $x < 0$, the gradient is 0; the gradient of these nodes and subsequent nodes are always 0, and no longer responds to any data, causing the corresponding parameters to never be updated

training. For this reason, the determination of key parameters is provided in Section 4.

4. Empirical Results and Analysis

We focused on the allocation of regional carbon quotas, which is a regression problem, and our goal was to minimize the loss rate while ensuring that the test set results meet the realistic expected range in China. Therefore, based on experience and BP training, we chose MSE as the loss function and Adam [37] as the gradient descent algorithm in the backpropagation process, the final learning rate was 0.009, and number of iterations were 5000. Other parameters are shown in Table 7.

The total loss rate of the benchmark model was 0.02419. And the comparison between the results of the training set and the historical quotas are shown in Table 8 (benchmark model in Table 8). The test results are shown in Table 9 (benchmark model in Table 9). After adding up the historical quotas in pilot regions and the estimated quotas in the nonpilot regions, national quotas were obtained. The resulting national quotas are displayed in Table 10 (benchmark model in Table 10).

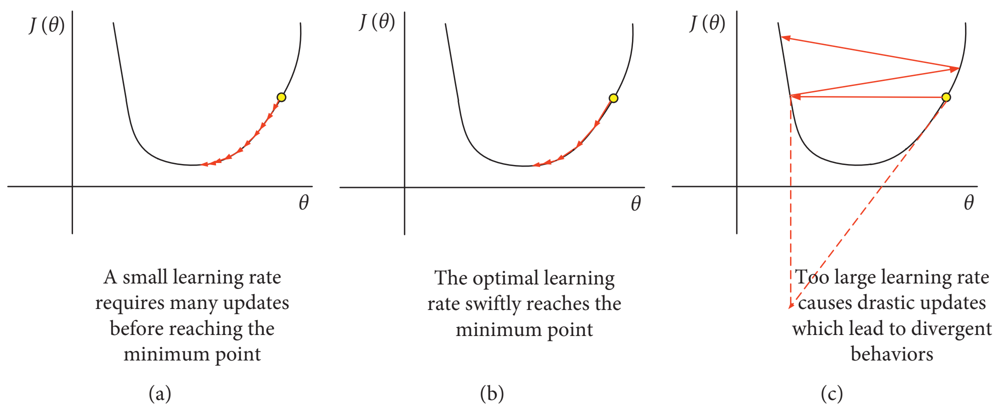


FIGURE 1: The loss function results of different learning rates.

TABLE 5: Gradient descent functions and parameter description.

Name	Gradient descent function	Parameter description
BGD	$\theta^* = \theta - \eta \cdot \widehat{g} = \theta - \eta \cdot \nabla_\theta J(\theta)$	$\widehat{g}$ denotes gradient estimate θ denotes weight
SGD	$\theta^* = \theta - \eta \cdot \widehat{g} = \theta - \eta \cdot \nabla_\theta J(\theta; x^{(i)}; y^{(i)})$	$\widehat{g}$ denotes gradient estimate θ denotes weight
Momentum	$v_t = \gamma v_{t-1} + \eta \cdot \nabla_\theta J(\theta)$ $\theta^* = \theta - v_t$	vt denotes gradient estimate θ denotes weight
AdaGrad	$\widehat{g}_{t,i} = \nabla_\theta J(\theta_i)$ $\theta^*_{t+1,i} = \theta_{t,i} - (\eta/\sqrt{\widehat{g}_{t,i} \cdot \widehat{g}_{t,i} + \varepsilon}) \cdot \widehat{g}_{t,i}$	$\widehat{g}_{t,i}$ denotes gradient estimate $\theta_{t,i}$ denotes weight
RMSprop	$g_t = \nabla_\theta J(\theta_{t-1})$ $E[g^2]_t = \gamma E[g^2]_{t-1} + (1-\gamma)g_t^2$ $\theta^*_{t+1} = \theta_t - (\eta/\sqrt{E[g^2]_t + \varepsilon})g_t$	$\gamma = 0.9E[g^2]_t$ denotes Exponential weighted average g_t denotes gradient estimate θ_t denotes weight
Adam	$g_t = \nabla_\theta J(\theta_{t-1})$ $m_t = \beta_1 m_{t-1} + (1-\beta_1)g_t$ $v_t = \beta_2 v_{t-1} + (1-\beta_2)g_t^2$ $\widehat{m}_t = (m_t/(1-\beta_1^t))$, $\widehat{v}_t = (v_t/(1-\beta_2^t))$ $\theta^*_{t+1} = \theta_t - (\eta/(\sqrt{\widehat{v}_t + \varepsilon}))\widehat{m}_t$	$\beta_1 = 0.9$ $\beta_2 = 0.999$ $\varepsilon = 10e-8$ g_t denotes gradient estimate θ_t denotes weight

TABLE 6: Advantages and disadvantages of gradient descent algorithms.

Name	Advantages	Disadvantages
BGD	The principle of gradient descent is simple	(1) Calculation is very slow (2) Difficult to handle a large dataset (3) Cannot add new data to update the model
SGD	(1) Compared with BGD, SGD training speed is faster (2) New data can be added to update the model	(1) Frequent updates may cause severe oscillations in the loss function
Momentum	(1) Consider the speed of the previous step and the new gradient (2) Can speed up the convergence and suppress the shock	
AdaGrad	(1) Compared with SGD, it adds a denominator (2) Handle the case where the number of gradient updates is small	(1) If gradient update is frequent, it may cause the subsequent gradient updates be slow or disappear
RMSprop	(1) Similar to momentum, it can reduce fluctuations (2) Overcome the problem of the sharp decrease or disappearance of gradient in AdaGrad (3) It performs better than SGD, momentum, and AdaGrad, based on the nonstationary objective function	
Adam	(1) Combine momentum and RMSProp (2) Integrate the contents of gradient descent, momentum, Adagrad, and RMSprop with certain improvement (3) Easy to use, insensitive to gradient scaling, can be used for large data, processing sparse data, easy to adjust hyperparameters, etc.	

Table 7: Parameters of the benchmark model.

Layer	Function	Activation
Import-hidden	(8, 6)	ReLU
Hidden-export	(6, 1)	$f(x) = x$

Table 8: Comparison between model results and historical quotas.

Year	Region	Historical quotas	Benchmark model		New model	
			Training results	Loss	Training results	Loss
2014	Beijing	0.5	0.5114	0.0114	0.502	0.002
	Tianjing	1.6	1.7397	0.1397	1.584	−0.016
	Shanghai	1.5	1.7795	0.2795	1.4785	−0.0215
	Hubei	3.24	3.0159	−0.2241	3.2485	0.0085
	Guangdong	3.88	3.6819	−0.1981	3.8754	−0.0046
	Shenzhen	0.33	0.3623	0.0323	0.3169	−0.0131
	Chongqing	1.3	1.3283	0.0283	1.2949	−0.0051
2015	Beijing	0.5	0.5441	0.0441	0.5144	0.0144
	Tianjing	1.6	1.7312	0.1312	1.6169	0.0169
	Shanghai	1.6	1.6728	0.0728	1.6214	0.0214
	Hubei	3.24	3.0162	−0.2238	3.1589	−0.0811
	Guangdong	4.08	3.8977	−0.1823	4.0595	−0.0205
	Shenzhen	0.33	0.3623	0.0323	0.3182	−0.0118
	Chongqing	1.25	1.419	0.169	1.3037	0.0537
2016	Beijing	0.5	0.4536	−0.0464	0.4882	−0.0118
	Tianjing	1.6	1.5162	−0.0838	1.5506	−0.0494
	Shanghai	1.5	1.4923	−0.0077	1.5099	0.0099
	Hubei	2.8	2.9384	0.1384	2.8715	0.0715
	Guangdong	4	4.0552	0.0552	4.0141	0.0141
	Shenzhen	0.3	0.3636	0.0636	0.3209	0.0209
	Chongqing	1.3	1.4176	0.1176	1.2564	−0.0436
2017	Beijing	0.5	0.3623	−0.1377	0.4946	−0.0054
	Tianjing	1.6	1.363	−0.237	1.6399	0.0399
	Shanghai	1.6	1.2582	−0.3418	1.5741	−0.0259
	Hubei	2.5	2.5891	0.0891	2.4883	−0.0117
	Guangdong	4.2	4.5011	0.3011	4.208	0.008
	Shenzhen	0.3	0.3623	0.0623	0.3043	0.0043
	Chongqing	1.3	1.2166	−0.0834	1.2883	−0.0117

Unit of quotas: 100 million tons.

Table 9: Test quotas of nonpilot regions.

Region	Benchmark model				New model			
	2014	2015	2016	2017	2014	2015	2016	2017
Hebei	4.4742	4.7213	4.6639	3.7958	4.2489	4.2512	3.9389	3.4056
Shanxi	8.7874	9.5514	9.9179	7.9321	6.7575	7.2639	7.4857	6.0955
Inner Mongolia	5.4541	5.3178	5.8367	6.0577	5.4135	5.2643	5.3196	5.8676
Liaoning	5.7035	5.2196	6.768	5.4917	5.9595	5.5083	5.8179	5.1618
Jilin	3.3295	3.1944	3.2286	2.8011	4.5732	4.651	4.5139	4.3528
Heilongjiang	4.5658	4.7666	5.3126	4.6711	5.1161	5.3088	5.5612	5.4386
Jiangsu	4.9033	4.9379	4.9276	4.5071	3.4851	3.6938	3.7096	3.6284
Zhejiang	4.3896	4.3216	4.0597	3.9379	4.2172	4.3214	4.0143	4.036
Anhui	3.6231	3.673	3.5841	3.1743	3.64	4.0833	3.8208	3.1997
Fujian	2.9149	2.9427	2.7986	2.5997	3.558	3.5671	3.3428	3.4086
Jiangxi	2.1524	2.3186	2.408	2.1446	3.5334	3.5535	3.4032	3.2111
Shandong	7.08	7.0777	7.2947	6.766	5.6471	5.6322	5.4154	5.0304
Henan	3.6061	3.7802	3.8308	3.2201	3.1582	3.2654	3.3632	2.937
Hunan	2.1693	2.3032	2.3678	2.1815	2.6993	2.8095	2.894	2.9341
Guangxi	2.0799	1.991	2.083	2.0383	3.2628	3.1519	3.0473	3.0321
Hainan	7.092	7.7761	7.7654	5.8029	6.1714	6.5658	6.6143	5.6134
Sichuan	3.0356	2.9589	2.8458	2.523	1.9639	2.1027	2.0326	1.9373

Table 9: Continued.

Region	Benchmark model				New model			
	2014	2015	2016	2017	2014	2015	2016	2017
Guizhou	3.3924	3.2882	3.5977	2.4839	3.4729	3.4067	3.4404	2.8913
Yunnan	1.1869	0.9986	1.1944	1.1038	2.698	2.6555	2.7258	2.5445
Shaanxi	4.4716	4.4987	4.5698	3.7151	2.7507	2.9126	2.689	2.6752
Gansu	3.9962	4.2983	4.357	3.2307	2.3042	2.6187	2.6927	2.0599
Qinghai	0.3623	0.3623	0.8758	0.3623	2.0744	1.9919	2.1466	2.0002
Ningxia	8.9951	9.0065	8.892	9.0385	4.9634	5.1076	5.0497	5.1049
Xinjiang	8.286	8.2076	9.1625	6.5578	3.6802	3.9403	4.236	3.8333

Unit of quotas: 100 million tons.

Table 10: National quota.

	National quota			
Benchmark model	2014	2015	2016	2017
	118.4012	120.1122	124.3424	108.137
New model	2014	2015	2016	2017
	107.6989	110.2274	109.2749	102.3993

Unit of quotas: 100 million tons.

The results of the benchmark model indicated that, during the initial stage of China's unified carbon emission market (3-4 years), the national quota will increase and, after an adaptation period, China's total quota and regional quotas will begin to decrease. This will stimulate enterprises to accelerate emission reduction and prove China's determination to achieve carbon peaks before 2030 and carbon neutrality by 2060.

5. Further Analysis: Build a Comprehensive Feature System

It is unreasonable to allocate regional quotas only considering the factor of corporate CO_2 emissions. In addition to corporate production factors, regional CO_2 emissions should also consider regional human activities and the role of forests in reducing those emissions. Thus, we believe that, in addition to the CO_2 emissions reported by enterprises, quotas in China's pilot regions should also take into account other features such as forest carbon sinks, population, and natural endowments. Based on the comprehensive factors of these three aspects, the regional quotas were determined. Therefore, we added other factors related to people and regions (see Section 3.1.2, for details) into the feature system. Subsequently, we chose MSE as the loss function, Adam as the gradient descent algorithm, and trained BP to obtain the final learning rate (0.003), number of iterations (5000), and other parameters (Table 11).

The total loss rate of the new model was 0.00089, and the results of the comparison between the training set and historical quotas are shown in Table 8 (new model in Table 8). The test results are displayed in Table 9 (new model in Table 9). After adding up the historical quotas in pilot regions and the estimated quotas in the nonpilot regions, the national quota was obtained (new model in Table 10).

Table 11: Parameters of the new model.

Layer	Function	Activation
Import-hidden	(21, 19)	ReLU
Hidden-export	(19, 1)	$f(x) = x$

While using the comprehensive feature system, the new model displayed a lower loss rate than that obtained with the benchmark model, and the calculated national quota was closer to the CO_2 emissions reported by China. These results indicated that the feature system has a certain degree of rationality and accuracy. Similarly, the calculation results of the model presented a trend, where the amount of national carbon quotas initially increased and later began to decrease.

6. Conclusions and Future Work

The whole world is expecting China to lead the economic recovery and green development after the global epidemic. It also expects China's 14th Five-Year Plan to become the Guide for green recovery. In the same year, China established a unified national carbon emissions market. This represents not only China's further exploration of the carbon emissions trading system to achieve green development but also one of the important tools for China to achieve two low-carbon goals. In addition, quota allocation is an important factor that determines the functionality of Chinese carbon market. In order to calculate other regional quotas, we trained a BP benchmark model. For this purpose, we considered historical quota data of China's 7-carbon emissions trading market pilot regions from 2014 to 2017 and selected suitable features that fit China's "bottom-up" total control method. Later, we built a feature system that included human, corporate, and regional factors, retrained the model, and recalculated quotas for other regions. The results are presented herein. First, both, the benchmark model results and the results obtained using the comprehensive feature system showed that within the sample interval, the amount of China's national carbon quotas displayed an initial increase to later decrease. Second, the model trained with the characteristic data of the feature system built in the present research displayed a lower loss rate as compared with the benchmark model. These results demonstrated that the feature system proposed in this paper fits not only the actual situation of

China's CO_2 emissions and quotas but also that the framework of the system is reasonable and accurate. Third, the feature system and training model proposed in the present article combined with the original "bottom-up" total control and post adjustment method can be used by Chinese CO_2 emission decision makers to obtain advanced predictions. We have provided the content of China's carbon emissions trading quota system, which can promote the operation of China's carbon emissions market, encourage participants in market transactions to reduce emissions, and accelerate China's low-carbon development.

Of course, we also admit that, in the future, the feature system and model proposed in this article can be further improved and perfected as follows. First, the indicators related to enterprises in the currently constructed feature system are substitute indicators because the specific transaction data and enterprise-related data of China's carbon emission market have not been unified and officially announced. Therefore, once the data is available, this part of the indicators will increase or decrease. Second, at present, China's national unified carbon emissions trading market has just started, and the main participants are enterprises, with less individual participation. At the same time, the central government has not issued a policy about people's low-carbon life. Therefore, when China's emission reduction program enters the critical stage in the future, the features related to people will increase or decrease. Third, the current BP neural network model has only three layers. In the future, with the improvement of feature data, a certain number of hidden layers may be further increased to train a model with low loss rate and stronger generalization ability.

Data Availability

Regional features (except that for Shenzhen) were obtained from the National Bureau of Statistics and China Statistical Yearbook (https://data.cnki.net/yearbook/Single/N2019110002). Features of Shenzhen were obtained from the Shenzhen Statistical Yearbook (https://data.cnki.net/yearbook/Single/N2020030065). The historical quotas data came from the 2014–2018 Beijing Carbon Market Annual Report (https://cbeex.com.cn/article/xxfw/xz/bjtscndhq/).

Conflicts of Interest

The authors declare no conflicts of interest.

Authors' Contributions

Y.Q. designed the model and the computational framework, analyzed the data, and wrote the manuscript. W.P. conducted empirical research. R.Y. collected the relevant literatures. G.P. collected the data. All authors discussed the results and contributed to the final manuscript. All authors have read and agreed to the published version of the manuscript.

Acknowledgments

This research was supported by the National Natural Science Foundations of China (71763010, 71463023, and 71803038) and Young Excellent Academic Talent Support Program of Jiangxi University of Finance and Economics.

References

[1] T. Balamurugan, L. Karunamoorthy, N. Arunkumar, and D. Santhosh, "Optimization of inventory routing problem to minimize carbon dioxide emission," *International Journal of Simulation Modelling*, vol. 17, no. 1, pp. 42–54, 2018.

[2] G. Nastase and A. Serban, "Experimental study on CO_2 capture in a residential space," *Environmental Engineering and Management Journal*, vol. 18, no. 5, pp. 1001–1011, 2019.

[3] J. Du, F. Qiao, and L. Yu, "Temporal characteristics and forecasting of PM2.5 concentration based on historical data in Houston, USA," *Resources, Conservation and Recycling*, vol. 147, pp. 145–156, 2019.

[4] L. Tang, H. Wang, L. Li, K. Yang, and Z. Mi, "Quantitative models in emission trading system research: a literature review," *Renewable and Sustainable Energy Reviews*, vol. 132, p. 110052, 2020.

[5] Y. Zhou, J. Jiang, B. Ye, Y. Zhang, and J. Yan, "Addressing climate change through a market mechanism: a comparative study of the pilot emission trading schemes in China," *Environmental Geochemistry and Health*, vol. 42, no. 3, pp. 745–767, 2020.

[6] A. Rose, B. Stevens, J. Edmonds et al., "International equity and differentiation in global warming policy," *Environmental and Resource Economics*, vol. 12, no. 1, pp. 25–51, 1998.

[7] S. Kverndokk, "Tradeable CO_2 emission permits: initial distribution as a justice problem," *Memorandum*, vol. 4, 1992.

[8] W. Y. Chen and Z. X. Wu, "Carbon emission permit allocation and trading," *Journal of Environmental Sciences (IOS Press)*, vol. 11, no. 4, p. 468, 1999.

[9] S. Kverndokk, E. Nævdal, and L. Nøstbakken, "The trade-off between intra- and intergenerational equity in climate policy," *European Economic Review*, vol. 69, pp. 40–58, 2014.

[10] C. O. Criado and J. M. Grether, "Convergence in per capita CO_2, emissions: a robust distributional approach," *Resource & Energy Economics*, vol. 33, no. 3, pp. 637–665, 2010.

[11] J. W. Park, C. U. Kim, and W. Isard, "Permit allocation in emissions trading using the Boltzmann distribution," *Physica A Statistical Mechanics & Its Applications*, vol. 391, no. 20, pp. 4883–4890, 2011.

[12] M. Ghiyasi, "Emission utilization permission based on environmental efficiency analysis," *Environmental Science and Pollution Research*, vol. 26, no. 21, p. 21295, 2019.

[13] H. H. Wang, H. C. Liu, X. J. He, and W. H. Zeng, "Allocation of carbon emissions right based on the intergenerational equity," *China Environmental Science*, vol. 36, no. 6, pp. 1895–1904, 2016.

[14] W. J. Yi, L. L. Zou, and J. Guo, "How can China reach its CO_2 intensity reduction targets by 2020? A regional allocation based on equity and development," *Energy Policy*, vol. 39, no. 5, pp. 2407–2415, 2011.

[15] Q. Dai, Y. Li, Q. Xie, and L. Liang, "Allocating tradable emissions permits based on the proportional allocation concept to achieve a low-carbon economy," *Mathematical Problems in Engineering*, vol. 2014, pp. 1–8, 2014.

[16] Y. J. Hu, X. Y. Li, and B. J. Tang, "Assessing the operational performance and maturity of the carbon trading pilot

program: the case study of Beijing's carbon market," *Journal of Cleaner Production*, vol. 161, no. 10, pp. 1263–1274, 2017.

[17] J. Zhu, H. Sun, N. Liu, D. Zhou, and F. Taghizadeh-Hesary, "Measuring carbon market transaction efficiency in the power industry: an entropy-weighted TOPSIS approach," *Entropy*, vol. 22, no. 9, pp. 973–984, 2020.

[18] T. Wang, X. Wang, Y. Gong et al., "Initial allocation of carbon emission permits in power systems," *Journal of Modern Power Systems and Clean Energy*, vol. 5, no. 2, pp. 239–247, 2017.

[19] Y. Wang, H. Zhao, F. Duan, and Y. Wang, "Initial provincial allocation and equity evaluation of China's carbon emission rights-based on the improved TOPSIS method," *Sustainability*, vol. 10, no. 4, pp. 982–996, 2018.

[20] J. K. Song, L. L. Liang, and D. P. Niu, "Allocation of carbon emission permits among provinces in China:based on environmental FCAM," *Technology Economics*, vol. 36, no. 10, pp. 100–106, 2017.

[21] Q. W. Yu and F. P. Wu, "Bi-level planning model of provincial carbon emission rights allocation from the perspective of equity and efficiency," *Soft Science*, vol. 32, no. 4, pp. 72–76, 2018.

[22] H. Q. Qian, L. B. Wu, and F. Z. Ren, "From "spurring a willing horse" to efficiency driven: a study of China's regional CO_2 emission permit allocation," *Economic Research Journal*, vol. 54, no. 3, pp. 86–102, 2019.

[23] J. Wu, Y. Fan, Y. Xia, and J. Y. Liu, "Impacts of initial quota allocation on regional macro-economy and industry competitiveness," *Management Review*, vol. 27, no. 12, pp. 18–26, 2015.

[24] S. H. Zeng and Y. Xu, "Research on China's provincial carbon emission reduction allocation efficiency based on zero-sum DEA model," *Modernization of Management*, vol. 34, no. 5, pp. 63–65, 2014.

[25] H. Jiang, X. Shao, X. Zhang, and J. Bao, "A study of the allocation of carbon emission permits among the provinces of China based on fairness and efficiency," *Sustainability*, vol. 9, no. 11, pp. 2122–2134, 2017.

[26] W. Pan and W. L. Pan, "Research on the allocation of China's provincial carbon emission rights based on energy efficiency," *Soft Science*, vol. 32, no. 6, pp. 45–48, 2018.

[27] W. Guo, T. Sun, and H. Dai, "Efficiency allocation of provincial carbon reduction target in China's "13·5" period: based on zero-sum-gains SBM model," *Sustainability*, vol. 9, no. 2, pp. 167–186, 2017.

[28] B. Ye, J. Jiang, L. Miao, and D. Xie, "Interprovincial allocation of China's national carbon emission allowance: an uncertainty analysis based on Monte-Carlo simulations," *Climate Policy*, vol. 17, no. 4, pp. 401–422, 2017.

[29] H. Alishiri, A. Taklif, H. Amadeh et al., "Efficient allocation of CO_2 emissions in selected OPEC member based on zero sum gains (ZSG-DEA) data envelopment analysis model," *Quarterly Journal of Applied Theories of Economics*, vol. 5, no. 1, pp. 213–236, 2018.

[30] Y. Huang and Y. Zhang, "Energy use and carbon emissions efficiency study of Chinese regions based on price factor," *Polish Journal of Environmental Studies*, vol. 27, no. 5, pp. 2059–2069, 2018.

[31] Y. W. Qi, "Decoupling effect and gravity center trajectory of regional economic growth and carbon emissions in China," *Modern Finance & Economics*, vol. 38, no. 5, pp. 17–29, 2018.

[32] The European Union, "EC guidance document no1 on the harmonized free allocation methodology for the EU-ETS post 2012," in *General Guidance to the Allocation Methodology*, 2011, https://ec.europa.eu/clima/policies/etshttp://ccap.org/assets/Tomas-Wyns_CCAP_Ch.pdf.

[33] F. Cucchiella, I. D'Adamo, M. Gastaldi, and S. C. Lenny Koh, "Assessment of GHG emissions in Europe: future estimates and policy implications," *Environmental Engineering and Management Journal*, vol. 19, no. 1, pp. 131–142, 2020.

[34] G. Liobikienė and M. Butkus, "The European Union possibilities to achieve targets of Europe 2020 and Paris agreement climate policy," *Renewable Energy*, vol. 106, pp. 298–309, 2017.

[35] I. Perissi, S. Falsini, U. Bardi et al., "Potential European emissions trajectories within the global carbon budget," *Sustainability*, vol. 10, no. 11, pp. 4225–4239, 2018.

[36] J. B. Kim, "Implementation of artificial intelligence system and traditional system: a comparative study," *Journal of System and Management Sciences*, vol. 9, no. 3, pp. 135–146, 2019.

[37] V. Laparra, A. Berardino, J. Ballé, and E. P. Simoncelli, "Perceptually optimized image rendering," *Journal of the Optical Society of America A*, vol. 34, no. 9, pp. 1511–1525, 2017.