sustainability

MDPI

Article

Iron Ore Price Prediction Based on Multiple Linear Regression Model

Yanyi Wang [1,2,3], Zhenwei Guo [1,2,3,*], Yunrui Zhang [4], Xiangping Hu [5] and Jianping Xiao [1,2,3,*]

[1] School of Geosciences and Info-Physics, Central South University, Changsha 410083, China
[2] Key Laboratory of Metallogenic Prediction of Nonferrous Metals and Geological Environment Monitoring, (Central South University), Ministry of Education, Changsha 410083, China
[3] Hunan Key Laboratory of Nonferrous Resources and Geological Hazard Exploration, Changsha 410083, China
[4] ESSEC Business School, 95021 Paris, France
[5] Industrial Ecology Programme and Department of Energy and Process Engineering, Norwegian University of Science and Technology (NTNU), 7034 Trondheim, Norway
* Correspondence: guozhenwei@csu.edu.cn (Z.G.); jpxiao@csu.edu.cn (J.X.)

Abstract: The fluctuation of iron ore prices is one of the most important factors affecting policy. Therefore, the accurate prediction of iron ore prices has significant value in analysis and judgment regarding future changes in policies. In this study, we propose a correlation analysis to extract eight influencing factors of iron ore prices and introduce multiple linear regression analysis to the prediction. With historical data, we establish a model to forecast iron ore prices from 2020 to 2024. Taking prices in 2018 and 2019 as samples to test the applicability of the model, we obtain an acceptable level of error between the predicted iron ore prices and the actual prices. The prediction model based on multiple linear regression has high prediction accuracy. Iron ore prices will show a relatively stable upward trend over the next five years without the effects of COVID-19.

Keywords: iron ore price; multiple linear regression; prediction; geophysics prospecting

Citation: Wang, Y.; Guo, Z.; Zhang, Y.; Hu, X.; Xiao, J. Iron Ore Price Prediction Based on Multiple Linear Regression Model. *Sustainability* **2023**, *15*, 15864. https://doi.org/10.3390/su152215864

Academic Editor: Grigorios L. Kyriakopoulos

Received: 12 September 2023
Revised: 3 October 2023
Accepted: 20 October 2023
Published: 12 November 2023

1. Introduction

As iron ore is an important industrial basic resource, a guaranteed capability is particularly important. As iron ore is the most important raw material for steel, countries around the world are highly dependent on it for production. According to the USGS (2017) report, the estimated global resource of crude iron is 170 billion tons, while iron ore is about 82 million tons. China ranks third in the production of iron ore after Australia and Brazil [1].

The price trend of iron ore plays an important role in corporate earnings and countries' economic development [2]. Iron ore price prediction holds importance for economic and environmental sustainability. Fluctuations in iron ore prices have a direct and profound impact on various industries, such as steel production and construction [3]. Iron ore price affects production costs and economic stability [4]. Accurate forecasts can help businesses and governments make informed decisions to mitigate economic risks. Moreover, these forecasts assist in responsible resource management, preventing environmental overexploitation and degradation, which aligns with environmental sustainability goals. By allowing companies to plan and implement cost-effective and environmentally friendly practices, iron ore price prediction plays a vital role in shaping a more sustainable and balanced economy and environment.

However, iron ore price is determined by several factors, which are analyzed based on the economic characteristics of the mineral resources industry [5]. For instance, gross national income (GNI), per capita gross domestic product (GDP), tariffs, fixed-asset investment, steel production, scrap consumption, iron ore production, and iron concentrate production costs can lead to iron ore price fluctuations [6]. As well as these routine factors, the Russo–Ukrainian War and COVID-19 are the most important factors currently affecting global iron ore prices. Forecasting iron ore prices is a rather difficult task. To maintain the stability of the iron ore market and prevent the risk of iron ore price fluctuations, accurate prediction models of iron ore prices are crucial for policymakers [7].

Scientists are constantly experimenting with predictions of iron ore prices. Trend methods are common models of price prediction. Due to the uncertainty of influencing factors, trend methods are not able to predict prices with adequate accuracy. The time series method, taking into consideration precious metal prices and exchange rates, has been applied for dynamic price analysis [8]. The Monte Carlo simulation approach has also been applied to forecast long-term iron ore prices [9], but this method has usually led to overestimation. Systems thinking and the scenario planning method have been used to forecast iron ore prices [10]. This method analyzes the main factors related to price. Then, the future price is estimated using the dynamic model.

Moghaddam et al. [11] suggested a model for predicting the monthly price of iron ore using artificial neural networks. The predictive method is complex and requires a large amount of data. Considering the chaos of iron ore price forecasts, Mobtaker et al. [12] introduced the entropy method for prediction. This method considers a historical survey of iron ore price behavior and the effective parameters of price fluctuation. The parameters are weighted, and the future trend of iron ore prices is predicted. Weng et al. [13] proposed a BIC-GA-RELM model to predict iron ore prices, which provided a new method with high accuracy. The Geometric Brownian Motion (GBM) model is one of the most used mathematical forecast models. Ramos et al. [14] simulated the forecast price of iron ore with the GBM method based on a historical iron ore commodity price series and compared it to the predictions made by financial agents. However, the model also has limitations on the prediction of short-term trends. Therefore, it should be used with other quantitative and qualitative models [11].

Established on the assumption of historical repetition, time series prediction is more suitable for stable historical data, limited in the domain of price forecasting. As proposed by Deng [15], the gray system shows excellent results when the time series is short, the statistical data are few, and the information is incomplete [16]. Gray prediction is essentially a trend prediction, which has certain requirements for the original data. However, it is often difficult for data on price prediction to meet these requirements. Therefore, gray prediction is limited in the field of price prediction.

To capture the nonlinear characteristics hidden in iron ore prices, artificial intelligence models have recently been favored by scholars because of their powerful generalization capabilities [17]. A typically and commonly used method is that of artificial neural networks (ANNs). Machine learning-based methods provide a powerful tool for nonlinear iron ore price analysis. Guo [18] established a neural network model for predicting iron ore prices, which considers six major factors: marginal cost of iron ore, market interest rate, inflation rate, domestic iron ore consumption, domestic iron ore supply, and internationally agreed-upon prices. In addition, the ANN methods require a great deal of historical data during training and at the validation stage for future data prediction. Therefore, in the case of limited data volume, the ANN methods have difficulty demonstrating good predictive performance [19].

The regression analysis prediction method is a classical statistical method for prediction and has a wide range of applications and various types. Regression analysis, as an important market forecasting method, has been widely used in metal price forecasting [20,21]. Via statistical analysis and multiple linear regression methods, Zhu et al. [22] constructed a univariate linear regression equation with iron ore sea freight price as the dependent variable and China's GDP growth rate as the independent variable. Gu et al. [23] emphatically analyzed the relationship between GDP, fixed-asset investment, steel output, scrap consumption, raw ore supply, the production cost of iron concentrate powder, and iron ore prices and constructed a multiple linear regression prediction model to predict iron ore prices from 2016 to 2020.

Regression analysis is a classical statistical method for making predictions, and its related research has been developing in recent years. Taking full account of the chaotic characteristics of iron ore price data, the regression method is proposed for predicting iron ore prices in this paper. In addition to considering some of the influencing factors proposed by Gu et al. [23], GNI and tariffs are added to the model in this paper. The steps of the regression analysis prediction method can be summarized as follows (Figure 1).

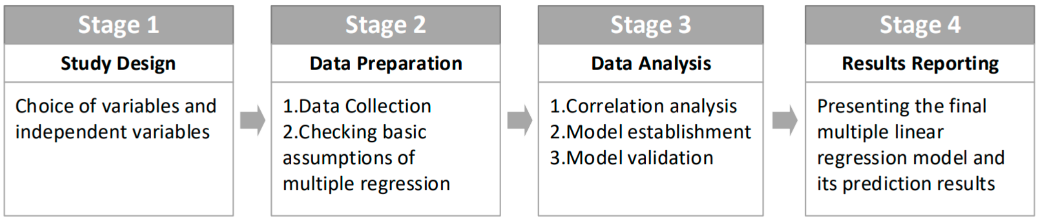

Figure 1. The strategy of regression prediction method.

In this study, we aim to establish a multiple regression prediction model for iron ore prices. We construct a regression equation with eight influencing factors as independent variables and apply historical iron ore price datasets from 1998 to 2019 to the proposed model. Our multiple regression method produces outputs with accurate predicted values. The innovations and potential contributions of this article can be summarized as follows: (1) By considering many influencing factors of iron ore prices, an iron ore price prediction model based on multiple regression is proposed, which effectively improves the accuracy of iron ore price prediction. (2) The model, based on multiple regression, is implemented, and the price of iron ore from 2020 to 2024 and the annual growth rate is forecast. However, COVID-19 and the Russian–Ukrainian War have changed the world significantly. As we live in a peaceful country (China), we hope that people will focus on peace and development. Sustainable development depends not only on a peaceful environment but also on global connectivity and communication.

The paper is organized as follows: The main theoretical methods and proposed multiple regression model are explained in Section 2. In Section 3, the results of the prediction model are presented. A detailed discussion of the results is given in Section 4. Finally, the conclusions of the paper are presented in Section 5.

2. Methods

The multiple linear regression method is constructed with the following steps. First, we identify the most important influencing factors influencing iron ore prices from several perspectives. Second, we build a regression model based on the identified factors. Finally, we evaluate the performance of the proposed model to ensure that it works.

2.1. Influencing Factors of Iron Ore Prices

A single factor is not enough for a regression model. Through economic analysis and characteristics analysis of the mineral resources industry, we can determine eight factors that are relevant to iron ore prices.

GNI is an important indicator of national income and labor value. Its change was basically consistent with the price of iron ore before 2006. After 2006, its growth remained stable while the price of iron ore fluctuated widely (Figure 2a). Using correlation analysis, the correlation coefficient between GNI and iron ore prices is around 0.7522.

GDP is an indicator used to measure the economic level of a country or region. In this paper, GDP is replaced by real GDP per capita to take into account population growth. Its relationship with the price of iron ore is similar to that between GDP and the price of iron ore (Figure 2b). A correlation analysis of GDP and iron ore prices gives a correlation coefficient of 0.7609.

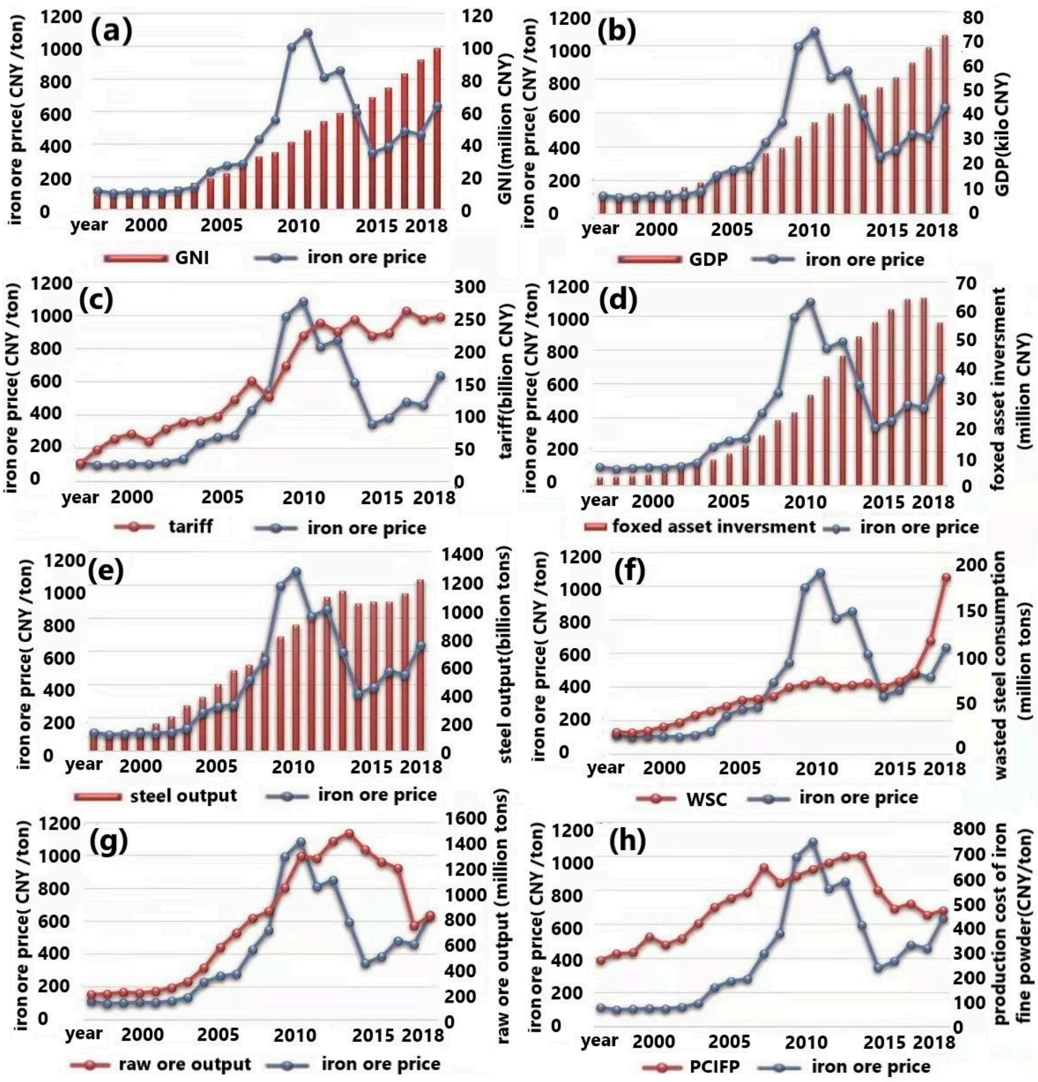

Figure 2. The change chart of iron ore price and eight factors: (**a**) GNI; (**b**) GDP; (**c**) Tariff; (**d**) fixed-asset investment; (**e**) steel output; (**f**) waste steel consumption; (**g**) raw ore output of iron ore; and (**h**) production cost of iron fine powder.

Macroeconomics plays an invaluable role in commodity trading and is a very important indicator that affects commodity prices. We can easily understand that when the economic situation is good, downstream industries such as the construction industry and the automobile machinery industry will have good development, which can increase the demand for iron ore and thus increase iron ore prices. Therefore, the above two factors can be seen as specific macroeconomic manifestations that affect iron ore prices [24,25]. Although these two factors do not perform well in terms of correlation coefficient, their impact on iron ore prices cannot be ignored. This phenomenon may be caused by the lagged effect of macroeconomic indicators on iron ore prices. Therefore, we consider these two factors in predicting iron ore prices.

The iron ore market is dominated by a few giants, meaning that international factors can significantly impact iron ore prices. Global prices directly affect domestic iron ore prices. Additionally, China primarily relies on imports because the volume and quantity of its imports are exceedingly high. Thus, tariff policies also influence its price [26]. Via correlation analysis, we calculated a correlation coefficient of 0.8524 between tariffs and iron ore prices. This indicates a strong correlation between the two variables (Figure 2c).

In addition to the above three factors, fixed-asset investment, steel production, waste steel consumption, raw iron ore production, and production cost of iron are significant determinants of iron ore prices [23,27].

The steel industry and construction industry have a close relationship with fixed assets. Therefore, fixed-asset investment is also an essential factor that affects the price of iron ore (Figure 2d). The correlation coefficient between fixed assets and the price is 0.7201.

As the iron ore industry is upstream of the iron and steel industry, steel production directly impacts iron ore prices. The trend of the two industries is similar from 1998 to 2010. Iron ore prices fluctuated considerably from 2010 to 2014, but the trend has become consistent again (Figure 2e). The calculated correlation coefficient between steel production and iron ore prices is 0.7729.

Waste steel is a substitute for iron ore. Its consumption affects the prices of iron ore [28]. Iron ore is primarily used to produce steel, and as a raw material for steelmaking, waste steel offers unique advantages. Increases in scrap steel consumption have led to higher demand. The downstream steel industry demand has remained strong, and the demand for iron ore has increased. Consequently, with unaltered supply, prices rise. The correlation coefficient is 0.6069 between waste steel consumption and the iron ore price trend (Figure 2f).

Based on the theoretical analysis of supply and demand, an increase in raw iron ore production is linked to a decrease in iron ore prices. The patterns of iron ore prices and raw ore output from 2013 to 2017 align with this principle. Although there are deviations in other periods, analysis shows a strong correlation with a coefficient of 0.8399 (Figure 2g).

Since cost determining prices is an objective law in commodity economics, the production cost of iron concentrate has a close relationship with the prices of iron ore. The comparison of the two figures shows that the trend of production cost of iron concentrate bears a strong resemblance to that of the price increase of iron ore (Figure 2h). Additionally, the high correlation coefficient of 0.8352 also proves the strong correlation between the two variables.

To sum up, these eight factors and their correlation coefficients were collected in Table 1.

Table 1. Summary of influencing factors and their correlation coefficients.

Index	GNI	GDP	Tariff	Fixed Investment	Steel Output	Waste Steel Consumption	The Raw Ore Output of Iron Ore	Production Cost
Correlation Coefficient	0.75	0.76	0.85	0.72	0.77	0.61	0.84	0.84

2.2. Data Preparation

Due to the missing data for certain influencing factors from 1980 to 1997, we have opted to utilize historical data ranging from 1998 to 2019 for our analysis.

As mentioned above, numerous factors influence iron ore prices, with reasons for increases and decreases varying. The price volatilities occur periodically, with ascents and descents following in intermittent patterns. It ascends during a certain period and then shows a downward trend at a certain time, and this intermittent situation continues. Figure 3 illuminates the changes in iron ore prices from 1998 up to 2019, with each period having its own distinct reasons and characteristics.

Prior to China's involvement in the negotiation of iron ore prices, there was a steady upward trend in iron ore prices during Period 1 (pre-2004). The key period for price analysis is Period 2 (2004–2008), which witnessed a significant increase in price due to the economic activities of BRICS nations (Brazil, Russia, India, China, and South Africa), as well as heightened demand for steel. In 2004, China joined iron ore price negotiations. Since then, prices have been steadily increasing due to the growth of China's iron and steel industry. Specifically, there was a rise of 71.5% and 65% in 2005 and 2008, respectively. Region 3 (2008–2011) shows a more significant price increase than Period 2. The increase in iron ore prices was significantly influenced by the global economic downturn, growth in the global economic situation, as well as the economic roles played by Southeast Asian countries and BRICS. During Period 4 (2011–2015), due to high production at low cost to big companies, steel demand decreased, and iron ore surplus was high. Then, the price decreased. Driven by China's economic recovery and supply-side reform, iron ore prices began to show an upward trend again during Period 5 (2015–2019).

To forecast iron ore prices, we utilized the data of the above eight influencing factors of iron ore prices from 1998 to 2019 as historical data (Table 2).

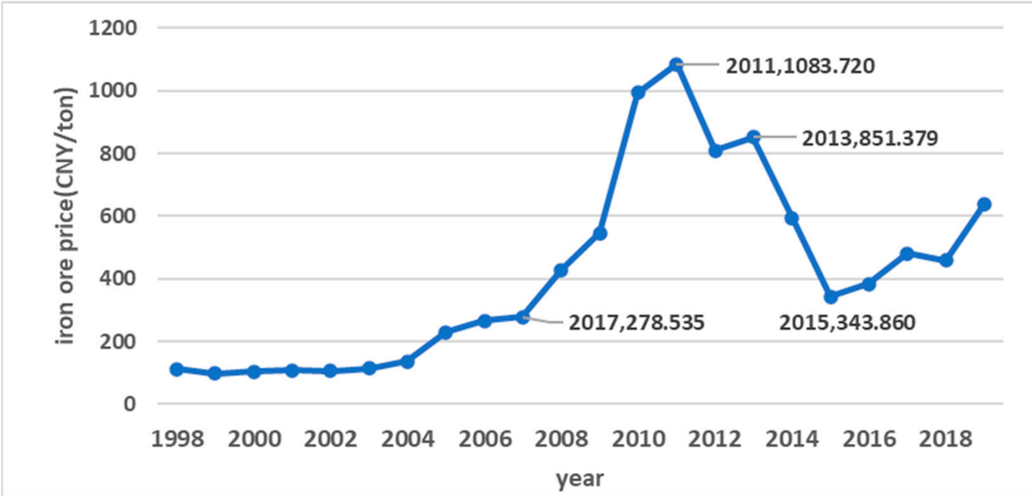

Figure 3. Iron ore price in China between 1998 and 2019 (Source: China Statistical Yearbook).

Table 2. Data on influencing factors of iron ore prices from 1998 to 2019.

Year	GNI (Billion CNY)	GPD (Billion CNY)	Tariff (Billion CNY)	Fixed Investment (Billion CNY)	Steel Output (Million Ton)	Waste Steel Consumption (Million Ton)	The Raw Ore Output of Iron Ore (Million Ton)	Production Cost of Iron Fine Powder (CNY/Ton)
1998	8380.0	6830.0	31.3	2840.0	107.0	27.5	206.0	260.0
1999	8940.0	7200.0	56.2	2990.0	121.0	26.7	209.0	285.0
2000	9910.0	7910.0	75.0	3290.0	131.0	29.2	224.0	290.0
2001	10,900.0	8690.0	84.1	3720.0	161.0	34.4	217.0	352.0
2002	12,000.0	9480.0	70.4	4350.0	193.0	39.2	231.0	320.0
2003	13,700.0	10,600.0	92.3	5560.0	241.0	48.2	261.0	345.0
2004	16,100.0	12,500.0	104.0	7050.0	320.0	54.3	310.0	404.0
2005	18,600.0	14,300.0	107.0	8880.0	378.0	60.0	420.0	468.0
2006	21,900.0	16,700.0	114.0	11,000.0	469.0	67.2	588.0	502.0
2007	27,100.0	20,400.0	143.0	13,700.0	566.0	68.5	707.0	525.0
2008	32,100.0	24,000.0	177.0	17,300.0	605.0	72.0	824.0	624.0
2009	34,800.0	26,100.0	148.0	22,500.0	694.0	83.1	880.0	563.0
2010	41,000.0	30,700.0	203.0	25,200.0	803.0	86.7	1070.0	589.0
2011	48,300.0	36,200.0	256.0	31,100.0	886.0	91.0	1330.0	615.0
2012	8380.0	6830.0	31.3	2840.0	107.0	27.5	206.0	260.0
2013	8940.0	7200.0	56.2	2990.0	121.0	26.7	209.0	285.0
2014	9910.0	7910.0	75.0	3290.0	131.0	29.2	224.0	290.0
2015	10,900.0	8690.0	84.1	3720.0	161.0	34.4	217.0	352.0
2016	12,000.0	9480.0	70.4	4350.0	193.0	39.2	231.0	320.0
2017	13,700.0	10,600.0	92.3	5560.0	241.0	48.2	261.0	345.0
2018	16,100.0	12,500.0	104.0	7050.0	320.0	54.3	310.0	404.0
2019	18,600.0	14,300.0	107.0	8880.0	378.0	60.0	420.0	468.0

2.3. Multiple Linear Regression Model

Multiple linear regression is a valuable statistical technique for examining the association between a group of independent variables and one dependent variable by fitting linear models to the observed data.

In this case, we regard the eight factors listed in Table 1 as independent variables. The equation of the proposed multiple linear regression model is outlined below:

$$\begin{aligned} F(t) = b_0 + b_1N(t) + b_2P(t) + b_3T(t) + b_4L(t) + b_5C(t) + b_6U(t) \\ +b_7V(t) + b_8E(t) \end{aligned} \quad (1)$$

where N(t)—Gross National Income (GNI), P(t)—Gross Domestic Product (GDP), T(t)—Tariff, L(t)—Fixed-Asset Investment, C(t)—Steel Production, U(t)—Waste Steel Consumption, V(t)—Raw Iron Ore Output, E(t)—The Production Cost of Iron Concentrate.

The R^2 (Coefficient of Determination) is 0.920, which indicates a good fit for the model. The Durbin–Watson value is 1.400, which indicates the presence of positive autocorrelation among the residuals, although it suggests the error terms in the model are not entirely independent.

In our analysis of the normal distribution, we constructed a Probability–Probability (P–P) plot (Figure 4) that exhibits a striking proximity to a nearly 45-degree diagonal line. This plot proves the normality of the residuals.

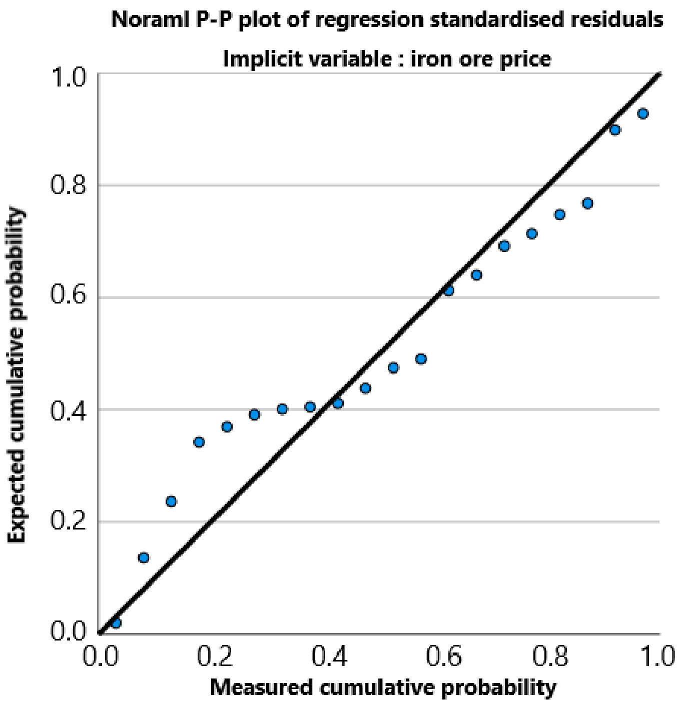

Figure 4. Normality test using normal Probability–Probability (P–P) plot for multiple linear regression.

We observed a correlation between the independent variables resulting from multicollinearity, and multicollinearity has an impact on the estimated coefficients. Our focus is to build a model for iron ore price prediction, and multicollinearity does not influence the prediction [29,30].

In accordance with the coefficients presented in Table 3, the regression model equation can be formulated as Equation (2):

$$\begin{aligned} F(t) = -488.81 - 0.038683N(t) + 0.59416P(t) + 0.12591T(t) - 0.0042319L(t) + \\ 0.00021895C(t) - 0.027253U(t) + 0.0032307V(t) - 0.74026E(t) \end{aligned} \quad (2)$$

Table 3. Summary of the estimates of regression coefficients.

Coefficients	Regressed Value
b_0	−488.81
b_1	−0.038683
b_2	0.059416
b_3	0.12591
b_4	−0.0042319
b_5	0.00021895
b_6	−0.027253
b_7	0.0032307
b_8	−0.740260

3. Results

3.1. Accuracy Evaluation

To assess the effectiveness of the proposed model, we conducted a test by estimating iron ore prices in years with previous iron ore price data. We then compared the estimated value and the actual value to calculate the model's performance error. We selected 2018 and 2019 as the test subjects based on this approach. The results of the prediction and error are presented in Table 4 and Figure 5. They illustrate the comparison of predicted and actual iron ore prices. Based on the graph and table, the predicted values closely approximate the true response values. The resulting outputs of the testing are considered satisfactory and confirm the reliability of the model.

Table 4. The forecast price and error for 2018 and 2019.

Year	Predicted Value	Actual Value	Error
2018	414.44	459.15	−9.74%
2019	678.80	636.11	6.71%

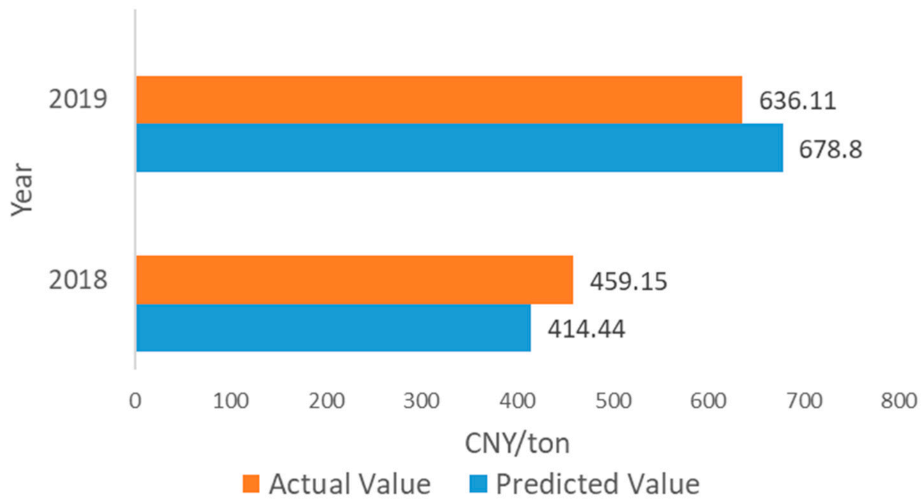

Figure 5. Forecast price and error for 2018 and 2019.

The predicted value of iron ore price in 2018 and 2019 is 414.44 CNY/ton and 678.79 CNY/ton, respectively. Compared with the actual value of 459.15 CNY/ton and 636.11 CNY/ton, the error is less than 10%. Therefore, the model has proven effective and accurate.

3.2. Forecasting Experiments

In this section, the multiple linear regression model is employed to forecast the actual case. Due to the large fluctuation of iron ore prices, it is not reliable to predict many years in the future. Therefore, this paper concentrates on predicting iron ore prices within the coming five years.

3.2.1. Independent Variable Prediction

Many prediction studies have proposed mathematical models to predict independent variables, including GDP and fixed-asset investment. Wu and Qiu [31] utilized a multi-dimensional gray model to accurately predict China's GDP in 2020. Gu [23] made predictions for variables including GDP, fixed-asset investment, and steel production between 2016 and 2020 and developed a five-year forecast for iron ore prices. Due to the regulation and control of the macroeconomy, the growth rates of influencing factors fluctuate greatly. If singular historical data are used to establish a function model, it may result in significant deviations. Therefore, the influencing factors are predicted based on the historical growth trend of each factor and regulatory control of the national macroeconomy. Following this approach, this paper predicts the values of the chosen independent variable for the next five years and calculates the change rates.

Based on Table 5 and historical data analysis, the forecasted factors that will impact iron ore prices over the forthcoming five-year period are shown in Table 6. Table 6 clearly shows the ascending nature of the diverse elements impacting iron ore prices, in addition to the anticipated tariff fluctuations.

Table 5. Annual growth rates of individual variables over the next five years.

Independent Variable	GNI	GDP	Tariff	Fixed Investment	Steel Output	Waste Steel Consumption	The Raw Ore Output of Iron Ore	Production Cost
Annual growth rate	7%	7%	*	9.5%	7%	3%	2%	4%

* Where tariffs are greatly affected by policies, the data fluctuates greatly when calculating the historical growth rate, reaching 99.1% at the highest. Therefore, the annual tariff growth rate in the next five years is predicted to be 1.5%, −5%, 3%, 5%, and 1.5%.

Table 6. Prediction of Factors Affecting Iron Ore Price in the Next Five Years.

Year	2020	2021	2022	2023	2024
GNI (billion CNY)	106,000	113,000	121,000	130,000	139,000
GDP (billion CNY)	7570	8100	8670	9280	9930
Tariff (billion CNY)	332	349	359	366	372
Fixed Investment (billion CNY)	61,400	67,200	73,600	80,600	88,300
Steel Output (million tons)	1290	1380	1480	1580	1690
Waste Steel Consumption (million tons)	227	233	240	248	255
The raw ore output of iron ore (million tons)	861	878	896	914	932
Production cost of iron fine powder (CNY/ton)	473	492	511	532	553

3.2.2. Prediction Results

The predicted value of iron ore prices in the following five years can be calculated by inputting the corresponding independent variable values into the regression equation. The resulting values are presented in Table 7.

Table 7. Prediction of iron ore prices and its growth rate in the following five years.

Year	2020	2021	2022	2023	2024
Prediction of iron ore prices (CNY/ton)	750.74	785.46	808.12	821.47	826.55
Growth rate (%)	18.02	4.42	2.88	1.63	0.62
Iron ore prices (CNY/ton)	750.74	1239	838.56	820.80	/

Figure 6 indicates an increase in iron ore prices from 2020 to 2024, with an average growth rate of 5.51%. The minimum price of iron ore in the next five years is expected to be 750.74 CNY/ton in 2020, while the maximum predicted price is 826.55 CNY/ton in 2024.

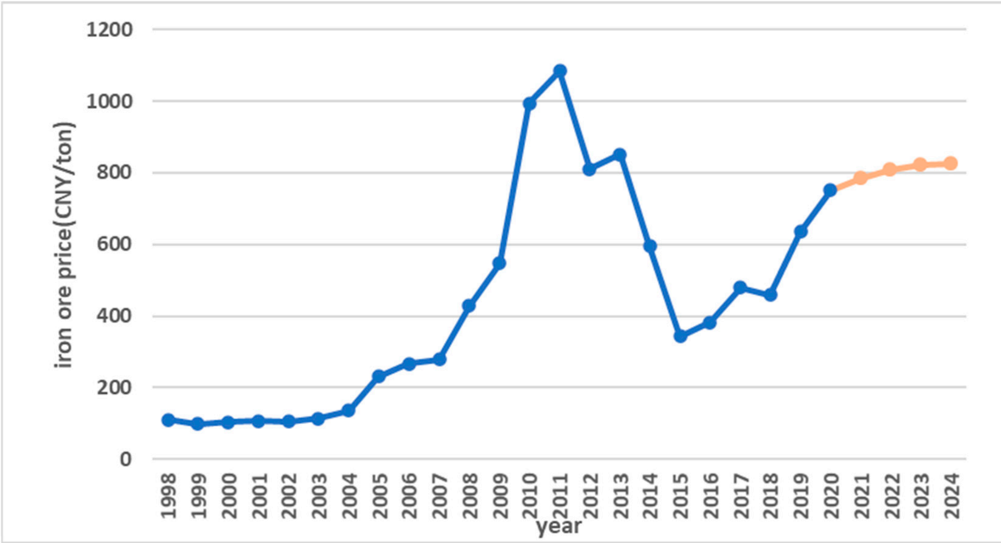

Figure 6. Trend of iron ore prices from 1998 to 2024, values for 2020–2024 are predicted (line in orange).

Figure 6 clearly describes the fluctuation of iron ore prices. Table 8 presents the correlation coefficients between the estimated value of iron ore prices and the anticipated value of each factor within the next five years. The correlation coefficients listed in Table 8 indicate a significant correlation between the iron ore price and the eight influencing factors. It is affirmed by the presence of a correlation coefficient with a value above 0.9 for every independent variable. With a certain increase in GNI, GDP, fixed investment, steel production, scrap consumption, iron ore production, and iron concentrate production costs, the price of iron ore can be expected to rise at a certain rate.

Table 8. The correlation coefficient between the predicted value of iron ore price and the predicted value of each factor.

Variables	Correlation Coefficient
GNI (billion CNY)	0.944
GDP (billion CNY)	0.944
Tariff (billion CNY)	0.996
Fixed Investment (billion CNY)	0.938
Steel Output (million tons)	0.943
Waste Steel Consumption (million tons)	0.950
The raw ore output of iron ore (million tons)	0.952
Production cost of iron fine powder (CNY/ton)	0.948

4. Discussion

Due to production factors, including mine safety accidents, adverse weather conditions, and geographical conflicts in recent years, the price of iron ore has fluctuated significantly. From a macro perspective, China's economic development remains relatively stable despite trade frictions and weak global economic growth. The steel industry is widely regarded as crucial to a country's prosperity. China's steel production and demand have not yet reached the same level as in developed countries. China is in the middle of rapid industrialization, with a surge in demand for steel [32]. Iron ore is a highly essential raw material for the steel industry that serves as vital for its development [33]. Iron ore has an important effect on the growth of China's economy [34]. Therefore, iron ore resources play a crucial role in the development of the national economy [35]. With the ongoing progress of China's supply-side reform, iron ore is expected to maintain strong demand, and rising prices will likely remain a long-term trend.

From the microscopic perspective, the implementation of environmental protection policies that align with the "innovation–coordination–green–opening–sharing" development pattern proposed by the supply-side reform has led to an improvement in social responsibility performance standards for mining enterprises. The absence of technical support and procedure development for ecological regulation in mining may escalate the expenditure of rectifying environmental damage after mining if external diseconomies emerge. Assuming all other external factors are consistent, the overall expenses associated with iron ore mining increase. Companies may shift some of the economic pressure onto the demand side, leading to price increases. Iron ore mining enterprises face increased operational risks and instability in future profits. In this way, mining enterprises tend to demand high-risk compensation rates. In particular, when the market is volatile and risky, the necessary rate of return required by mining companies may be higher, making iron ore prices rise further. The estimates for return and risk estimates of investments in projects for mineral resource exploitation are usually closely tied to the iron ore price [36]. Therefore, the prediction interval in this study can be considered a continuation of the fifth period discussed in Section 2.2.

Multiple linear regression methods are often preferred over other methods due to their interpretability, low computational cost, and easy-to-implement and maintain modeling approach [37]. Compared to Gu's [23] multiple regression model, this model takes into account additional factors that influence the price of iron ore and improves interpretive ability. However, these findings have limitations. One limitation is the wide timeframe of the data search, combined with limited available data for the regression analysis. These influences may affect the generalizability and accuracy of the prediction. The accuracy of independent variable prediction plays a pivotal role in the process of predicting iron ore prices using multiple linear regression. When it is inaccurate, the accuracy of the predicted values of iron ore will also be affected. Additionally, despite the significant correlation between the eight independent variables considered and the iron ore price, changes in iron ore prices are dependent on various economic, political, and other factors working in conjunction. To enhance predictive accuracy, our future research may encompass additional factors.

Technological innovation has a certain impact on iron ore prices within the demand industry. On one hand, innovation drives improvements in steel smelting technology. It may lead to increased steel production, expanded advantages in economies of scale, and reduced production costs. On the other hand, the advancement of technology has stimulated the growth of associated manufacturing industries, potentially leading to an increase in waste steel. As a substitute for iron ore, the quantity of waste steel is steadily rising in China's present development situation. Simultaneously, its utilization rate may also profit from technological breakthroughs in recycling and be further enhanced. The price of iron ore may be reduced by its substitution effect. Meanwhile, in a state of stable demand for finished iron ore products, the implementation of technological innovations could potentially reduce the overall demand for iron ore to some degree due to the resulting production savings. This could result in a reduction of iron ore prices without any change in supply. However, the method explained in Section 2 is straightforward to follow and produces reliable results. This approach is a helpful tool for mitigating investment risks and uncertainties.

Despite being geographically distant from the Russo–Ukrainian War, the iron ore market in China has been impacted by COVID-19. In this paper, we assume a prediction of the price without factoring in the Russo–Ukrainian War and COVID-19. The iron ore price in 2021 has exceeded predictions, possibly due to the impact of the COVID-19 outbreak. Upon analyzing the anticipated results against actual data, we speculate that the discrepancy in iron ore prices from the projected figures in 2021 was influenced to some degree by the epidemic. Amid COVID-19, the economies and markets of major global powers have experienced a slowdown or decline in growth, resulting in reduced or even negative GDP growth rates in numerous countries. In this context, as one of the four pillars driving GDP growth, consumer demand was inevitably affected. This was specifically manifested as a contraction in consumer demand, which led to a decrease in the number of consumer goods purchased by the entire society, including automobiles. Iron ore is a vital raw material for various consumer goods. Its demand has been declining due to a slump in its downstream industries, resulting in price fluctuations. In addition, following the outbreak of the pandemic, countries globally implemented frequent control measures to restrain the spread of the contagion, which significantly impacted iron ore trading. Restrictions on imports and exports of iron ore and variations in transportation costs have caused a disparity between the predicted and actual price of iron ore.

5. Conclusions

Considering the volatile fluctuations in iron ore prices, we propose a multiple linear regression model to predict the price of iron ore. The results reveal that GNI, GDP, Tariff, Fixed-Asset Investment, Steel Production, Waste Steel Consumption, Raw Iron Ore Output, and the Production Cost of Iron Concentrate are critical influencing factors of the price of iron ore. The proposed model incorporates the above eight factors as independent variables. The model is adequately developed and verified using actual data based on the data of the eight independent variables from 1998 to 2019.

Several conclusions and implications from the study are summarized as follows:

(1) In this paper, correlation analysis was used in the influencing factors of the price of iron ore. Then, we identified various factors that are highly correlated with iron ore prices, including not only fundamental factors (supply and demand) but also non-fundamental factors such as GNI, GDP, Tariff, Fixed-Asset Investment, Steel Production, Waste Steel Consumption, Raw Iron Ore Output, and the Production Cost of Iron Concentrate. In addition, we further illustrated that the price of iron ore is the result of many factors.

(2) The proposed model can be easily applied in forecasting the price of iron ore with a high degree of precision. Its effectiveness was validated through a prediction test for 2018 and 2019. The predicted iron ore prices in the next five years (2020–2024) are, respectively, 750.74 CNY/ton, 785.46 CNY/ton, 808.12 CNY/ton, 821.47CNY/ton, and 826.55 CNY/ton with an average annual growth rate of 5.51%. A detailed procedure for constructing multiple regression modeling has been presented, ensuring its application in any price prediction system. The predicted price has deviated slightly from the actual price, potentially due to the impact of COVID-19.

Although the model proposed in this paper possesses impressive predictive performance, there remains scope for refinement. The price of iron ore is influenced by numerous factors, albeit the input variables incorporated in this study only provide a partial understanding. Nonetheless, the model showcases commendable predictive capability and boasts potential for further application in other domains.

Author Contributions: Conceptualization, Y.W. and Y.Z.; methodology, Y.W. and Y.Z.; validation, Y.W. and Y.Z.; formal analysis, Y.W. and X.H.; resources, Y.W. and Y.Z.; writing—original draft preparation, Y.W.; writing—review and editing, Z.G., X.H. and J.X.; supervision, Z.G.; funding acquisition, J.X. and Z.G. All authors have read and agreed to the published version of the manuscript.

Funding: This research is supported by the Basic Science Center Project National Natural Science Foundation of China (NSFC), 42074171, and the Hunan Science and Technology Major Project of the Ministry of Science and Technology of China (2020GK1021).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data were collected from the Internet.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Mishra, D.P.; Swain, S.K. Global trends in reserves, production and utilization of iron ore and its sustainability with special emphasis to India. *J. Mines Met. Fuels* **2020**, *68*, 11–18.
2. Astier, J. Evolution of iron ore prices. *Miner. Econ.* **2015**, *28*, 3–9. [CrossRef]
3. Upadhyay, R.K. Iron Ore Resources: Their Conservation, Value addition and Impact on Iron and Steel Industry. *Tata Search* **2008**, *1*, 197–205.
4. Wang, L.; Lai, M.; Zhang, B. The transmission effects of iron ore price shocks on China's economy and industries: A CGE approach. *Int. J. Trade Glob. Mark.* **2007**, *1*, 23–43. [CrossRef]
5. Su, C.W.; Wang, K.H.; Chang, H.L.; Peculea, A.D. Do iron ore price bubbles occur? *Resour. Policy* **2017**, *53*, 340–346. [CrossRef]
6. Wu, J.; Yang, J.; Ma, L.; Shen, X. A system analysis of the development strategy of iron ore in China. *Resour. Policy* **2016**, *48*, 32–40. [CrossRef]
7. Li, H.; Zhang, Z.; Li, L.; Zhang, Z.; Chen, J.; Yao, T. Types and general characteristics of the BIF-related iron deposits in China. *Ore Geol. Rev.* **2014**, *57*, 264–287. [CrossRef]
8. Sari, R.; Hammoudeh, S.; Soytas, U. Dynamics of oil price, precious metal prices, and exchange rate. *Energy Econ.* **2010**, *32*, 351–362. [CrossRef]
9. Pustov, A.; Malanichev, A.; Khobotilov, I. Long-term iron ore price modeling: Marginal costs vs. incentive price. *Resour. Policy* **2013**, *38*, 558–567. [CrossRef]
10. Morandi, M.I.W.M.; Rodrigues, L.H.; Lacerda, D.P. Foreseeing iron ore prices using system thinking and scenario planning. *Syst. Pract. Action Res.* **2014**, *27*, 287–306. [CrossRef]
11. Moghaddam, M.R.; Manjezi, M.; Danesh, A.H.M.; Kakha, G. Prediction of monthly price of iron ore by using artificial neural network. *Indian J. Sci. Res.* **2014**, *7*, 1200–1204.
12. Mobtaker, M.M.; Osanloo, M. Chaos in iron ore price prediction. In Proceedings of the Southern African Institute of Mining and Metallurgy MPES 2015—Smart Innovation in Mining, Johannesburg, South Africa, 9–11 November 2015.
13. Weng, F.; Hou, M.; Zhang, T.; Yang, Y.; Wang, Z.; Sun, H.; Luo, J. Application of regularized extreme learning machine based on BIC criterion and genetic algorithm in iron ore price forecasting. In Proceedings of the 3rd International Conference on Modelling, Simulation and Applied Mathematics, Shanghai, China, 22–23 July 2018.
14. Ramos, A.L.; Mazzinghy, D.B.; Barbosa, V.D.S.B.; Oliveira, M.M.; Silva, G.R.D. Evaluation of an iron ore price forecast using a geometric Brownian motion model. *REM-Int. Eng. J.* **2019**, *72*, 9–15. [CrossRef]
15. Wang, W.M.; Wang, W.K.; Du, D. Application of Grey Prediction Model GM (1.1) for Forecasting Hydrology Kas an example of Manasi river. *Ground Water* **2007**, *2007*, 10–12.
16. Zhao, M.; Zhao, D.; Jiang, Z.; Cui, D.; Li, J.; Shi, X. The gray prediction GM (1, 1) model in traffic forecast application. *Math. Model. Eng. Probl.* **2015**, *2*, 17–22. [CrossRef]
17. Lv, J.; Tang, W.; Hosseinzadeh, H. Developed multiple-layer perceptron neural network based on developed search and rescue optimizer to predict iron ore price volatility: A case study. *ISA Trans.* **2022**, *130*, 420–432. [CrossRef] [PubMed]
18. Guo, L. Analysis of Factors Affecting Iron Ore Price and Research on Price Prediction. Master's Thesis, Xi'an University of Architecture and Technology, Xi'an, China, 2008. (In Chinese)
19. Daut, M.A.M.; Hassan, M.Y.; Abdullah, H.; Rahman, H.A.; Abdullah, M.P.; Hussin, F. Building electrical energy consumption forecasting analysis using conventional and artificial intelligence methods: A review. *Renew. Sustain. Energy Rev.* **2017**, *70*, 1108–1118. [CrossRef]
20. Escribano, A.; Granger, C.W.J. Investigating the relationship between gold and silver prices. *J. Forecast.* **1998**, *17*, 81–107. [CrossRef]
21. Kearney, A.A.; Lombra, R.E. Gold and platinum: Toward solving the price puzzle. *Q. Rev. Econ. Financ.* **2009**, *49*, 884–892. [CrossRef]
22. Zhu, Y.H. Research on the Relationship between the Cost of New Shipbuilding Materials and Changes in International Iron Ore Prices. Master's Thesis, Shanghai Jiao Tong University, Shanghai, China, 2015. (In Chinese).
23. Gu, X.W.; Wang, Z.K.; Xu, X.C.; Huang, X.H.; Wang, Q. Linear regression prediction of iron ore price considering multiple factors. *Nonferrous Met. (Min. Part)* **2017**, *69*, 98–103. (In Chinese)
24. Wårell, L. An analysis of iron ore prices during the latest commodity boom. *Miner. Econ.* **2018**, *31*, 203–216. [CrossRef]
25. Omundsen, J. Forecasting the Price of Iron Ore. 2019. Available online: https://core.ac.uk/download/pdf/286812523.pdf (accessed on 19 October 2023).
26. Yan, X.D.; Yan, J.; Wang, Y.W.; Cui, J.H. The Impact of Tariff Policy Adjustment on the Development of China's Steel Industry. *Modern Industrial Economy and Informationization.* **2022**, *12*, 16–19+39. (In Chinese)
27. Li, D.; Moghaddam, M.R.; Monjezi, M.; Jahed Armaghani, D.; Mehrdanesh, A. Development of a group method of data handling technique to forecast iron ore price. *Appl. Sci.* **2020**, *10*, 2364. [CrossRef]
28. Liu, Y.; Li, H.; Guan, J.; Liu, X.; Guan, Q.; Sun, Q. Influence of different factors on prices of upstream, middle and downstream products in China's whole steel industry chain: Based on Adaptive Neural Fuzzy Inference System. *Resour. Policy* **2019**, *60*, 134–142. [CrossRef]
29. Kutner, M.H.; Nachtsheim, C.J.; Neter, J.; Li, W. *Applied Linear Statistical Models*; McGraw-Hill Irwin: Boston, MA, USA, 2005.

30. Huang, B.; Li, Y.; Liu, Y.; Hu, X.; Zhao, W.; Cherubini, F. A simplified multi-model statistical approach for predicting the effects of forest management on land surface temperature in Fennoscandia. *Agric. For. Meteorol.* **2023**, *332*, 109362. [CrossRef]
31. Wu, P.; Qiu, S.P. Macroeconomic development prediction based on optimized multi-dimensional grey model. *Stat. Decis.* **2020**, *36*, 42–45. (In Chinese)
32. Gan, Y.; Griffin, W.M. Analysis of life-cycle GHG emissions for iron ore mining and processing in China—Uncertainty and trends. *Resour. Policy* **2018**, *58*, 90–96. [CrossRef]
33. Akbary, P.; Ghiasi, M.; Pourkheranjani, M.R.R.; Alipour, H.; Ghadimi, N. Extracting appropriate nodal marginal prices for all types of committed reserve. *Comput. Econ.* **2019**, *53*, 1–26. [CrossRef]
34. Zhang, Q.; Wang, Y.; Zhang, W.; Xu, J. Energy and resource conservation and air pollution abatement in China's iron and steel industry. *Resour. Conserv. Recycl.* **2019**, *147*, 67–84. [CrossRef]
35. Yan, L. Research on China's Macroscopic Countermeasures against Iron Ore Price Fluctuation. Master's Thesis, China University of Geosciences, Beijing, China, 2009. (In Chinese)
36. Lima, G.A.C.; Suslick, S.B. Estimativa da volatilidade de projetos de bens minerais. *Rev. Esc. Minas* **2006**, *59*, 37–46. [CrossRef]
37. Yan, X.; Su, X. *Linear Regression Analysis: Theory and Computing*; World Scientific: Singapore, 2009.