

Testing the Quality of Forecasting Using the Theils U Statistics for Forecasting the Production Capacity of Crude Oil in Iraq Until 2033

Mariam Jumaah Mousa ^{a*}, Mustafa Kamil Rasheed ^b

^{a*} *Instructor, Department of Financial and Banking Sciences, College of Islamic Science, Imam Al-Kadhum University, Iraq. E-mail: maram.alamy@alkadhum-col.edu.iq*

^b *Professor, Doctor in Macroeconomic, College of Administration and Economics, Mustansiriyah University, Iraq. E-mail: dr_mustafa_kamel@uomustansiriyah.edu.iq*

Abstract

Since 2004, Iraq has been living on crude oil revenues in financing economic activities, which made it captive to the fluctuations of the global oil market, which has mortgaged the capacity of the macroeconomic Iraqi at the threshold of crude oil production capacity, at a time when Iraq is losing its local economic efforts in harnessing crude oil to its advantage, especially since All exploration, investigation, exploration and other works are carried out with foreign efforts. The research sheds light on the production capacity of crude oil in Iraq until the year 2033 in order to review what the production capacity of crude oil will lead to after the entry into force of the global climate agreement. Which constitutes a new challenge for the Iraqi economy. The research assumes that the decline in the production capacity of crude oil after 2030 could be a motive for abandoning the heavy dependence on crude oil in financing the Iraqi economy. Several advanced statistical techniques were used in order to reach the accuracy of the required statistical predictions according to the (Theils U) statistic, and predictions were made inside in order to verify the predictive power of the approved statistical models, the data used for the period (2005-2019), the research found (survival of Iraq is a large producer and exporter of crude oil until the year 2033), the research recommended (investment in the oil sector in order to develop local production and diversify the national economy).

Keywords: Crude Oil Production Capacity, Statistical Forecasts, Macroeconomics, Oil Market.

Introduction

The Iraqi economy has been characterized as a rentier economy, which means that it has witnessed, and still is, a significant weakness in the supply and links of its productive sectors in favor of the oil sector. Perhaps the easy financial surpluses prompted decision-makers to adopt them as a basis for financing the Iraqi state budgets, without regard for what would happen to the wheel of industrial, agricultural and service production, And the losses that the Iraqi economy will pay as a result of continuing these imperfect macroeconomic policies.

The production capacity of crude oil is also facing many challenges that led to the inability to reach the production capacities specified in OPEC. The most important of these challenges are the security and political instability, external interference, and the lack of vision among decision makers in building a healthy and sound economy, which prompted many oil installations to reserve many stages of oil production and wait for many expansion decisions. This was reflected negatively in the production capacity of crude oil and the export capacity, as the Iraqi economy did not seek to benefit from this resource in feeding the productive sectors and increasing the added value, and then increasing the economic surpluses that were supposed to be achieved in return for consuming this resource.

Research Problem: The Iraqi economy faces a global historical turn after 2030 when the global climate agreement is implemented, which puts the Iraqi economy at a crossroads, and it has to hedge from now on for that, in order to protect the economy from those events that will change its production pattern.

Research Importance: Forecasting the production capacity of crude oil in Iraq until 2033 puts a proposal for decision-makers in order to take the necessary measures to maintain the sustainability of the economy, support local industries, and promote economic diversification.

Research Hypothesis: The decline in the production capacity of crude oil after 2030 could be a motive to abandon the heavy dependence on crude oil to finance the Iraqi economy.

Research Objective

1. Forecasting from inside the time series of crude oil production capacity.
2. Forecasting the production of crude oil in the Iraqi economy until 2033.

Research Methodology: Using the deductive method in realizing the facts related to the research problem using advanced statistical methods to find out what will result in the future, relying on official data from the relevant institutions, as well as books, periodicals and reports.

Crude Oil Production Capacity and Statistical Forecasting, A Conceptual Introduction

Oil Production

The world witnessed rapid changes in most of the global macroeconomic variables, and after 2014 until 2021 unleashed the era of crises to undermine the unity of the contemporary economy and digital technology. The cartels took over the rest of the organizational unit, whether in the production, distribution or consumption of crude oil.

The recurrence of crises proved to the world that the OPEC cartel was not steadfast in front of it. As long as the primary goal of OPEC is to stabilize oil prices by adopting a set of strategies (Smith, 2008), a goal that OPEC was unable to maintain, especially during the Corona pandemic.

Geopolitical tensions, the recurrence of global crises, and the shutdown caused by the Corona pandemic have caused economic and social turmoil with serious repercussions in various countries of the world. This led to the emergence of new challenges for oil producers and consumers around the world, which made it difficult to predict oil demand or at least to stabilize the global oil price. (Majid Ibrahim, 2014).

Today, the production of crude oil has become a function of the strength of the link in the same cartel and the ability to form a new cartel that is stronger than the previous one, to allow the oil producer to better market his production, which is extremely difficult in light of the volatile and unstable international environment, which causes concern about turmoil. Which will affect oil production in the short or medium term.

An Introduction to Statistical Forecasting

There are many methods of forecasting in time series, and perhaps the most important of them refers to (Hammoud Manaf Youssef and others, 2019).

a) Linear General Trend T_t

The general trend component measures the changes that occur on the phenomenon at the long level, and this component is one of the most important components of the time series, which is relied upon in calculating estimates. It is increasing in periods and decreasing in other periods, which can be linear or non-linear to estimate this component. There are several methods that have been relied on the method of least squares (Safa Majeed Al-Kalabi, 2018).

Least Squares Method LSM: The general trend component model is similar to the simple linear regression model.

$$Y_t = A + Bt + e_t \dots (1)$$

$$\hat{Y}_t = a + bt \quad t = 1, 2, 3, \dots, n \dots (2)$$

Where: Y_t the dependent variable and t is the time independent variable.

By the LSM method:

$$b = \frac{\sum y_t t - n \bar{t} \bar{y}}{\sum t^2 - n \bar{t}^2} \dots (3)$$

$$a = \bar{y} - b \bar{t} \dots (4)$$

Where:

(a): the constant term of a general trend component equation.

(b): the general trend component.

b) He Averages Method

1. Simple Averages Method

The simple average method is used to predict the values of the time series that are affected by the general trend, as the predicted value represents the arithmetic mean of the last T of the data that precedes the prediction.

$$\hat{Y}_{t+1} = \frac{Y_t + Y_{t-1} + \dots + Y_{t-T+1}}{T} \dots (5)$$

Where:

$\hat{Y}_{t+1}$: predicted value for the upcoming period.

(Y_{t-1}): the real value of period t.

(T): the number of periods of moving media.

2. The Double Moving Averages Model

Moving averages are taken for the initial moving averages, and to extract predictions, the algorithm is used, which assumes, as a special case, that the two averages have the same periods used to calculate the average, as follows:

$$S'_t = \frac{Y_t + \dots + Y_{t-N+1}}{N}$$

$$S''_t = \frac{S'_t + \dots + S'_{t-N+1}}{N}$$

$$\hat{Y}_{t+1} = a_t + b_t * L \quad -L \text{ is the predicted number of steps.}$$

$$a_t = 2S'_t - S''_t, b_t = \frac{2}{N-1} (S'_t - S''_t)$$

The method can be generalized to MA(M*N), which means calculating a moving average to degree N, then calculating the vector S'_t over N, then calculating a moving mean of degree M/ for the previous media, and dividing S''_t by M.

3. Simple Exponential Smoothing Method

Its standard form is as follows (Anwar Rashid Al-Salmani et al., 2018):

$$\hat{Y}_{t+1} = \alpha Y_t + (1 - \alpha)Y_{t-1} \quad 0 < \alpha < 1 \dots (6)$$

Where:

$\hat{Y}_{t+1}$: Estimated observing value in the subsequent period.

α : a constant that is automatically estimated based on the prediction error calculation and has a value ranging from zero to one.

Y_t : observing value for the current period.

Y_{t-1} : observing value in the previous period.

The Quality of Prediction Test Using Theils U Statistic

Theil proposed two U statistics used in finance the first is a measure of forecast accuracy the second is a measure of forecast quality by calculating the theil u-statistic, we compare VAR model forecasts against the forecasts generated assuming a random walk model (efficient markets are assumed to have a random walk type of behavior) the statistic in excess of one means our model did not forecast we compared to the random walk model (the naïve forecast) theils U statistic is calculated as follows to arrive at the theils U-statistic we compare several error statistics followed by calculating the theils U-statistic.

1. Calculate the sum of squared forecast Errors (SSE).

$$SSE_t = \sum_{i=1}^{N_t} e_{it}^2 \dots (7)$$

e_{it}^2 : the actual value minus the predicted Value for the dependent variable.

2. Root mean Error (RME) is calculated:

$$RME_t = \sqrt{\frac{SSE_t}{N}} \dots (8)$$

3. Sum of squared Error No-change Forecasts if we assume the difference between the dependent variable value assuming the random walk model and the actual observed value.

4. Root mean square of the SSE_{NCFt} :

$$RMS_{NCFt} = \sqrt{\frac{SSE_{NCFt}}{N_t}} \dots (9)$$

Where: (N_t) the total number of sample observations theil U For equation (8,9) the u-statistic is (REF: Dharmasena & A. Bessler, 2004).

$$U_t = \frac{RMS_t}{RMS_{NCFt}} \dots (10)$$

If $U=1$ the naïve method is as good as the forecasting technique being evaluated.

If $U<1$ the forecasting technique being used is better than the naïve method.

If $U>1$ there is no point to using the formal forecasting method. naïve produces better result.

Results and Discussion

The data of the development of Iraqi oil production for the period 2005-2019 was relied on in order to analyze it with the general linear trend model, the exponential smoothing model, the moving averages of the series data from their relationship to time as an independent variable that expresses Trend of the development of oil production.

Table 1: Numbering the statistical characteristics of the development of oil production for the period 2005-2019

Mean	3074
Standard Error	261.4914
Median	2942
Mode	#N/A
Standard Deviation	1012.752
Sample Variance	1025666
Kurtosis	-1.24485
Skewness	0.538372
Range	2795
Minimum	1853
Maximum	4648
Sum	46110
Count	15

From the above statistical characteristics, it was found that the production rate during the period was 3074, and production was not repeated for more than a year, and this is clear from the value of the mode, that the rate of production difference from its average during the period is 1025666 and with a deviation equal to 261.4914. As for the value of the skewness, the frequency distribution is skewed towards the right, i.e. concentration the values of the variable are towards small values more than their concentration towards large values compared to their arithmetic mean. As for the flatness value -1.24485, it indicates that the distribution is flat compared to the normal distribution and there is no significant effect of flattening on the results of statistical analyzes because it is between [-3,3] and that the highest production during The fifteen years was equal to 4648 in 2019 and the lowest production equal to 1853 in 2005 and the difference between the highest production and the lowest production is 2795.

A. General Trend

It shows the development of oil production as a dependent variable Y_i and time t_i as an independent variable affecting oil production. The model was estimated based on the least squares method, and the results were as follows:

$$\hat{y}_i = 1338.229 + 216.9714 t$$

$$R^2 = 0.91797 \text{ adj } R^2 = 0.91166$$

Table 2: Analysis of Variance of the General Trend Equation

ANOVA					
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Significance F</i>
Regression	1	13181448	13181448	145.4805	1.97E-08
Residual	13	1177882	90606.29		
Total	14	14359330			

The results indicate that there is an effect between the dependent variable, the development of oil production and the time-independent variable, with a direct relationship, and this means that over time the development of oil production rises, and the values of the coefficient of determination show that the explanatory power of the model is high, as the value of the coefficient of determination reached 92%, indicating that the independent variable has Explain the changes that occur in the development of oil production and that the estimated model as a whole is significant because $F = 145.48$ which is greater than the tabular, which confirms the significance of the model, that the value of b is positive and its value is 216.9714 means that the increase by one unit in time (one year) leads to an increase in the total development of oil production by b .

B. Methods of Averages

Simple Average Form

The simple mean method was applied for N from different periods and the best results were when the value of $N = 2$.

Double Moving Averages Model

$MA(N*M)$ was applied for M, N from different periods and the best results were ($M=2, N=2$).

C. Single Exponential Smoothing Model

It was model applied with different coefficients. The best results were when the value of $\alpha = 0.995$, which indicates the impact of its importance in the prediction model because its value is high and close to the correct one.

Table 3: Accuracy Measurements for MA (2)

Naïve	Exp smothing	MA (2*2)	MA (2)	
128360.1	128683.4	50434.1	32090.02	MSE
	1.0033>1	0.59796<1	0.5<1	Thuil's u

Decision: When comparing prediction models, we find that the moving averages model is the best model for forecasting, then the prediction model using double averages. MSE, however, it is better to use the U-statistic because it aims to measure the quality and accuracy of the appropriate model for forecasting. And the prediction using the exponential smoothing method was the result of the test statistic $U=1.033$ almost one, that is, the forecasting efficiency using the naïve method is similar to the prediction using the exponential smoothing method.

The following table shows the real value and the predicted value within the time series.

Table 4: Real and Predicted Value Within the Time Series

Years	ProdYT	Ma (2)
2005	1853	#N/A
2006	1957	1905
2007	2035	1996
2008	2281	2158
2009	2336	2308.5
2010	2358	2347
2011	2652	2505
2012	2942	2797
2013	2979	2960.5
2014	3110	3044.5
2015	3504	3307
2016	4648	4076
2017	4469	4558.5
2018	4410	4439.5
2019	4576	4493

Therefore, the MA (2) model was then relied upon in forecasting until 2033.

Table 5: Forecasting Off Time Series to 2033

Years	MA (2)
2019	4493
2020	4466.25
2021	4479.625
2022	4472.938
2023	4476.281
2024	4474.609
2025	4475.445
2026	4475.027
2027	4475.236
2028	4475.132
2029	4475.184
2030	4475.158
2031	4475.171
2032	4475.164
2033	4475.168

We note from the above table a decrease in oil production for the period (2020-2033) from what it was in 2019 and fluctuation in production and there is no jump in production in the coming years, Whatever, that Iraq will not give up its rent during the coming period, which naturally requires preparation of Macroeconomic policies to enhance the structure of local production, the economy and its productive sectors, depending on the financial surpluses generated by the oil sector during the next ten years.

Conclusions

1. The evolution of the geopolitical scene around the world and the repercussions of the Corona pandemic and the Russian-Ukrainian war have caused many countries to reconsider the use of crude oil, not even coal, as primary sources of energy in the short, medium and even long term, and this applies to the prediction results that indicate that Iraq's oil production will remain Crude oil is almost stable until 2033, which undermines the implementation of the 2030 climate agreement.
2. Through the analysis of the time series data, there is an increasing general trend for the time series under study, and this is evident by estimating the general trend component using the least squares method that the value of b is positive and equal to 216.9714.

3. It is preferable to use the U-statistic because it compares two prediction models together, which is the naïve model with the rest of the prediction models.
4. Through the value of the U statistic, the best model was found, the moving averages model at $N = 2$.
5. Lower oil production for the period (2020-2033) than it was in 2019.

Recommendations

1. Develop scenarios for the future of the Iraqi economy dependent on crude oil, which would put the possibilities of losses and gains that will be achieved at the level of the local economy.
2. The necessity of concerted efforts of decision makers through mobilizing macroeconomic policies towards developing the oil industry and benefiting from the accompanying investment projects that would support the local industry and enhance the interdependence between the sectors of the national economy.
3. Developing the oil derivatives industry to be a tributary to local industries and to compensate for the import of fuels and the waste of foreign exchange.
4. Investing gas associated with crude oil in order to secure sustainable sources of fuel for the production of electric power, as well as the economic gains achieved from selling natural gas globally.
5. Relying on the U-statistic in testing the quality of prediction to come up with good predictions that decision-makers will use in order to build correct strategies.

References

- Abdel, S.B., & Suleiman, B. (2019). What's driving the price of crude oil? Analysis of determinants and causes of shocks. *Dirasat Journal*, 16(1).
- Dharmasena, S., & Bessler, D. (2004). Weak-form efficiency vs semi-strong form efficiency in price discovery: an application to international black tea markets. *Sri Lankan Journal of Agricultural Economics*, 6(1381-2016-115717), 1-24.
- Hammoud, M.Y. (2019). Time Series Analysis, Part One, First Edition, Publishing House, Al-Daad Library for Printing and Publishing, Baghdad, Iraq.
- Hammoud, M.Y. (2020). Time Series Analysis, Part One, Second Edition, Publishing House, Al-Daad Library for Printing and Publishing, Baghdad, Iraq.
- Majed, I.A. (2014). Reality and future prospects of demand for crude oil and natural gas in industrialized countries and the implications for member states. *Journal of Oil and Arab Cooperation*, 40(149).
- Muhammad, A.H. (2009). The prospective role of the oil sector in rebuilding the economy of Iraq. *Al-Ghari Journal of Economic and Administrative Sciences*, 2(11).
- Safaa, M.A. (2018). *Using different prediction methods to analyze the number of patients with malignant tumors*. Master's thesis, Statistics Sciences, College of Administration and Economics, University of Karbala.
- Smith, J.L. (2008). World oil: market or mayhem?. CEEPR.