Int. J. Nonlinear Anal. Appl. In Press, 1–16
ISSN: 2008-6822 (electronic)
http://dx.doi.org/xxx

Designing a model for financial forecasting using the integration of neural networks, Box Jenkins and Holt Winters methodology

Kobra Hosseini, Mohammad Ali Keramati*, Mohammad Ali Afshar Kazemi, Zadallah Fathi

Department of Management, Central Tehran Branch, Islamic Azad University, Tehran, Iran

(Communicated by name of the Editor)

Abstract

Each of the methods used in the forecasting process has advantages and disadvantages based on the nature of the input data. The use of a new approach in the field of forecasting, which has recently received attention. The use of a hybrid approach is to increase the accuracy of forecasting methods. The combination of several forecasting methods, if done professionally, always improves the accuracy of the forecasting methods. The current research is a retrospective comparison that uses average OPEC oil price data from 2003 to 2022 to forecast for 6/2022 to 5/2024. The methods used in this research in single variable forecasting are time series method. It was based on the methodology of Box Jenkins and Holt Winters, who in the second step entered the hybrid model based on artificial neural network. Based on the results of error analysis, Box Jenkins methodology between ARIMA time series processes, ARIMA5.1.5 ARIMA 3.1.5, ARIMA 4.1.5 and ARIMA 5.1.3 You have the best accuracy with MSE ,61.86 , 63.29 ,63.21,63.62 respectively They were. Due to the nature of the data, the accuracy of the Holt- Winters method was not suitable compared to the time series method. To combine prediction methods, the best artificial neural network was designed, this neural network included a 5-neuron input layer, a 5-neuron middle layer and a single-neuron output layer, based on Levenberg – Marquardt Algorithm and the linear sigmoid function was used, the output results showed that the designed hybrid neural network was able to significantly improve the accuracy of the prediction methods and improve theMSE ,MAPE ,AIC andBIC indices .

Keywords: prediction, time series, ARIMA, neural network
2020 MSC: 68T07, 37M10

1 Introduction

Researchers have published sophisticated prediction techniques that are related to various fields of mathematics, statistics and artificial intelligence. Each of the methods used in the forecasting process has advantages and disadvantages based on the nature of the input data. Using a new approach in the field of forecasting that has recently received attention. The use of a hybrid approach is to increase the accuracy of forecasting methods. The combination of several forecasting methods, if done professionally, always improves the accuracy of the forecasting methods. The

*Corresponding author
Email addresses: first.author@email.address (Kobra Hosseini), second.author@email.address (Mohammad Ali Keramati), third.author@email.address (Mohammad Ali Afshar Kazemi), forth.author@email.address (Zadallah Fathi)

Received: xxx yyyy *Accepted:* xxx yyyy

Arima time series method based on Box Jenkins methodology has been used in the forecasts of Moore's economic and non-economic variables and the evaluation by accuracy indicators shows that among the methods based on financial statistics, Box Jenkins methodology has high accuracy. In relation to forecasting methods based on financial mathematics, the method based on Halt Winters has a good accuracy in forecasts. Regarding the methods based on machine learning, recently different neural networks have been used by researchers, the most important of which is ANN. univariate and multivariate ANN.
In the global economy, crude oil is considered one of the most important strategic goods that plays a significant role in determining many regional and international equations. Therefore, economic researchers and political decision makers are always trying to know the correct forecast of crude oil prices. Oil markets are considered one of the most complex, turbulent and non-transparent international financial markets, the price of crude oil is one of the most important key variables that influence the performance strategy of financial markets between International has a significant impact. Hence, pre Oil price forecasting not only plays an effective role in government policy, but also on the optimal It is also very effective to increase production in the long term.
According to the topics mentioned in the field of forecasting, the main question of this research is:
To what extent can the artificial neural network improve the accuracy of the predicted values by Box Jenkins and Holt Winters methodology by integrating the OPEC crude oil price forecasting process?

2 Forecasting and future research

Futuristic research is a science, because the futurist must act according to the rules in all stages, from problem understanding and data analysis to research methods and critical reasoning. On the other hand, future research is considered a kind of science because the findings of future researchers have the ability to transform into various forms of social and economic values. (Song et al., 2020). Since the Second World War, scientists, sociologists and activists in the field of operational research and many other people of science who called themselves futurists, started to establish and expand quantitative and qualitative methods in order to rationally predict the future. Therefore, forecasting methods have always been an important tool in the hands of future researchers. Quantitative forecasting methods are divided into two categories of using mathematical and intelligent knowledge, each of which has advantages and disadvantages.

one of the key elements of management and future research, is a tool for future Forecasting, as research of parameters and necessary variables in a system scope, the process of forecasting provides information for the decision and scenario making process. Forecasting is one of the important tools for making strategic decisions among managers and decision makers (Neto et al., Forecasting methods are used in various economic and social fields, and the main use of . (2021 Energy is one of the main .these methods is in the analysis and future research of energy systems axes of the development and progress of human societies and for this reason it has strategic importance in relations between countries. (Lee et al., 2021). One of the basic goals in estimating a regression model is to be able to predict changes in the endogenous variable with a certain value of the exogenous variable. Forecasting is the process of using an objective or mental model to estimate a variable for the past or future. In order to predict a variable, one should first predict the variable within the sample and choose the best method. It .then predicted the variable based on the best model for the future (Conde Vela et al., 2020)

3 Basics of artificial neural networks

networks are computational models that, by processing experimental data, transfer the knowledge or law hidden behind the data to the network structure. For this reason, they are called intelligent systems; Because based on calculations on numerical data or examples, they learn general rules.
networks are a simulation of the human nervous system. So, the main unit of the network - are artificial neurons that are simulated based on natural neurons. Based on their structure and type, neural networks can consist of one or more layers of artificial neurons. In fact, the neuron or node is the smallest information processing unit in the network with independent input and output, which is the basis of neural networks. Each of these neurons, by receiving input information and processing them, produces an output message.

Neuron

4 Advantages and disadvantages of artificial neural networks

networks are among the methods capable of estimating multiple nonlinearities in the data. These methods are considered a flexible calculation framework for a wide range of nonlinear problems. One of the obvious advantages of such models compared to other non-linear models is that artificial neural networks are a universal approximation that can approximate any function with desired accuracy. The strength of neural networks is processing the information in the data. Such networks do not require any assumptions about the shape of the pattern in the pattern making process and are generally a data-based pattern.
Despite all the unique advantages that are considered for artificial neural networks, these - types of models also have disadvantages, including the limitation of the number of required data or the limitation of linear modeling of this model. species of patterns mentioned (Franco et al., 2020).

5 Neural network prediction methods

It has been more than twenty years that the artificial neural network has been known as a general numerical tool It has different uses. McClock and Pitts presented the generality of theANN model in 1943 to solve a numerical problem in the way that the human brain solves it. Their article showed remarkable results that with a suitable combination of simple processing units, the ANN model is able to calculate any computable function (Tan et al., 2020). With this claim, the methods of using neural network started and progressed.
In order to know howANN is known, first we have an introduction to the human brain. The biological brain is a powerful processor that has the ability to perform physical tasks such as pattern recognition in a very short time . The study of neuroscience for the human brain shows that this powerful function starts from a series of basic components called neurons, each of which is connected to another neuron with a complex connection . (Sivansen et al., 2021). Figure 2-1 It cell body, axon and ,dendrite shows the basic implementation of a neuron which includes synapse. Information from an external source enters the neuron through dendrites . This information is weighted by another component called synapse . The synapse has a function to convert the electrical potential into a chemical signal to stimulate or silence any input . Then another organ called the cell body combines the inputs and fires (activates) an output , provided that the level of the combined input is greater than or equal to a certain threshold value. The axon, which acts as a transmission path , ultimately conveys the processed output to another unit sends (Delisa et al., 2021).

6 Artificial neural network structure for prediction

ANN is a combination of basic computational processing elements connected with each other in the form of layers. These processing elements, called neurons , model a multi-input nonlinear processor . Connections between neurons have There are different coefficients . Each neuron receives all inputs and combines them with different weights. The combined inputs are passed .through an activation function to produce an output, which can be used as an input to other units

7 Types of neural networks in terms of methods of learning

networks are divided into three categories based on the learning method (Fathi, 2011).
In this method, the weights are calculated once and updated. : Training with fixed weights .1 These networks are used in optimizing and compressing information as well as retrieving patterns.
the weights are 4 In these networks , there is no target output so that : Unsupervised training .2 corrected based on the error value obtained by comparing the output of the network with it.
Therefore, the network receives its training patterns through its inputs and places them in different categories as desired. When the network receives a series of inputs, by discovering a correlation and statistical relationship between different inputs, it corrects the weights and an .answer appears in the output that indicates the classes to which that input belongs In such networks, for each category of input training patterns, the : Supervised training .3 corresponding output is also given to the network and the change of weights continues until the difference between the value of the network output and the target output is minimized.

Fixed Weighted
Unsupervised Learning
Target Output
Supervised Learning

8 Types of neural networks in terms of structure and information processing

Artificial neural networks are divided into different types based on the direction of information entering the network, the structure and how the network is processed . Each of these models are used in different modeling problems according to the background and conditions of the desired problem as well as their capabilities and limitations . This type of neural networks can be summarized as follows:

1. Feedforward multilayer neural MLP networks
2. Hopfield neural Hopfild networks
3. self-organizing neural SOM networks
4. radial axis neural RBF networks
5. Probabilistic neural PNN networks
6. General Regression Neural GRNN Networks

9 Forecasting time series based on Box Jenkins methodology

The purpose of time series analysis is to describe, describe and predict the future values of a process. Describe the process including data graphing , identifying its stability and instability and checking the autocorrelation of the series. before Forecasting involves estimating future values of Time series are divided into two types of .the series based on observed data (Istan et al., 2021) stationary and non-stationary series. A series is stationary when there is no regular change in its mean and variance and strong periodic changes are removed. Unstable series can be transformed into stationary series by differentiating or stabilizing its variance (Hall et al., 2020). Changes in time series can be caused by natural factors or other factors, and therefore, its components should be known and measured . Time series Models are in fact random A model in which a time series containsN observations from an infinite population is a random process. The most important goal of time series analysis is to find the model of changes and predict the future. The purpose of time series forecasting is to estimate the values of the data set whose value is unknown at the time of the analysis. One of the techniques for predicting the behavior of time series is Box Jenkins method. The basis of Box Jenkins' approach in time series forecasting is based on a wide range of forecasting models for a time series. The general group of models for a time series in the Box Jenkins methodology are integrated autoregressive and moving average models, which are known asARIMA models in statistics The time series model is written as .ARIMA(p,d,q) In this form, p determines the autoregressive order orAR and expresses the dependence of an element in the present time on the previous effective values. In this model, the regression of each element is determined according to its previous values. Autoregressive processes are useful in expressing situations where the current value of the time series depends on its previous values plus a random coefficient. q also determines the order of the moving average, by which the dependence of the series on the random elements of the present and the past is defined. The autoregressive pattern combined with the moving average of orderp ,d ,q is shown below. Construction and forecasting of time series models includes four stages of pattern identification, parameter estimation, pattern correctness diagnosis and forecasting.

10 Box Jenkin's methodology based on time series

econometrics Econometric modeling
one of the new and important branches of economics, the familiarity and use of Econometrics is which can be a great help to better understand the relationships between economic and social phenomena. Econometric models describe an economic system with the help of a set of simultaneous equations that express mutual relationships between measurable variables. Equations are based on theory or empirical observations of behavior or technical relationships between variables, and if the analyzed model includes one or more elements with a break from ,the dependent variable as an explanatory variable, then that model is called dynamic. (Gujarati 2017)

Multilayer Perceptron
Hopfield
Self-Organizing Map
Radial Basis Function
Probabilistic Neural Network
General Regression Neural Network
Econometrics

Characteristics of a good forecasting model

C. AC Harvey provides the following criteria to judge the quality of a model (Gujrati and Damodar, 2016)

1) Logical lack of explanatory variables

A model is never able to accurately describe the reality as it is, that is, to accurately describe the reality, we will have to present such a complex model that lacks the slightest scientific value. Simplification and abstraction are inevitable in any modeling program. In this regard, the principle of logical scarcity of explanatory variables dictates that a model should be considered as simple as possible, so that only key variables should be included in the analysis, thereby erased.removing all random and minor effects, except for disturbance or error

Recognizability

This principle dictates that for a given data set, the estimated parameters must yield unique values. In other words, only one estimate is obtained for each specific parameter.

Good value

Considering that the main purpose of regression analysis is to explain the changes in the dependent variable by means of explanatory variables considered in the model, therefore a model measured by it is even as high as R^2 is considered as a good model when the explanation should not be considered alone, but this criterion can be used R^2 possible. Of course, the criterion in parallel with other criteria.

Compatibility with theory

high value, due to the wrong signs of one R^2 A model may not be a good model despite having a or more of its coefficients. Then the results should be interpreted with skepticism.

The power of generalization and prevention

According to Friedman, "(... the only appropriate test for the validity of a hypothesis (model) is The above implies the generalization power of R^2 .((to compare its prediction with experiences only in relation to one sample R^2 the model, and it should be said that this power expressed by is and for that sample, and what is intended here is the power of prediction and generalization for a period outside the sample period.

11 The nature of regression analysis

Regression is the main tool of econometrics. In general, regression analysis deals with the study on one or more other variables (of the dependence of one variable (dependent variable by estimating or predicting the average value or average values of the (explanatory variable first type variable in the case that the values of the second type variable are known or have been determined (in repeated sampling), for example, an agricultural economic expert uses regression analysis to study the dependence of the yield of a product, such as wheat, on temperature, rainfall, light, and productivity. Therefore, a dependency analysis can make it possible to predict or predict the average yield of the product according to the assumed information about the explanatory variables . (Niromand et al., 1400).

11.1 Nature of data for regression analysis

The accuracy of any econometric analysis will ultimately depend on the availability of the correct data. Generally, three types of data are accessible for empirical analysis:

1. Time series data
2. Cross-sectional data

Parsimony
Disturbance or error term
Disturbance or error term
Identifiability
Goodness of fit
Dependent-Variables
Explanatory - Variables
time series data
Cross-sectional data

3. or tabular data Composite

Time series data: values of one or more variables collected over a period of time , such as data on gross national product) GNP (employment, unemployment, money supply, etc., denoted by , the symbolt Such data can be collected at regular intervals such as daily, weekly, monthly, . quarterly and annually. They can also be quantitative (price, income and money supply) or qualitative (male and female, employed and unemployed, married and single, white and black).

Cross-sectional data: the values of one or more variables for several units (household, firm, state) or the values of one or more variables for a sample case (inflation rate data collection) at a specific (same) time, which are denoted by the symboli Such as the five-year population census . by the statistics center, etc.

Composite or panel data: There are elements of both cross-sectional data and time series, and in fact, this data is a combination of time series and cross-sectional data that has spatial (spatial) and time dimensions denoted by the symbolt i .show In these relationships, n is ,the number of predictions e_i the prediction error, which is obtained from the difference between the predicted values and the actual values, and y_i is the actual value. These three criteria will be used to measure the power of prediction in this research.

12 Prediction or time series based on Box Jenkins methodology

12.1 Data mean test

The purpose of Box Jenkins is to identify and determine the statistical model, which can be interpreted as the model that generates real sample data from a random process. If we want to use this model for prediction, the features of this model should be constant over time . If the observed series is unstable compared to the average, then the series can be differentiated so that the desired tip The main thing in time series models is that if the series . series becomes a stationary series obtained from the model will not used in the model are unstable or have a time trend, the results be valid. Therefore, before evaluating the presentation models done, the variables should be tested first in terms of significance and then according to the results obtained, selected their suitable face in the model. To determine the significance of the variables, the Dickey-Fowler test is used . In this test, if the absolute value of theT statistic is greater than the critical values, the assumption of the significance of the time series is confirmed. As can be seen in Table1 the oil , price time series data became stationary after one differentiation.

Table 1: Dickey-Fuller test of the data used in the research

data	t statistic Before differentiation	Dickey-Fuller's test statistic before differentiation critical value at the (level 5%)	Differentiation order	t statistic After differentiation	Dickey-Fuller's test statistic after differentiation (critical value at the 5% level)
Oil prices	1/021	1/5146	1	1678/5	3/5107

Figure1 .shows the differenced data, which is the characteristic of Dadard

Pooled data
Panel data

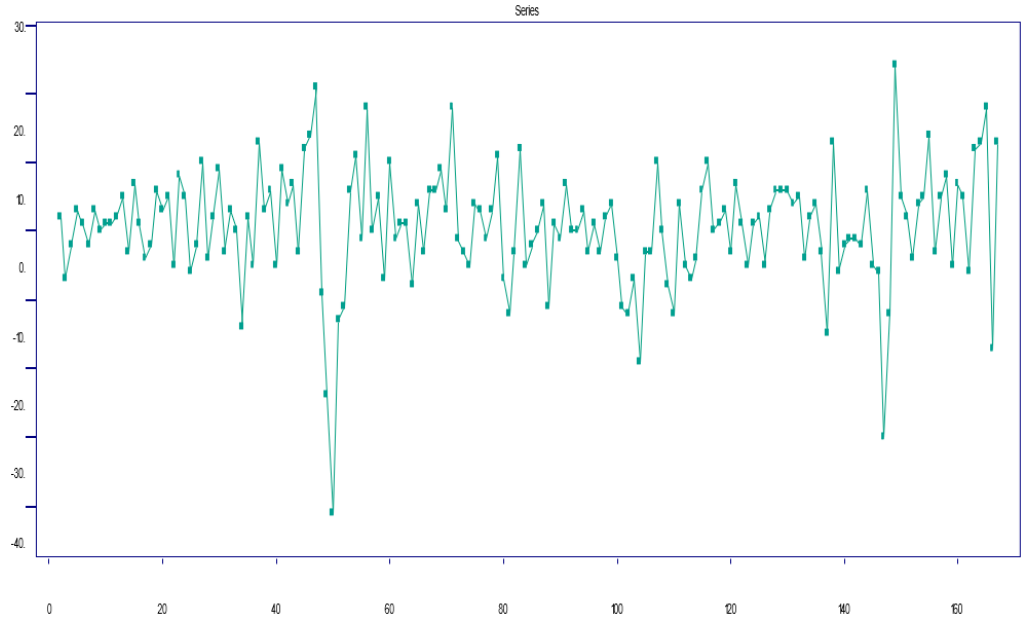

Figure 1: First order difference of oil price data

13 The first stage of diagnosis based on Box Jenkins methodology

After stationing the data using the form ofACF functions autocorrelation and partial autocorrelation PACFAppropriate model for identification data may be madeACF PACF Oil prices are shown in Figure 2 .

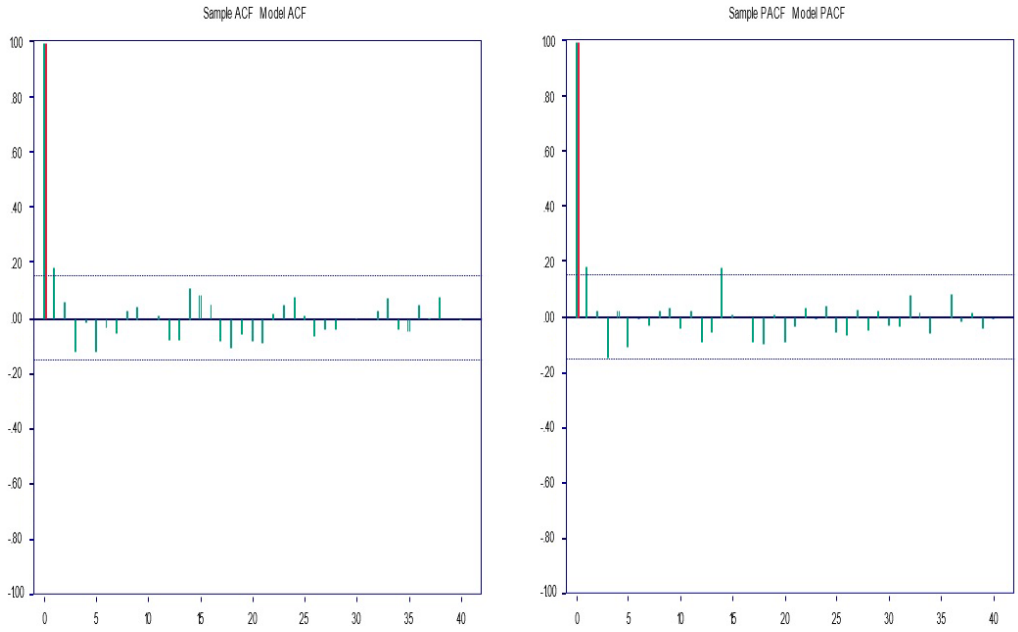

Figure 2: ACF functions and PACFOil prices

ARIMA method or Box Jenkins model, a model is prepared for a time series that considers the behavior of a variable or groups of variables based on their past behavior, these models are basically models based on a specific and relatively small number of variables and in Data trend forecasting does not accept emphasis on theoretical foundations as a priority. To use the Box- Jenkins model to forecast oil prices for the coming months, the appropriate model for each series should be selected based on the lowestMSE value based on significance . According to the generalized Dickey-Fuller's stationarity test that was performed in the previous part, the time series of the amount of Oil prices

at the level and width from the origin was equalized after one differentiation , thereforeARIMA is . used to predict the series with the Box-Jenkins method

Table 2: ARIMA(p.d.q) models for oil price

Model type	AR process rating	MA process rating	order of difference	MSE index
ARIMA (1.1.1)	1	1	1	69.7565
ARIMA (1.1.2)	1	2	1	67.6026
ARIMA (1.1.3)	1	3	1	68.1315
ARIMA (1.1.4)	1	4	1	66.9464
ARIMA (1.1.5)	1	5	1	66.4456
ARIMA (2.1.1)	2	1	1	68.3238
ARIMA (2.1.2)	2	2	1	66.4269
ARIMA (2.1.3)	2	3	1	65.9390
ARIMA (2.1.4)	2	4	1	65.9677
ARIMA (2.1.5)	2	5	1	65.9891
ARIMA (3.1.1)	3	1	1	68.1033
ARIMA (3.1.2)	3	2	1	67.1036
ARIMA (3.1.3)	3	3	1	65.8350
ARIMA (3.1.4)	3	4	1	68.0145
ARIMA (3.1.5)*	3	5	1	63.2919
ARIMA (4.1.1)	4	1	1	68.0992
ARIMA (4.1.2)	4	2	1	65.9736
ARIMA (4.1.3)	4	3	1	66.0377
ARIMA (4.1.4)	4	4	1	63.5430
ARIMA (4.1.5)*	4	5	1	63.2150
ARIMA (5.1.1)	5	1	1	66.8641
ARIMA (5.1.2)	5	2	1	66.8177
ARIMA (5.1.3)*	5	3	1	63.6299
ARIMA (5.1.4)	5	4	1	63.7311
ARIMA (5.1.5)*	5	5	1	61.8630

Based on the obtained results and comparison of mean squared errorMSE , ARIMA model It has the lowest error square value (5.15)(61.8630) among different ARIMA models then ARIMA (4.1.5) ,ARIMA (3.1.5) models andARIMA (5.1.3)had the least error, the analysis of each of them is given according to Box Jenkins methodology.

14 model stimation(5.15)

To estimate the coefficients of the model, theMLE maximum likelihood method is used, but when the model is non-linear with respect to the parameters, we resort to non-linear methods. **ARIMA .It is in the form of the following equation** process equation (5.1.5)

$$\begin{aligned} Z(t) &= 0.4213Z(t-1) + 0.1355Z(t-2) - 0.4843Z(t-3) - 0.4666Z(t-4) + 0.3264Z(t-5) \\ &+a(t) - 0.2527a(t-1) - 0.1470a(t-2) + 0.3925a(t-3) + 0.7050a(t-4) - 0.4252a(t-5) \end{aligned} \tag{14.1}$$

14.1 diagnostic control(ARIMA (5.1.5))

After selecting a specificARIMA modeland estimating its parameters, we are looking for whether the selection model fits the data well ? In other words, is the selection model the most appropriate model to describe the data. Because it may be anARIMA modelprovide a better fit of the data . A simple test to check this point is that the

Diagnostic checking

residuals from this disturbance model should be white, i.e. have a normal distribution, constant variance and zero mean. The results showed that the residuals from the processARIMA (5.1.5) White disturbance .are the remaining diagram of the ,ARIMA process (5.1.5).it has been shown In Figure3

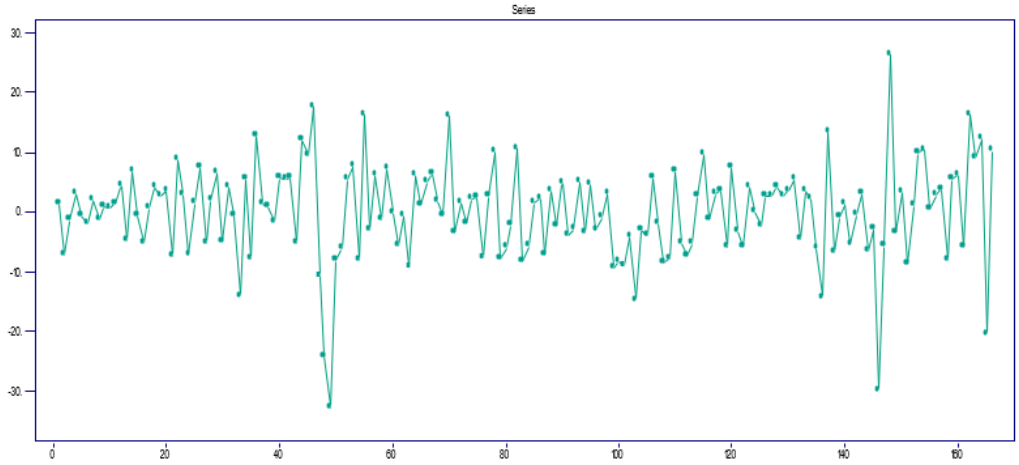

Figure 3: the remainder of the ,ARIMA process (5.1.5)

14.2 ForecastingARIMA process (5.1.5)

In this step, 24 predictions are made out of sample for theARIMA process 5.1.1it has been shown

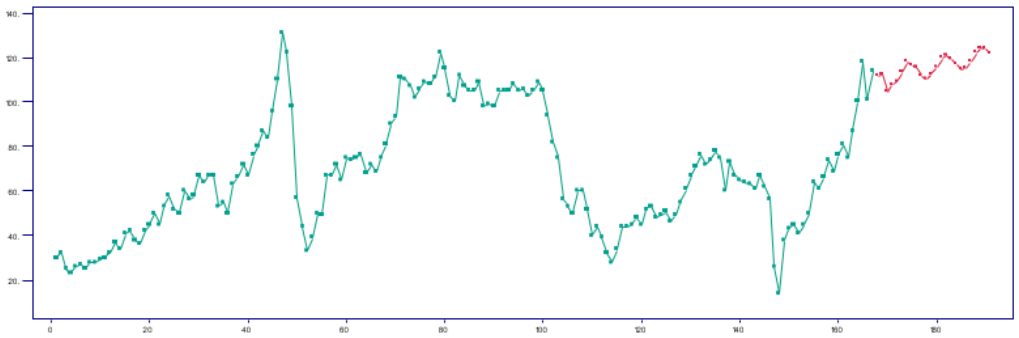

Figure 4: process prediction (5.1.5)

White noise

Table 3: ARIMA(p.d.q) models for oil price

Prediction step	Forecast month	Estimated price
1	2022/6	111.56
2	2022/7	112.34
3	2022/8	104.74
4	2022/9	107.70
5	2022/10	109.07
6	2022/11	113.11
7	2022/12	117.90
8	2023/1	116.50
9	2023/2	115.47
10	2023/3	111.62
11	2023/4	110.15
12	2023/5	112.27
13	2023/6	115.39
14	2023/7	119.70
15	2023/8	120.90
16	2023/9	119.55
17	2023/10	116.83
18	2023/11	114.46
19	2023/12	115.13
20	2024/1	117.97
21	2024/2	121.78
22	2024/3	124.20
23	2024/4	123.83
24	2024/5	121.59

14.3 ARIMA model estimation (4.1.5)

$$Zt = 0.2251Z(t-1) + .05325Z(t-2) - .5725Z(t-3) - .3627Z(t-4) + a(t) - .05883a(t-1) - .03714a(t-2) + .4618a(t-3) + 0.6256a(t-4) - .08309a(t-5) \tag{14.2}$$

process control(4.1.5)

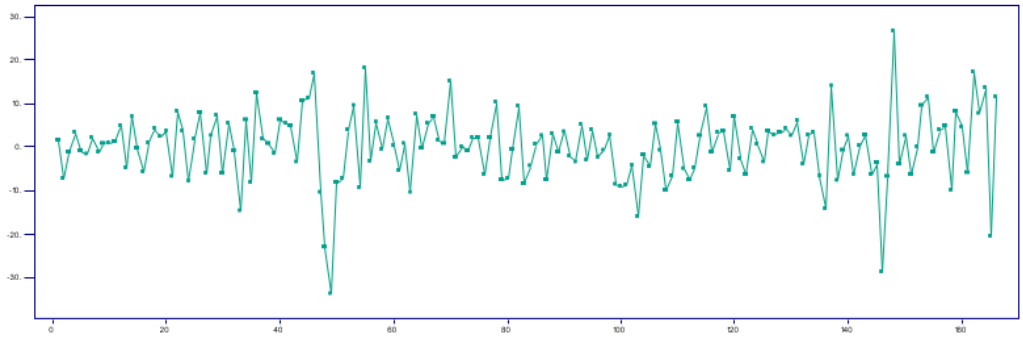

Figure 5: ARIMA process residual(4.1.5)

Data points = 166
Sample Mean = -0.0206
Sample Variance = 63.214203
St. Error (Sample Mean) = 0.574345

14.4 ForecastARIMA process 4.1.5

Table 4: ARIMA forecast 4.1.5

Prediction step	Forecast month	Estimated price
1	2022/6	116.53
2	2022/7	113.85
3	2022/8	105.14
4	2022/9	109.66
5	2022/10	109.68
6	2022/11	114.96
7	2022/12	117.56
8	2023/1	117.61
9	2023/2	115.57
10	2023/3	112.55
11	2023/4	111.62
12	2023/5	113.24
13	2023/6	116.87
14	2023/7	120.23
15	2023/8	121.43
16	2023/9	120.05
17	2023/10	117.40
18	2023/11	115.67
19	2023/12	116.33
20	2024/1	119.24
21	2024/2	122.72
22	2024/3	124.75
23	2024/4	124.33
24	2024/5	122.13

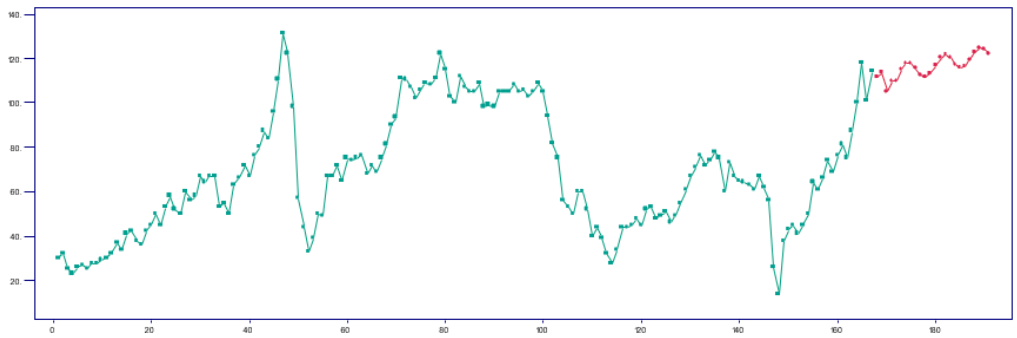

Figure 6: ARIMA process prediction(4.1.5)

15 ARIMA model estimation (3.1.5)

$$Z(t) = .6195Z(t-1) - .09256Z(t-2) - .6607Z(t-3) + a(t) - .4575a(t-1) + .06614a(t-2) + .5387a(t-3) + .2857a(t-4) - .1396a(t-5) \tag{15.1}$$

16 ARIMA process control (3.1.5)

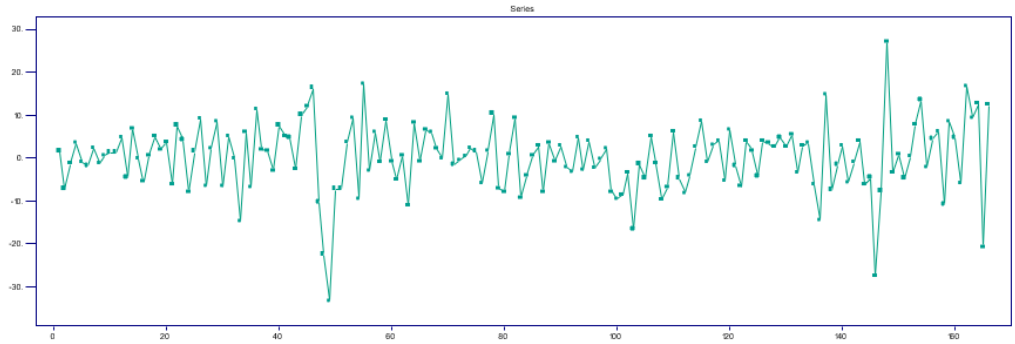

Figure 7: the remainder of the ,ARIMA process (3.1.5)

Sample Mean = -.0197
Sample Variance = 63.291775
Std.Error (Sample Mean) = .576532

17 ARIMA process forecasting (3.1.5)

Table 5: ARIMA forecast 4.1.5

Prediction step	Forecast month	Estimated price
1	2022/6	112.46
2	2022/7	114.06
3	2022/8	106.12
4	2022/9	109.09
5	2022/10	109.45
6	2022/11	115.21
7	2022/12	117.37
8	2023/1	118.50
9	2023/2	115.77
10	2023/3	113.13
11	2023/4	111.56
12	2023/5	113.22
13	2023/6	116.71
14	2023/7	120.33
15	2023/8	121.72
16	2023/9	120.52
17	2023/10	117.83
18	2023/11	115.93
19	2023/12	116.36
20	2024/1	119.16
21	2024/2	122.69
22	2024/3	124.90
23	2024/4	124.76
24	2024/5	122.56

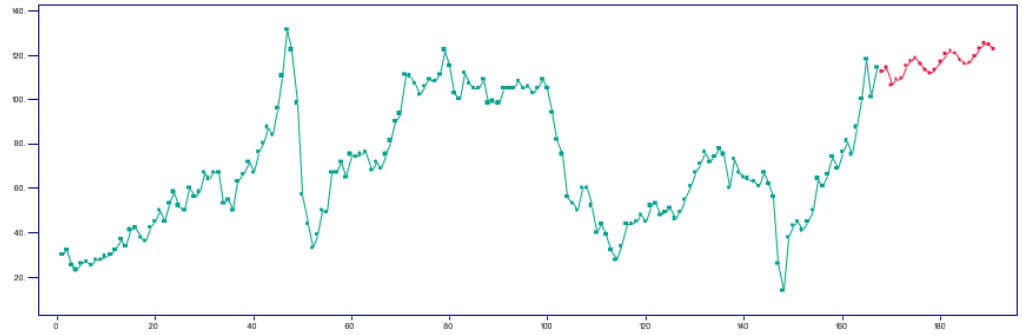

Figure 8: Forecast byARIMA process (3.1.5)

18 model estimation(5.1.3)

$$Z(t) = .7881Z(t-1) - .2189Z(t-2) - .7618Z(t-3) + .2800Z(t-4) - .1449Z(t-5) + a(t) - .6256a(t-1) + .1715a(t-2) + .6454a(t-3) \tag{18.1}$$

19 process control(5.1.3)

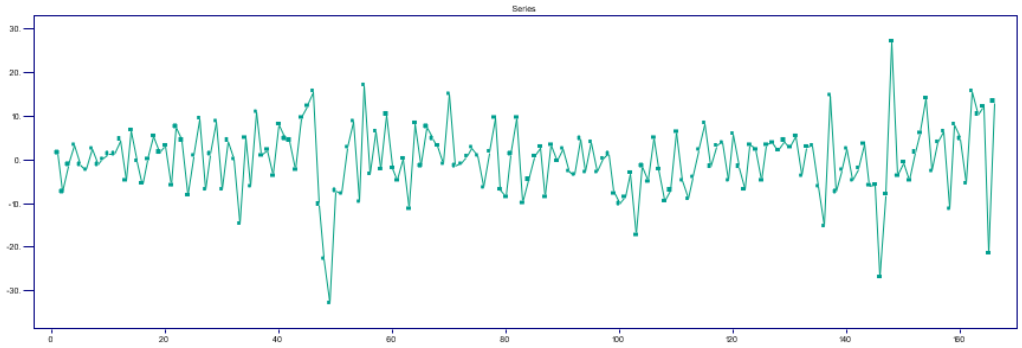

Figure 9: Remaining ARIMA process(5.1.3)

Sample Mean = -.0195
Sample variance = 63.629162
Std . Error (Sample Mean) = .581731

20 ForecastARIMA process (5.1.3)

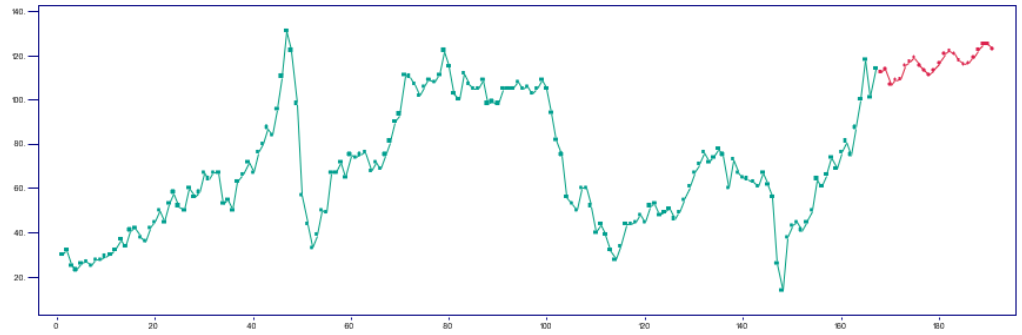

Figure 10: ARIMA process prediction(5.1.3)

Table 6: forecastingARIMA process (5.1.3)

Prediction step	Forecast month	Estimated price
1	2022/6	112.16
2	2022/7	113.74
3	2022/8	106.76
4	2022/9	108.7
5	2022/10	109.18
6	2022/11	115.49
7	2022/12	117.29
8	2023/1	119.07
9	2023/2	115.65
10	2023/3	113.42
11	2023/4	111.19
12	2023/5	113.30
13	2023/6	116.46
14	2023/7	120.60
15	2023/8	121.80
16	2023/9	120.88
17	2023/10	117.85
18	2023/11	115.98
19	2023/12	116.15
20	2024/1	119.11
21	2024/2	122.64
22	2024/3	125.09
23	2024/4	124.86
24	2024/5	122.79

21 Conclusion

Forecasting is one of the successful management tools and a key element in management and economic planning. Oil prices in OPEC member countries a A very important and influential macro-economic variable on various internal and external sectors of a country's economy, such as the balance of payments and the strength of international competition , plays a decisive role in economic policies. Oil price changes affect different sectors of a country's economy. Therefore, modeling and predicting the future trend of this variable seems to be Considering the importance of- . necessary for providing economic policies and guidelines predicting oil prices, the main goal current study Designing a hybrid computational intelligence.model for OPEC oil price forecasting
OPEC oil price data was used in order to design a computational intelligence model . for ,After stating the introduction and generalities of the research in the first chapter, the ,this purpose literature of the subject was examined in terms of theoretical and empirical foundations in the second chapter. In the third chapter, the research method was presented in detail. The fourth chapter to estimate the proposed model and test.related to it was allocated

1- What is the accuracy of forecasting methods based on Box Jenkins methodology for predicting the price of OPEC crude oil?

Based on the results obtained from the research among theARIMA models models 4

ARIMA 5.1.5, ARIMA 4.1.5, ARIMA 3.1.5, ARIMA 5.1.3They had better accuracy than other methods. Quasi analysis showed thatMSE forARIMA model 5.1.5was obtained equal to 61.86, theAIC index was equal to 1184 and theBIC index was equal to 1175, theMSE for the Arima process was4.1.5 It was equal to 63.215 and the AIC index was equal to 1194andBIC was equal to 1184. The accuracy of the ARIMA 3.1.5 process was equal to 63.2919, and theAIC and BICvalues were equal to 1202 and 1196, as well as theMSE error rate ,5.1.3ARIMA IS equal to 63.6299, and theAIC and BICvalues were equal to 1214 and 1198. After differentARIMA processes four models entered the ,

combinathon stage.

The results of this part are in line with the results obtained in the studies of Karakart et al. (2021), Zhang and Hong (2022) and Smich et al. (2021).

2- Among the models based on Box Jenkins methodology, which model is more accurate?

Four indices were used to measure the accuracy of forecasting processes under Box Jenkins methodology, which included MSE squared error, MAPE average percentage of absolute value of error , andAIC andBIC indices . Among the 25 models that were studied in this .researchARIMA process 5.1.5 It had better accuracy than other models. Thus, the amount of MSE,errorMAPE error, and AIC andBIC .index WERE lower.

The results of this section are also consistent with the results of the studies of Karakart et al. (2021), Smich et al. (2021), Liu et al. (2021), Zhang and Hong (2022), and Abd et al. (2020).

3- How accurate is the Holt Winters forecasting method for predicting the price of OPEC crude oil?

In order to predict using Holt-Winters method in this research, alpha coefficient equal to0.96 beta equal to ,0.34 and gamma equal to1 The accuracy of the prediction .were obtained method showed that the error rate of Holt-Winters method based on theMAPE index is equal to 17.25%. which is relatively high. In the studies ofAbdullah et al. (2020) and Fan et al. (2021), this high index was obtained

4- Among the forecasting methods of Bucks Jenkins and Halt Winters, which one is more accurate in predicting OPEC crude oil prices?

Comparing the accuracy of Halt Winters and Box Jenkins methodology showed that the error rate of Halt Winters is higher than that of Box Jenkins methodology.

In the research of Edward and et al. (2022), Gerrard et al. (2022), Yousefi et al. (1400), Omidi et al. which is consistent with the results of this research.

5- What is the proper design of the neural network for combination to achieve high Accuracy?

MSE index showed that the best design network for prediction includes 5 input neurons, 1 output neuron and 5 intermediate neurons, the best activation distribution is the linear sigmoid function and the optimal number of layers includes 1 outer layer, 1 intermediate layer and 1 layer is the output. In the research of Frank et al. (2021), Jistou et al. (2021), Azar et al. (1401), Omidi et al., the most suitable network designed to combine prediction had an input layer, an intermediate layer, and an output layer. which is consistent with the present .research

6- To what extent does creating the second stage of forecasting and combining forecasting methods improve the accuracy of forecasting methods?

The use of theAnn-Arima - Holt combined method was more accurate than the single use of forecasting methods. in a way that compared to the best forecasting process that isarima 5.1.5 Error index MSE went from 61.86 to 46.69. Also , theAIC index went from 1184 to and the 1014BIC index went from 1175 to 1012. Also meanMAPEreached 6.41%, that is, the amount of predicted values using the combined method is equal to 93.59%, which shows .the high accuracy of this method In their research, Frank et al. (2022) showed that the use of a combined algorithm has a higher accuracy than the use of individual methods. In their research, Gerard et al. mentioned above. Omidi et al. (1401) named the accuracy of artificial neural hybrid .forecasting methods higher than the time series based on Box Jenkins

References

[1] Haqit, F., July, F. (2019). Prediction of passenger movement indicators using the Helthwinters Markov chain model.

[2] Khajovi, S., Ghiori-Moghadam, A. (1400). Investigating the application of data coverage analysis in comparing and evaluating the performance of business units. Accounting Research, 16-1 ,(0)2.

[3] Sinaii, H., Gustasbi-Maharloui, R. (2019). Evaluating the efficiency and relative performance of companies with the data coverage analysis approach in order to form a stock portfolio. .Accounting Knowledge, 11(3), 105-132

[4] Shakrian, H. (1400). Designing a new model using the combination of two balanced scorecard models and data coverage analysis.

[5] Mousavi Jahormi, Y. (1400). Predicting value added tax on gasoline consumption. Scientific- .Research Quarterly, Economic Growth and Development Research, 6(21), 107-119

[6] Al-Musaylh, MS, Deo, RC, Adamowski, JF, & Li, Y. (2018). Short-term electricity demand forecasting with MARS, SVR and ARIMA models using aggregated demand data in Queensland, Australia. Advanced Engineering Informatics, 35, 1-16. doi: https://doi.org/10.1016/j.aei.2017.11.002

[7] Cook, WD, Ruiz, JL, Sirvent, I., & Zhu, J. (2017). Within-group common benchmarking using DEA. European Journal of Operational Research, 256(3), 901-910. doi: https://doi.org/10.1016/j.ejor.2016.06.074

[8] Debnath, KB, & Mourshed, M. (2018). Forecasting methods in energy planning models. Renewable and Sustainable Energy Reviews, 88, 297-325. doi: https://doi.org/10.1016/j.rser.2018.02.002

[9] Farhad, H., Lotfi, Hatami-Marbini, A., Agrell, P., Aghayi, N., & Gholami, K. (2012). Allocating fixed resources and setting targets using a common-weights DEA approach (Vol. 64).

[10] Graefe, A., Armstrong, J., J. Jones, R., & Cuzán, A. (2014). Combining Forecasts: An Application to Elections (Vol. 30).

[11] Kalate Ahani, I., Salari, M., & Shadman, A. (2018). Statistical models for multi-step-ahead forecasting of fine particulate matter in urban areas. Atmospheric Pollution Research. doi: https://doi.org/10.1016/j.apr.2018.11.006

[12] Li, Z., Rose, J., & Hensher, D. (2010). Forecasting automobile petrol demand in Australia: An evaluation of empirical models (Vol. 44).

[13] Ruiz, JL, & Sirvent, I. (2016). Common benchmarking and ranking of units with DEA. Omega, 65, 1-9. doi: https://doi.org/10.1016/j.omega.2015.11.007

[14] Sen, P. ,Roy, M., & Pal, P. (2016). Application of ARIMA for forecasting energy consumption and GHG emission: A case study of an Indian pig iron manufacturing organization. Energy, 116, 1031-1038. doi: https://doi.org/10.1016/j.energy.2016.10.068

[15] Shukur, OB, & Lee, MH (2015). Daily wind speed forecasting through hybrid KF-ANN model based on ARIMA. Renewable Energy, 76, 637-647. doi: https://doi.org/10.1016/j.renene.2014.11.084

[16] Toloo, M., & Mensah, EK (2018). Robust optimization with non-negative decision variables: A DEA approach. Computers & Industrial Engineering. doi: https:// doi.org/10. 1016/j.cie.2018.10.006

[17] Wang, X., & Meng, M. (2012). A Hybrid Neural Network and ARIMA Model for Energy Consumption Forecasting (Vol. 7).