ISSN: 2688-9323

Prediction of Crude Oil Price Change based on SPSS Multiple Regression Analysis

Chao Zhang[a], Jiawei Yang, Qibin Zhang

MBA Education and Management Center, Shanghai University, Shanghai, China

[a]lawnghyy@icloud.com

Abstract

Crude oil is the blood of modern industry, known as black gold. The change of crude oil price has experienced many ups and downs, which is difficult to figure out. With the continuous development of economy and the continuous increase of demand, the price of crude oil fluctuates greatly in recent years. What factors will affect the change of oil price and how will these changes affect the change of international oil price. In this paper, the trend of crude oil price will be predicted based on multiple regression analysis of SPSS software, and the linear regression equation will be obtained.

Keywords

Crude Oil; Price Forecast; Multiple Regression.

1. Analysis of Influencing Factors

1.1. World Crude Oil Production

Production determines the price to a certain extent. In the report, we selected the world oil production from 2008 to 2018 as the data parameter for analysis.

1.2. National Security of Major Producing Countries

For example, in the Iraq war in 2003, Iraq's crude oil production decreased significantly, but Saudi Arabia and Kuwait increased production immediately, which filled the decline in Iraq's exports and successfully prevented the sharp fluctuation of oil prices. The Syrian civil war and Libyan war in 2011 led to the reduction of domestic crude oil production and the rise of crude oil price to a certain extent. In this paper, we selected whether there was war in major oil producing countries from 2008 to 2018 as the data parameter for analysis.

1.3. Export Volume of Major Producing Countries

The so-called major producing countries refer to major crude oil exporting countries such as Russia, the United States and Saudi Arabia. Before 1990, the output of the Soviet Union was the largest, but after the disintegration, the output decreased greatly. Saudi Arabia was able to rapidly expand its production capacity due to the powan incident. In 1991, Saudi Arabia accounted for 35% of the total output of OPEC members and became the world's largest oil producer. However, OPEC countries still hold most of the oil sources, and the exploitable time is still about 100 years, so they are still a pivotal force. In this paper, we select the world oil export volume from 2008 to 2018 as the data parameter for analysis.

1.4. US Crude Oil Inventory

Every Wednesday, the American crude oil Association (API) and the U.S. Department of energy (EIA) will publish the weekly change table of energy supply and demand to reflect the increase and decrease of U.S. crude oil inventory balance, which plays an important role in the impact on crude oil prices. At present, the United States is the world's largest crude oil importer and consumer. In terms of demand, the importance of these two data can be imagined. In this paper,

ISSN: 2688-9323

we select the crude oil inventory of the United States from 2008 to 2018 as the data parameter for analysis.

1.5. US Dollar Index

The price of crude oil has always been closely linked to the US dollar, and its delivery and pricing are settled in US dollars, so the US dollar index will also have an impact on the price of crude oil. In the report, we selected the exchange rate of US dollar against RMB from 2008 to 2018 as the data parameter for analysis.

1.6. Global Economic Situation

The improvement of economic conditions led to the prevalence of industrial activities, and the demand for crude oil was also boosted, driving the oil price higher; On the contrary, when the economy is in a downturn, industrial activities are weak and the demand for crude oil is reduced, resulting in the weakening of oil prices. In the report, we selected the consumer price index of major countries and regions in the world from 2008 to 2018, the economic growth rate of major countries and regions over the years, the federal fund interest rate of the United States and other data parameters for analysis.

1.7. Seasonal Factors

Generally speaking, the price of unleaded gasoline tends to rise during the peak summer tourism season from May to September every year; The demand for heating oil tends to rise from October to April every year. The temperature in northern Europe and the east of the United States in winter has a great impact on the price of crude oil. In the report, we analyzed the dummy variables in 12 months of each year from 2008 to 2018.

2. SPSS Software and Multiple Linear Regression Model

2.1. SPSS Software

SPSS means statistical products and service solutions. It has powerful statistical analysis and data preparation functions, convenient chart display functions, broad compatibility and friendly interface. It is deeply loved by the majority of application statistical analysts.

2.2. Multiple Linear Regression Model

Multiple linear regression model means that in practical problems, a variable is often affected by multiple variables. The general form is:

$$Yi = \beta_0 + \beta i \times 1i + \cdots + \beta k \times ki + ui \quad i = 1,2 \ldots n \tag{1}$$

Where k is the number of explanatory variables, βj (j=1,2, …, k) is called the regression coefficient. The above formula is also called the random expression of the overall regression function, and its non-random expression is:

$$E\left(Y\middle|x_1\mathrm{i}, x_2 i, \ldots x_{k_i}\right) \quad i = 1,2 \ldots n \tag{2}$$

βj (j=1,2,… ,k) is also called partial regression coefficient. Multiple linear regression not only needs to test the regression coefficient, estimate the confidence interval of the regression coefficient, and discuss the prediction and hypothesis test, but also needs to consider the relationship between various variables, such as whether there is collinearity between them.

ISSN: 2688-9323

3. Linear Regression Analysis of Crude Oil Price

A	B	C	D	E	F	G	H	I	J	K	L
year	price (WTI)	production (unit: 100000kt)	economic 1	economic 2	war	D_war	EX rate	export value (unit: 100000kt)	EIA	IR	gold price
2008	101.82	36.4	94.3	3	no	0	6.97	21	0.000286	1.9275	867.093
2009	60.95	35.5	96.5	-0.1	no	0	6.83	20	0.000273	0.16	967.937
2010	79.61	36.3	100	5.4	no	0	6.78	20	0.000273	0.175	1236.732
2011	95.84	36.8	105.1	4.3	yes	1	6.47	21	0.000286	0.101667	1541.955
2012	94.54	37.3	109.6	3.5	yes	1	6.31	21	0.000286	0.14	1668.616
2013	97.26	37.5	114.3	3.5	yes	1	6.16	21	0.000286	0.1075	1410.907
2014	94.42	38.1	118.4	3.6	yes	1	6.16	20	0.000273	0.089167	1266.262
2015	49.99	39.1	122.6	3.5	yes	1	6.25	21	0.000286	0.1325	1153.431
2016	42.79	39.2	126.7	3.4	yes	1	6.62	22	0.0003	0.395	1248.264
2017	51.03	44.1	131.5	3.9	yes	1	6.77	22	0.0003	1.001667	1257.846
2018	65.55	44.5	138.2	3.6	yes	1	6.67	23	0.000314	1.5075	1269.453

Fig 1. Linear regression analysis chart

Find out what factors affect the price of crude oil. Through simple linear regression, this paper determines the influence degree of the following factors on crude oil price. The dependent variable is the price of crude oil. The independent variables include: (1) world crude oil production, (2) consumer price index of major countries and regions over the years, (3) economic growth rate of major countries and regions over the years, (4) whether there is war in the Middle East this year, (5) US dollar exchange rate, (6) world crude oil export volume, (7) US crude oil inventory this year, (8) US federal fund interest rate, (9) gold price, Regression equations will be developed using annual data.

Data Description:

1. Price - crude oil price (unit: USD)
2. Production - world crude oil output (unit: 100 million tons)
3. Economic 1 - consumer price index of major countries and regions in the world over the years
4. Economic 2 - economic growth rate of major countries and regions in the world over the years
5. War --- whether there is war in the Middle East this year
6. Ex rate - US dollar exchange rate
7. Export value - world crude oil export volume (unit: 100 million tons)
8. EIA - crude oil inventory of the United States this year (unit: 100 million tons)
9. IR - US federal funds rate
10. Gold price - gold price (unit: USD / oz)

As can be seen from figure 2, $R^2 = 0.923$, we use the decision coefficient R^2 To estimate the goodness of fit of the regression equation. R^2 It can be understood as the percentage of the total sum of squares that can be explained by the estimated regression equation. For the influencing factors of crude oil price, when using the regression equation to predict the crude oil price, we can conclude that 92.3% of the total square sum can be explained by the regression equation. In other words, 92.3% of the crude oil price can be explained by the world crude oil production,

ISSN: 2688-9323

the consumer price index of major countries and regions over the years, the economic growth rate of major countries and regions over the years, whether there is war in the Middle East this year, the US dollar exchange rate Explained by the linear relationship between world crude oil export volume, US crude oil inventory this year, US federal fund interest rate and gold price. Therefore, we are satisfied with the fitting effect of the estimated regression equation, and there is a high correlation between the selected independent variables and dependent variables.

Model summary

Model	R	R^2	After adjustment R^2	Errors in standard estimates
1	.961[a]	.923	.615	13.79562

a. Predictive variables:(Constant) gold price, production (unit: 100000kt) , economic 2, EX rate, IR, EIA, D_war, economic 1

Fig 2. Model summary

Coefficient[a]

Model		Non standardized coefficient B	Non standardized coefficient Standard error	Standardization coefficient Beta	t	Significance
1	(Constant)	710.069	225.003		3.156	.087
	production (unit: 100000kt)	−1.412	7.383	−.192	−.191	.866
	economic 1	−.492	1.971	−.323	−.250	.826
	economic 2	−.033	4.362	−.002	−.008	.995
	D_war	−6.318	35.971	−.133	−.176	.877
	EX rate	−48.002	47.797	−.623	−1.004	.421
	EIA	−1119169.3	1297738.21	−.648	−.862	.479
	IR	42.264	20.562	1.243	2.055	.176
	gold price	.076	.053	.775	1.432	.289

a. Dependent variable: price (WTI)

Fig 3. Model coefficient

Excluded variables[a]

Model		Enter beta	t	Significance	Partial correlation	Collinearity statistical tolerance
1	export value (unit: 100000kt)	.[b]	.	.	.	.000

a. Dependent variable: price (WTI)

b. Predictive variables in the model: (Constant) , gold price, production (unit: 100000kt) , economic 2, EX rate, IR, EIA, D_war, economic 1

Fig 4. Excluded variables

A It can be seen from the figure 3 and figure 4 that the variable "world crude oil export volume" is excluded in the linear regression, indicating that its impact on the dependent variable is negligible in the linear regression. Therefore, the price of crude oil is mainly affected by eight factors: world crude oil production, consumer price index of major countries and regions over the years, economic growth rate of major countries and regions over the years, whether there is war in the Middle East this year, US dollar exchange rate, US crude oil inventory this year, US federal fund interest rate and gold price. Therefore, we can get the regression equation:

ISSN: 2688-9323

Price=710.069-1.412*production-0.492*economic1-0.033*economic2-6.318*D_war-48.002*EX rate-1119169.3*EIA+42.264*IR+0.076*gold price

4. Conclusion

By analyzing the regression equation, we can see that the coefficient of the independent variable production is -1.412, the coefficient of economic1 is -0.492, and the coefficient of economic2 is -0.033, D_ The coefficient of war is -6.318, the coefficient of ex rate is -48.002, the coefficient of EIA is -1119169.3, the coefficient of IR is 42.264 and the coefficient of gold price is 0.076, among which, production, economic1, economic2 and D_ The coefficients of war, ex rate and EIA are negative, indicating that they are negatively correlated with the dependent variable crude oil price. The world's crude oil output, the consumer price index of major countries and regions in the world over the years, the economic growth rate of major countries and regions in the world over the years, the probability of no war in the Middle East this year, the higher the US dollar exchange rate, the higher the world's crude oil exports and the higher the US crude oil inventory this year, the lower the price of crude oil. The coefficients of IR and gold price are positive, indicating that they have a positive correlation with the dependent variable crude oil price.

The higher the federal funds rate and the price of gold, the higher the price of Plateau oil. In the aspect of significance analysis, the significance of production is 0.866, the significance of economic1 is 0.826, and the significance of economic2 is 0.995, D_ The significance of war is 0.877, the significance of ex rate is 0.421, the significance of EIA is 0.479, the significance of IR is 0.176, and the significance of gold price is 0.289. The closer the significance is to 0, the higher the significance is.

However, the significant values of the above data are too large, which may be caused by the following reasons: (1) the influencing factors considered are incomplete. The change of crude oil price is not a simple commodity problem, nor is the crude oil market a simple commodity market, which contains many complex factors such as finance and politics. Because crude oil has strategic material properties, its price will be affected by many factors, and only 9 factors are selected as independent variables in the above data. However, with the development of world multi polarization, economic globalization and production internationalization, competing for crude oil resources and controlling the crude oil market have become an important reason for the turbulence of the oil market and the soaring oil price. These factors have an important impact on oil prices, but it is difficult to simply measure with figures. (2) Calculation error. The above analysis only selects the multiple linear regression model as the analysis tool, but the crude oil pricing involves the international oil price pricing power, and its pricing complexity determines that its price is difficult to estimate with a simple linear regression model.

In the long run, the price trend of crude oil and gold is basically the same. There is a positive correlation between the price of gold and crude oil. The rise of crude oil price also indicates the rise of gold price. Through linear regression analysis, the factors affecting the change of oil price are finally determined: the impact of global economic conditions (including consumer price index of major countries and regions, economic growth rate of major countries and regions, federal fund interest rate of the United States), whether there is war in the Middle East this year, US dollar exchange rate, gold price and crude oil inventory of the United States this year.

ISSN: 2688-9323

References

[1] Klein T, Walther T. Oil price volatility forecast with mixture memory GARCH[J]. Energy Economics. 2016, 58: 46-58.

[2] Jammazi R, Aloui C. Crude oil price forecasting: experimental evidence from wavelet decomposition and neural network modeling[J]. Energy Economics. 2012, 34(3): 828-841.

[3] Zhang L J, Zhang Y J, Zhang L. A novel hybrid method for crude oil price forecasting[J]. Energy Economics. 2015 (49): 649-659.

[4] Li Naliao. Prediction of urban groundwater consumption based on SPSS multiple regression analysis [J]. Hebei water conservancy, 2021(08): 42-43.