Iraqi Statisticians Journal

https://isj.edu.iq/index.php/isj

ISSN: 3007-1658 (Online)

Application of Neural Networks for Predicting the Exchange Rate of the Iraqi Dinar Against the US Dollar and Comparison with the Box-Jenkins Method for Time Series 2015-2022

Fatimah Abdul – Hammeed jawad AL-Bermani

University of Baghdad, Administration &Economy College, Baghdad, Iraq

ARTICLE INFO

Article history:
Received 30 November 2024
Revised 30 November 2024 ,
Accepted 06 January 2025,
Available online 06 January 2025

Keywords:
Training set,
Validation set,
Neural Network architecture,
Time Series

ABSTRACT

Recently, there have been changes in the exchange rate of the Iraqi dinar against foreign currencies, which are considered important financial indicators that affect the labor market and the currency exchange market. In order to monitor the changes in the exchange rate of the Iraqi dinar against the US dollar and to anticipate future stages and the direction of the exchange rate changes, the aim of the research is to predict the exchange rate of the Iraqi dinar against the US dollar for the coming years by applying the Box-Jenkins methodology and neural networks. This is done to compare the traditional predictive models, such as ARIMA, with the neural network model, which demonstrated its prediction accuracy by reducing the Mean Squared Error (MSE) through training the network, selecting the appropriate model, and choosing the best architecture to represent the time series.

The study included a time series representing the exchange rates of the Iraqi dinar against the US dollar from January to December for the years 2015-2022. The data was sourced from the 2021/2022 annual statistical group issued by the Central Statistical Organization of Iraq, with data from the Central Bank of Iraq. For the analysis of the time series, the GRETL statistical program was used, and the Matlab R 2019b program was utilized for forecasting when using neural networks.

1. Introduction:

Forecasting is one of the important statistical methods for decision-making and has a significant impact on future planning. It encompasses various fields (health, agriculture, industry, finance, and population studies) as well as numerous phenomena .[11]

Given the fluctuations in the exchange rate of the Iraqi dinar against the US dollar and their impact on the national economy, these changes in prices can be observed over previous years to develop future plans for predicting the values of this phenomenon. Time series analysis is considered one of the methods used to study the behavior of phenomena and forecast changes in prices, enabling the formulation of future plans.

Therefore, the Box-Jenkins methodology is considered one of the most important approaches used in time series analysis and forecasting for future years. Recently, some studies have focused on the application of the Box-Jenkins methodology as well as employing neural networks for forecasting and comparing the two methods.

Corresponding author E-mail address: Fatimah.a@coadec.uobaghdad.edu.iq
https://doi.org/10.62933/egnts624

The researcher Imad Yacoub Hamid (2011) compared the Box-Jenkins models with neural networks, and the results showed the superiority of neural networks in forecasting. The application was conducted on data from the Sudanese agricultural sector, represented by the time series of wheat productivity.[15]

The researchers Salioua and Matar (2019) conducted a comparison between the Box-Jenkins methodology and artificial neural networks using the monthly average data of maximum temperatures for the city of Mosul (1983–2009). The results demonstrated the accuracy of artificial neural networks in prediction.[25]

Researcher Majd Naama (2023) compared traditional predictive models, including the multiple regression model, the ARIMA model, and the artificial neural network model, in terms of predictive capability for tobacco production in Latakia Governorate. The study utilized data from the annual agricultural statistical reports on production and cultivated area for the period (1991–2019). The results revealed the accuracy of artificial neural networks in prediction.[18]

2. Time Series

There are many economic and social phenomena that occur over successive or equal time periods, showing the effects of time, which may be increasing or decreasing. A time series can be defined as the values of a phenomenon arranged according to time; the time intervals can be consecutive and equal (annual, quarterly, monthly, daily, etc.). The nature of the changes that occur in the values of the phenomenon makes it possible to analyze, estimate, and forecast the time series [22].

The goal of time series is to accurately describe the phenomenon and explain the changes occurring in the data and the influencing factors, thus building a model for forecasting based on the changes occurring over a period of time according to the data of the phenomenon [22, 1].

3. Time Series Components

A time series has four components that are influenced by economic, environmental, social, and political factors, as follows:

1. Secular Trend

The secular trend describes the general effect of the phenomenon over a period of time, where the series may be increasing, decreasing, or stable [6].

2. Seasonal Variations

These are changes that occur within a year in a regular pattern. Climate and weather conditions are among the factors causing seasonal changes, as temperature or rainfall can influence seasonal variation [3].

3. Cyclical Variations

These variations manifest as rises or falls in the general trend of the time series values and occur over a long period in an irregular manner, but they may recur [6].

4. Randomness Variations

These are irregular changes that occur in the general trend of the time series and happen randomly by chance, making them unpredictable and uncontrollable, such as earthquakes, volcanoes, floods, and wars [3].

4. Box-Jenkins Models

The Box-Jenkins method for time series analysis is an advanced statistical approach for forecasting. Box and Jenkins (1970) introduced a method consisting of several stages, including identifying the time series, diagnosing and estimating, and finally forecasting. This method allows for the selection of the appropriate model based on autoregressive (AR), moving average (MA), and mixed (ARMA) models [7].

1) **Autoregressive Model (AR)**
The Autoregressive model (AR) defines the current value of the time series y_t as a function of its previous values plus random error terms. It can be expressed in the following form:

$y_t = \mu + \vartheta_1 y_{t-1} + \vartheta_2 y_{t-2} + \ldots + \vartheta_p y_{t-p} + e_t$ …(1)

Where:

y_t: Represents the value of the time series at time t

y_{t-i} : Represents the value of the time series at time t , where i=1,2,…,p

ϑ_i : Model parameters to be estimated, ranging from (1,-1) ,where i=1,2,…,p

e_t : Represents random errors, which have a mean of zero and variance σ_u^2

μ: The constant term

If the PACF displays a sharp cutoff while the ACF decays more slowly (i.e., has significant spikes at higherlags), we say that the stationarized series displays an AR [26].

2) **Moving Average Model (MA)**

The Moving Average model (MA) expresses the current value of the time series y_t in terms of random error terms (e_t, e_{t-1},…) upon which the model relies. The general form of the model of order q is denoted as MA(q) and can be written as follows [26]:

$y_t = u + e_t - \vartheta_1 e_{t-1} - \vartheta_2 e_{t-2} - \ldots - \vartheta_q e_{t-q}$ …(2)

Where :

u : The constant term

y_t : Represents the value of the time series at time t

q : model order

e_{t-i} : Random variations that are independent of each other at time t (i=.1,2,…,q)

ϑ_i : Model parameters to be estimated, ranging from (1,-1) ,where i=1,2,…,q

The (PACF) decreases exponentially or in a damped sinusoidal pattern, and (ACF) for the model MA (q) cuts off at zero after the lag, determining the model order.

3) **Mixed Autoregressive-Moving Average Model (ARMA)**

The ARMA model combines the AR(p) and MA(q) models, incorporating characteristics of both types to achieve a more flexible model . It is denoted as ARMA(p, q) and its general form is as follows:

$y_t = u + \vartheta_1 y_{t-1} + \vartheta_2 y_{t-2} + \ldots + \vartheta_p y_{t-p} + e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \ldots - \theta_q e_{t-q}$ … (3)

Where:

u : The constant term

y_t : Represents the value of the time series at time t

y_{t-i} : The value of the time series at time t , (i=1,2,…,p)

ϑ_i *:* The parameter autoregressive coefficients i=1,2,…,p

θ_j : The parameter moving average coefficients j=1,2,…,q

u_t : Represents random errors with a mean of zero and a variance of σ_u^2
The PACF values are decreasing exponentially or in diminishing sinusoidal waves, while the ACF for the MA(q) model cuts off at zero after lag q, which determines the order of the model [23].

4) Autoregressive Integrated Moving Averages Model (ARIMA)

This model transforms the non-stationary time series into a stationary time series after taking differences d, which is the degree of integration to stabilize the series. It is one of the most commonly used models for forecasting, denoted as ARIMA(p,d,q), where ppp represents the order of the AR model and q represents the order of the MA model [20].

$y_t = u + \vartheta_1 y_{t-1} + \vartheta_2 y_{t-2} + \ldots + \vartheta_p y_{t-p-d} + e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \ldots - \theta_q e_{t-q}$ … (4)

Where :

y_t : The value of the time series at time t

μ : The mean or constant term of the series

ϑ_i: The autoregressive coefficients i=1,2,…,p

θ_j: The moving average coefficients j=1,2,…,q

d: The degree of differencing to make the series stationary

5. **Reasons for using Box-Jenkins models** [4,14]:

1. The Box-Jenkins methodology addresses both univariate models and multivariate models stationary and non – stationary .
2. They suit complex time series and forecasting situations that provide various patterns, enabling the selection of an appropriate model and minimizing error as much as possible.

6. **Box-Jenkins Methodology**

To forecast using the Box-Jenkins methodology, there are four stages to reach the forecast as follows:

1. Checking the stationarity of the time series, and applying the necessary
2. Model identification
3. Checking the model's adequacy
4. Diagnosis checking the model's adequacy and forecasting

Stage One: Testing for the stationarity of the time series

A time series is considered stable if it oscillates around a constant mean with constant variance. Conversely, if it oscillates around a non-constant mean or has non-constant variance, it is deemed unstable. There are three conditions for achieving complete stationarity:[16]

1 - $E(y_t) = \mu$,The mean value must be constant
2 - $Var(y_t) = \sigma^2 = \hat{\gamma}_0$,The variance must be constant

where

$\hat{\gamma}_0 = \frac{1}{n}\sum_{t=1}^{n}(y_t - \bar{y})^2$, represents the estimate of the variance of the time series, which is constant and does not depend on the values of t
3 - The autocovariance function must also be constant

For seasonal time series, the autocovariance function at seasonal lag k (where the seasonality period is m) can be expressed similarly, but with seasonal lags:

$\gamma_k = E[(y_t - \mu)(y_{t+k} - \mu)]$ …(5)

where :

γ_k is the **autocovariance** at lag k

y_t is the value of the time series at time t

y_{t+k} is the value of the time series at time t+k

μ is the **mean** of the time series

However, in seasonal data, the lag k is usually associated with the seasonality period m.

For a seasonal period of m, the lag k will typically be a multiple of m, such as m,2m,3m,…

However, in seasonal data, the lag k is typically associated with the period of the seasonality, so if the time series exhibits seasonal patterns with a period m, the lag k corresponds to a multiple of m (e.g., after m periods, the series is expected to repeat its pattern). For seasonal series, this could involve seasonal differencing or adjustments to account for this seasonality.

In cases of non-stationarity, which are common in many models, the series may exhibit a general trend or seasonality. This can be observed through the autocorrelation and partial autocorrelation functions, where their values approach zero after the second or third lag, while remaining large for several lags. However, it is possible to transform them into stationary time series, which can be done in two ways:

1. **Stationarity in Mean**:

   Non-stationarity around the mean indicates that the time series does not fluctuate around a constant mean, which can be removed after appropriate differencing. A stable time series can be achieved after taking differences d, such that

$$W_t = \nabla^d \, y_t \qquad \ldots(6)$$

∇ represents the backward difference operator defined as:

$$\nabla y_t = (1\text{- B})y_t = y_t \text{ - } y_{t-1} \qquad \ldots(7)$$

The general form of the differencing d can be written as:

$$\nabla^d y_t = (1 - B)^d \, y_t \qquad \ldots(8)$$

2. **Stationarity in Variance**:

Stationarity is achieved when there are no varying fluctuations in the form of the time series. When the fluctuations are non-constant, the series is considered unstable. Therefore, it is transformed into a stable series through logarithmic or exponential transformations (power transformations), which are the simplest transformations, defined as follows:

$$y_t^{\lambda} = \begin{bmatrix} y_t^{\lambda} & if\ \lambda \neq 0 \\ Lny_t & if\ \lambda = 0 \end{bmatrix} \qquad \ldots(9)$$

Where:

y_t^{λ} : The transformed series at time t

y_t : The original untransformed series value

λ : The transformation parameter, usually ranging from (1≤ $\lambda \leq -1$)

The purpose of transforming the original data is to obtain residuals with constant (homogeneous) variance.

To check for stationarity, several tests can be conducted to assess the stability of the time series, as follows:

First: The Autocorrelation Function (ACF)

The ACF measures the degree of relationship between values of the same variable over time at different lags. Its values range from(-1≤ $\rho_k \leq 1$), where ρ_k is the autocorrelation coefficient.

$$\rho_k = \frac{cov(y_t, y_{t+k})}{\sqrt{var(y_t)var(y_{t+k})}} \qquad \ldots(10)$$

Where:

ρ_k : The autocorrelation coefficient

K : The maximum lag (where k=1,2,…,K , which can be determined as :

K= $12*(\frac{n}{100})^{\frac{1}{4}}$,and n represents the number of observations.

To achieve stationarity, the autocorrelation coefficients must fall within the confidence interval at a 95% level and a significance level of 0.05. If they fall outside this interval, the series is considered unstable. The autocorrelation coefficients are normally distributed, expressed as:

$\rho_k \sim N(0,\frac{1}{\sqrt{n}})$, The test formulation is as follows:

$H_0 : \rho_k = 0$
$H_1 : \rho_k \neq 0$

If ρ_k falls within the confidence interval, the null hypothesis is don't reject ; if it falls outside the confidence interval, the alternative hypothesis is accepted [11].

Second: The Partial Autocorrelation Function (PACF)

The PACF measures the relationship between the autocorrelations (y_t, y_{t+k}) and helps determine the order and type of the model. It can be estimated using the least squares method or a set of approximate equations. The mathematical formula for estimating the partial autocorrelation coefficients is as follows:

$\widehat{\emptyset}_{k+1,k+1} = \frac{\hat{\rho}_{k+1}-\sum_{j=1}^{k}\widehat{\emptyset}_{kj}\hat{\rho}_{k+1-j}}{1-\sum_{j=1}^{k}\widehat{\emptyset}_{kj}\hat{\rho}_{j}}$, j=1,2,…,k , and when k=1 …(11)

$\widehat{\emptyset}_{11} = \hat{\rho}_1$

$\hat{\rho}_j$: Estimates of the autocorrelation coefficients

The PACF can be used to determine whether the time series is stable and to specify the degree of the AR model, the MA model, or the appropriate ARIMA model to represent the time series data [10].

Third: Unit Root Test

The condition for stationarity is satisfied when the unit roots of the series lie inside the unit circle. One of the most commonly used methods to detect the stationarity of a time series is the Augmented Dickey-Fuller (ADF) test, which is based on three mathematical equations that assume the existence of a random process Y_t .

The first equation does not include a constant term (test without constant):

$\Delta\text{--}Y_t = \emptyset_1\text{---}Y_{t-1} + \sum_{j=1}^{p} B_j\,\Delta Y_{t-j} + e_t$ … (12)

P : changes in time series values for time periods

The second equation includes a constant term (test with constant):

$\Delta Y_t = \emptyset_0 + \emptyset_1 Y_{t-1} + \sum_{j=1}^{p} B_j\,\Delta Y_{t-j} + e_t$ …(13)

The third equation includes both a constant term and a time trend (with constant and trend):

$\Delta - Y_t = \emptyset_0 + \emptyset_1 - Y_{t-1} + \sum_{j=1}^{p} B_j\,\Delta Y_{t-j} + \delta t + e_t$ …(14)

Notations:

Y_t : The random process

Δ : The differencing operator (i.e , $\Delta Y_t = Y_t - Y_{t-1}$)

e_t : A series of random variables

$\emptyset$,B: Test parameters

$\emptyset_0$: The constant term

δt : The time trend

7. Hypothesis Testing

The hypotheses are expressed as follows:

$H_0: \emptyset_1 = 0$: (unstable, presence of a unit root)

$H_1: \emptyset_1 \neq 0$: (stable, absence of a unit root)

The test statistic is compared as: $t = \frac{\emptyset_1}{SE\emptyset 1}$ with tabulated values (Dickey-Fuller tables). If the computed t value is greater than the tabulated value, the null hypothesis is rejected, and the alternative hypothesis $H_1: \emptyset_1 \neq 0$ is accepted, indicating that the series is stable. Alternatively, if the p-value is less than 0.05, the alternative hypothesis is accepted, suggesting the series is stable [13, 26].

8. Determining the Order of the Model

There are several criteria for determining the order of the model, as follows:

1) Akaike Information Criterion (AIC)

The AIC, proposed by the Japanese scientist Hirotugu Akaike in 1974, aims to minimize the difference between the model density and the true density (observations), reducing the model's variance relative to the increase in the number of estimated parameters, expressed as [8]:

$$AIC = n\text{Log}\hat{\sigma}_u^2 + 2V \quad \ldots(15)$$

Where:

V: The total number of estimated parameters

$\hat{\sigma}_u^2$: The variance of the error

n : The number of observations

2) Hannan-Quinn Criterion (H-Q)

Proposed by researchers Hannan and Quinn in 1979, this criterion is abbreviated as H-Q, and its formula is as follows [1]:

$$H - Q(M) = \text{Ln}\sigma_u^2 + 2MC\text{Ln}(\text{Ln}(n))/n \quad , C > 2 \quad \ldots(16)$$

M : Represents the number of parameters in the model

C : Constant

3) Schwartz Bayesian Criterion (SBC)

Abbreviated as SBC, this criterion was proposed by researcher Schwartz in 1978, similar to the Bayesian Information Criterion (BIC). Its formula is as follows [2]:

$$SBC(P) = n\text{Ln}(\hat{\sigma}_u^2) + p\text{Ln}(n) \quad \ldots(17)$$

Where:

P : Represents the order of the chosen model or the number of model parameters

9. Estimating Model Parameters

After diagnosing the model and determining its order, the parameters are estimated. In the case of the ARIMA model, estimating the parameters can be complex and may not be straightforward; thus, the Maximum Likelihood Method is employed. This method is used when dealing with moving average processes that have unknown error bounds, aiming to minimize the sum of squared errors [13].

10. Testing Model Significance

This involves testing the error term and verifying the validity and adequacy of the specified model. The residuals resulting from applying the model should be randomly distributed. To determine whether the autocorrelation is randomly distributed, the Ljung-Box test is applied [3].

$$Q = m(m+2)\sum_{i=1}^{k} \frac{r_i^2\,(e)}{(m-i)} \quad \ldots(18)$$

Where :

m=(n-d)

n : number of time series views

d : number of differences

k : largest displacement

r_i(e) : autocorrelation of statistics at displacement

11. Residual Testing

This is used to assess whether the model is adequate based on the values of the autocorrelation coefficients of the residuals within a 95% confidence interval [22].

$$p_r\{\hat{r}_k(u)<1.96\frac{1}{\sqrt{r}}\} = 1- \alpha \qquad \ldots(19)$$

12. Forecasting

To predict future values after determining the appropriate model and estimating its parameters, which is considered the final stage in the Box-Jenkins methodology. The forecasted values can be obtained by taking the expectation at time t, expressed as [4]:

$$\hat{y}_{t+L} = \text{E}\ [y_{t+L}] \quad , \quad y_t,\ y_{t-1},\ y_{t-2},\cdots \quad \ldots(20)$$

13. Neural Networks

Neural networks have achieved significant advancements in many applications. Their concept revolves around simulating the human brain to make decisions and improve relationships between the elements used, as well as for forecasting [21].

In general, neural networks consist of the following layers [12, 21]:

1. **Input Layer**: This layer receives data through the neurons that comprise the network and contains only one layer.
2. **Output Layer**: This is where the final output is generated, and the architecture of the network depends on whether there is one or more processing units.
3. **Hidden Layer**: This layer performs processing and mathematical operations and sends the results to the output layer.
4. **Weights**: Weights are responsible for connecting the layers to each other, facilitating the transfer of data between the units.

14. Processing Units (Neurons) [19]

Weights are the primary element connecting the layers. The activation function is a value constrained within a specific range, comparing the sum with a threshold value, and its range is between [0, 1] and [-1, 1]. The following are common activation functions [17, 24]:

1. **Sigmoid Function**: This function transforms the outputs into a value bounded between [0, 1] and [-1, 1]. It is known as the sigmoid activation function and is one of the most widely used.[9]
2. **Step Function**: This function makes the output value equal to 0 or 1, known as the binary activation function.
3. **Linear Function**: The outputs equal the weighted inputs for the processing unit.
4. **Sign Function**: The output value from the processing unit equals 1 or -1, used in classification tasks to distinguish patterns.
5. **Output Function**: This function adjusts the result of the activation function.

There are several types of neural networks based on the types of layers [27]:

1- Single-Layer Neural Network

A Single-Layer Neural Network is one of the simplest types of neural networks. It consists of a single layer of neurons (or units) that process inputs and generate outputs. This network is composed of three main components:

1. **Input Layer:** It contains a set of neurons that

represent the inputs. Each neuron in this layer represents a specific value from the input data.

2. **Output Layer:** It contains one or more neurons to represent the results or predictions. The network generates output values based on the inputs it receives.
3. **Weights:** Each connection between neurons in the input layer and the output layer has a weight, which is a value that is adjusted during the training process to improve the network's accuracy. These weights are modified using algorithms like backpropagation.

2- Multilayer Neural Network

A **Multilayer Neural Network (MLP)** is a more complex type of neural network compared to a single-layer network. It consists of multiple layers of neurons: an input layer, one or more hidden layers, and an output layer. MLPs are designed to solve more complicated problems and can capture nonlinear patterns in data. Here's how it works:

Components:

1. **Input Layer:**

   This layer represents the raw input data. Each neuron in the input layer corresponds to a feature of the input data.

2. **Hidden Layers:**

   These are intermediate layers that lie between the input and output layers. The number of hidden layers can vary, and each layer can have multiple neurons. The hidden layers allow the network to capture complex relationships and nonlinearities in the data. The output of each neuron in these layers is passed through an activation function, which helps the network learn and generalize.

3. **Output Layer:**

   The output layer provides the final prediction or classification based on the learned weights from the previous layers. The number of neurons in the output layer corresponds to the number of possible outputs or classes.

4. **Weights and Biases:**

   Weights determine the strength of the connections between neurons. During training, the network adjusts these weights to minimize the error in the predictions. Biases are additional parameters that help shift the activation function, improving the model's flexibility.

15. Practical Framework

The Box-Jenkins methodology was applied to forecast the exchange rate of the Iraqi dinar against the US dollar. The time series data represents monthly exchange rates from January 2015 to December 2022, with a sample size of 96 observations collected from the Central Statistical Organization (Iraqi Central Bank). The statistical software Gretl was used to obtain the results.

16. Application of the Box-Jenkins Methodology

1. **Stationarity of the Time Series**

   We start by plotting the time series to understand its behavior regarding stationarity in mean and variance, as shown in Figure (1):

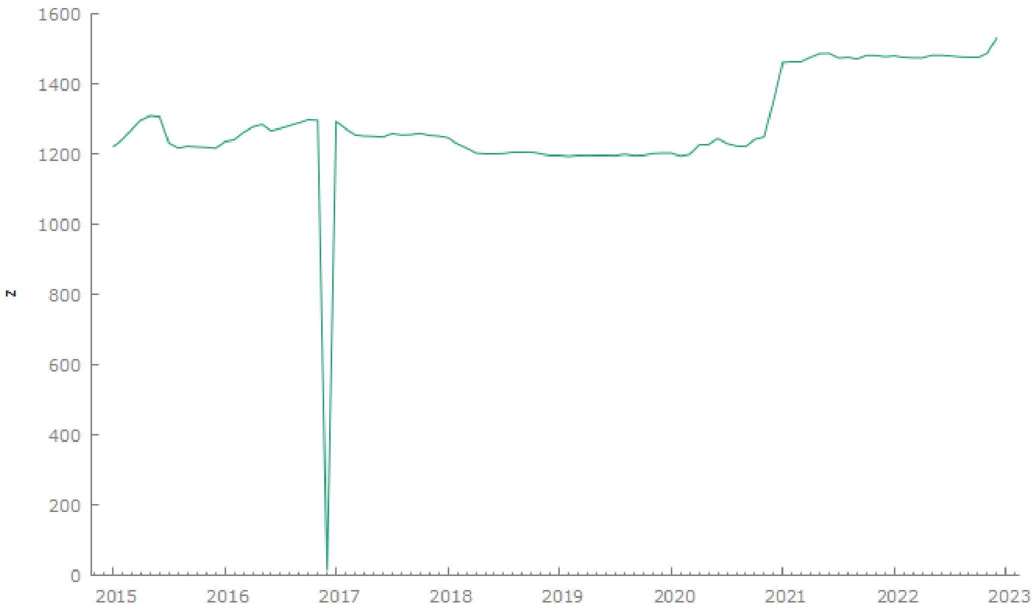


Figure (1) represents the time series plot.

Figure 1 illustrates the instability of the time series in terms of both mean and variance. A stationarity test can be conducted by plotting the Autocorrelation Function (ACF) and the Partial Autocorrelation Function (PACF) for increased accuracy. Figure (2) sequentially shows the plots of both functions based on the values of the autocorrelation and partial autocorrelation coefficients provided in Table (1) below:

Table (1): Autocorrelation and Partial Autocorrelation Functions

Lag	ACF	PACF
1	0.3919 ***	0.3919 ***
2	0.3798 ***	0.2672 ***
3	0.3724 ***	0.2018 **
4	0.3579 ***	0.1492
5	0.3431 ***	0.1115
6	0.3290 ***	0.0843
7	0.2966 ***	0.0380
8	0.2815 ***	0.0288
9	02686 ***	0.0243
10	0.2556 **	0.0200
11	0.2394 **	0.0113
12	0.2268 **	0.0092
13	0.2055 **	-0.0053
14	0.1901 *	-0.0081
15	0.1726 *	-0.0138
16	0.1615	-0.0087
17	0.1353	-0.0267
18	0.0794	-0.0803
19	0.0547	-0.0728

*Significance of the data

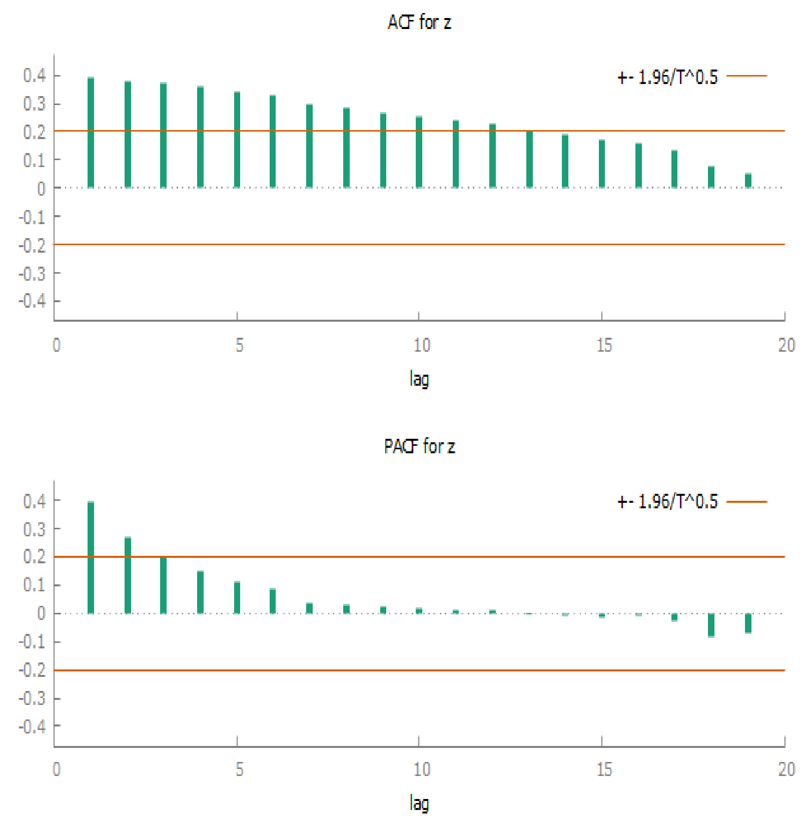


Figure (2) shows the plots of the Autocorrelation Function (ACF) and the Partial Autocorrelation Function (PACF).

Table (2) illustrates the autocorrelation and partial autocorrelation coefficients. Figure 2 indicates that the coefficient values are outside the confidence limits (±0.2), which suggests the instability of the series.

The first difference is taken to eliminate the instability of the series in terms of the mean as follows:

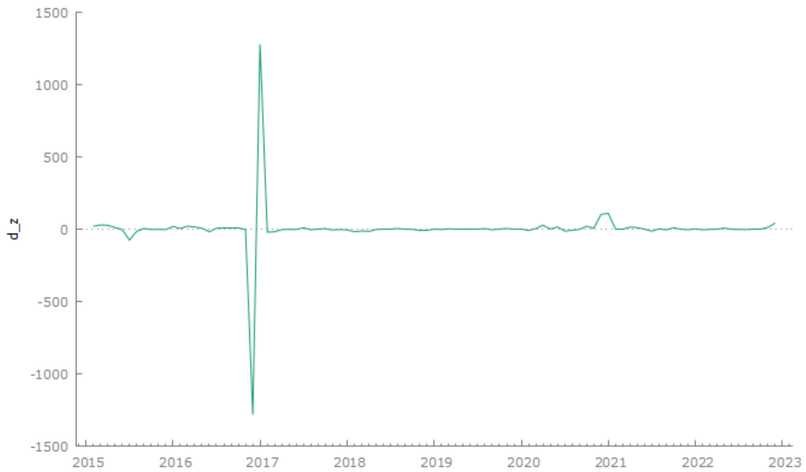
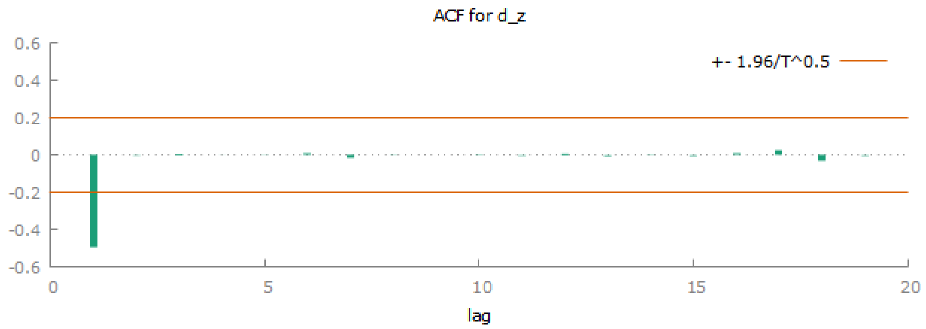


Figure (3) shows the time series plot after taking the first difference.

To confirm the stationarity of the series, the Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) are plotted, with their values illustrated in Table (2) as follows:

Table (2): Values of Autocorrelation and Partial Autocorrelation After the First Difference

Lag	ACF	PACF
1	- 0.4964 ***	- 0.4964 ***
2	- 0.0027	- 0.3307 ***
3	0.0063	- 0.2380 **
4	0.0002	- 0.1810 *
5	-0.0010	- 0.1438
6	0.0135	- 0.0942
7	-0.0151	- 0.835
8	- 0.0015	- 0.0771
9	- 0.0002	- 0.0718
10	0.0023	- 0.0623
11	- 0.0038	- 0.0605
12	0.0076	- 0.0448
13	- 0.0055	- 0.0413
14	0.0020	- 0.0345
15	- 0.0046	- 0.0372
16	0.0123	- 0.0174
17	0.0244	0.0401
18	- 0.0274	0.0337
19	- 0.0039	0.0209



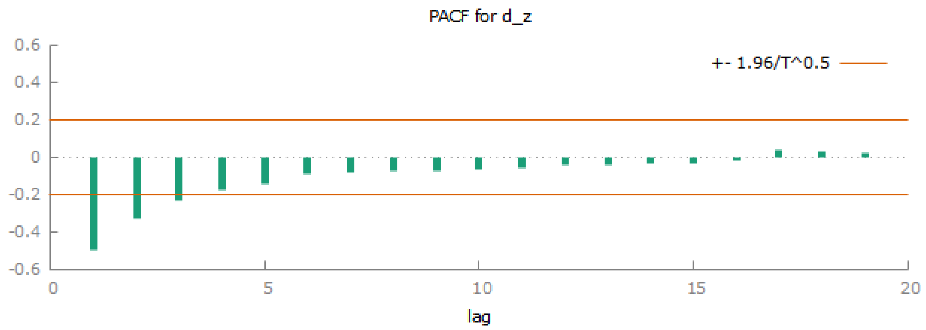


Figure (4) shows the plots of the Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) after taking the first difference.

To confirm stationarity, the Augmented Dickey-Fuller (ADF) unit root test is conducted without a constant, with a constant, and with a constant and trend. The results are presented in Table (3) as follows:

Table (3): Unit Root Test Results

	Estimated value	Test statistic	p – value
Test without constant	-2.91793	-7.88051	9.738e-14
Test with constant	-2.92688	-7.85498	1.324e-12
With constant and trend	-2.94965	-7.87206	5.129e-12

The results of the unit root test indicate that the p-value is less than the significance level of 0.05 for the equation without a constant, with a constant, and with a constant and trend. This suggests that the series is stationary after taking the first difference.

2. Diagnosis, Estimation, and Selection of the Best Model

To diagnose the suitable model for comparison, statistical criteria such as the Akaike Information Criterion (AIC), Bayesian Schwartz Criterion (SBC), and Hannan-Quinn Criterion (H-Q) are relied upon. The results are presented in Table (4) below:

Table (4): Comparison of Box-Jenkins Models

	Model	AIC	SBC	H-Q
1	(2,1,0)	1281.029	1291.202	1285.138
2	(1,2,0)	1383.021	1390.618	1383.021
3	(1,1,2)	1214.569	1227.285	1219.706
4	(1,1,1)	1238.007	1248.180	1242.116
5	(1,1,0)	1305.809	1313.439	1308.891
6	(0,1,2)	1212.919	1223.093	1217.029

The results indicate that the ARIMA (0,1,2) model is the best, as it has the lowest errors among the three criteria. The estimation of the parameters for the best model is shown in Table (5) below:

Table (5): Estimation of Parameters for the ARIMA(0,1,2) Model and Their Significance Values

Model Parameters	coefficient	Std.error	Z	p- value
Const	0.197434	0.0379785	5.199	2.01e-07***
Theta	-1.99942	0.00282689	-707.3	0.0000***
Theta - 2	0.999978	0.0278446	359.1	0.0000***

The results indicate the significance of the model parameters when comparing the p-value, which is less than the significance level of 0.05.

To test the model and confirm its adequacy after estimation, the Ljung-Box test is applied. The calculated Q statistic (Q_(L-B)) is 1.54865, and when compared to the critical Chi-square value, it is greater than the Q statistic. Furthermore, all autocorrelation and partial autocorrelation coefficients fall within the confidence limits, indicating that the model used is suitable for forecasting. The following figure 5 illustrates the coefficients of the Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) for the residual series of the ARIMA(0,1,2) model, which fall within the confidence interval.

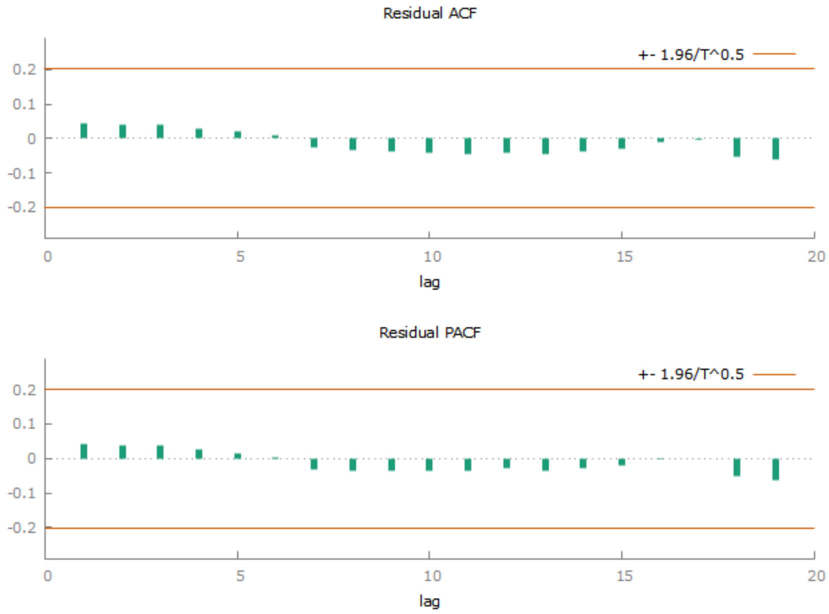


Figure (5): Autocorrelation and Partial Autocorrelation Coefficients for the Residual Series of the Significant ARIMA(0,1,2) Model

17. **Forecasting**

This is the final stage in building Box-Jenkins models after conducting all tests on the time series and selecting the significant model at the 95% confidence level, with z(0.025) = 1.96. Table (6) presents the forecasted values for the time series from (2023:1 to 2025:11) as follows:

Table (6): Forecasted Values for the Time Series Using Box-Jenkins

Obs	prediction	std. error	95% interval
2023:01	1500.26	141.870	(1222.20, 1778.32)
2023:02	1503.56	143.474	(1222.36, 1784.77)
2023:03	1506.30	145.384	(1221.36, 1791.25)
2023:04	1509.04	147.269	(1220.40, 1797.68)
2023:05	1511.78	149.129	(1219.49, 1804.07)
2023:06	1514.52	150.967	(1218.63, 1810.41)
2023:07	1517.26	152.783	(1217.81, 1816.71)
2023:08	1520.00	154.578	(1217.03, 1822.97)
2023:09	1522.74	156.352	(1216.29, 1829.18)
2023:10	1525.48	158.106	(1215.60, 1835.36)
2023:11	1528.22	159.840	(1214.93, 1841.50)
2023:12	1530.95	161.557	(1214.31, 1847.60)
2024:01	1533.69	163.255	(1213.72, 1853.67)
2024:02	1536.43	164.935	(1213.17, 1859.70)
2024:03	1539.17	166.599	(1212.64, 1865.70)
2024:04	1541.91	168.246	(1212.15, 1871.67)
2024:05	1544.65	169.877	(1211.70, 1877.60)
2024:06	1547.39	171.493	(1211.27, 1883.51)

Corresponding author E-mail address: Fatimah.a@coadec.uobaghdad.edu.iq
https://doi.org/10.62933/egnts624

2024:07	1550.13	173.094	(1210.87, 1889.39)
2024:08	1552.87	174.680	(1210.50, 1895.23)
2024:09	1555.61	176.251	(1210.16, 1901.05)
2024:10	1558.35	177.809	(1209.85, 1906.85)
2024:11	1561.09	179.354	(1209.56, 1912.61)
2024:12	1563.82	180.885	(1209.30, 1918.35)
2025:01	1566.56	182.403	(1209.06, 1924.07)
2025:02	1569.30	183.909	(1208.85, 1929.76)
2025:03	1572.04	185.402	(1208.66, 1935.42)
2025:04	1574.78	186.884	(1208.50, 1941.07)
2025:05	1577.52	188.354	(1208.35, 1946.69)
2025:06	1580.26	189.812	(1208.23, 1952.28)
2025:07	1583.00	191.260	(1208.14, 1957.86)
2025:08	1585.74	192.696	(1208.06, 1963.41)
2025:09	1588.48	194.122	(1208.00, 1968.95)
2025:10	1591.22	195.537	(1207.97, 1974.46)
2025:11	1593.95	196.943	(1207.95, 1979.96)

Table (6) shows the forecasted values for the years 2023–2025, and we observe an increase in the price of the Iraqi dinar.

The time series of the forecasted values can be plotted, as illustrated in Figure (6).

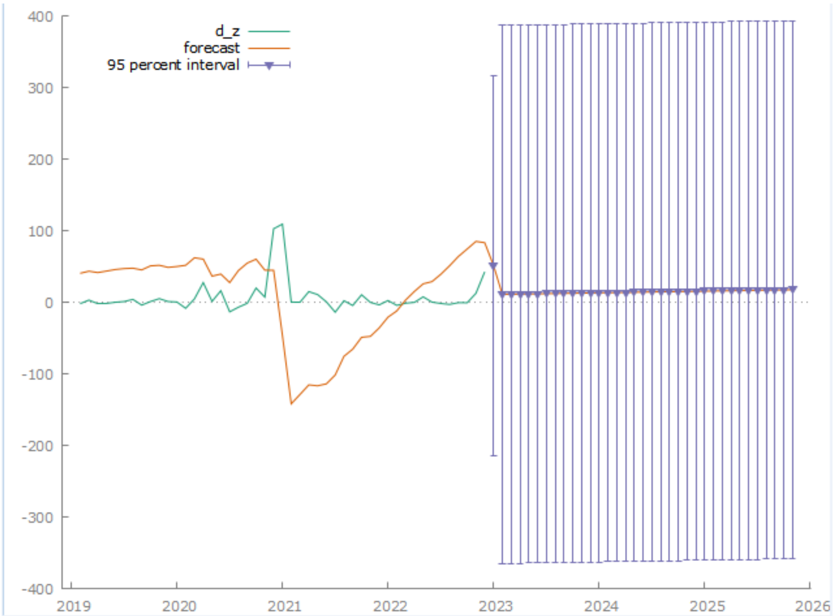


Figure (6): Plot of Forecasted Values

18. Artificial Neural Network

The neural network is used to forecast changes in the price of the Iraqi dinar against the US dollar. The artificial neural network is selected to build the time series model using a dynamic network, utilizing the software package (MATLAB R2015b).

The network is defined, consisting of inputs, hidden layers, and outputs. Data is processed using the NAR neural network, with the data divided into three groups.

The first group, the training set, includes 68 observations, representing 70% of the total observations. The network is trained by calculating the difference between the actual outputs and the outputs predicted by the network, determining the error level, and adjusting the weights to minimize the error.

The second group, the validation set, includes 14 observations, representing 15%, to assess the network's performance and its predictive ability during training.

The third group, the test set, also includes 14 observations, accounting for 15% of the data, for conducting the final test. These three groups determine the architecture of the network, where the price of the Iraqi dinar serves as the independent variable input neuron, consisting of 68 observations. The number of hidden layers is set to one, while the number of neurons is determined through the training of the network.

A nonlinear activation function is used in the hidden layer, while a linear activation function is applied in the output layer, with one output neuron. The testing is conducted to obtain the best architecture using the hidden neurons, as illustrated in Figure (7).

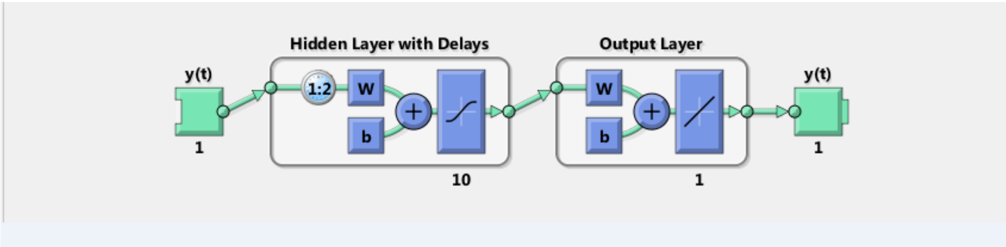


Figure (7): Architecture of the Neural Network

The number of hidden layers is 10, with one output layer. To train the network, the TRANLM backpropagation algorithm is used to minimize the Mean Squared Error (MSE) for improved output efficiency. The correlation coefficient (R) is calculated to evaluate the efficiency of the network training, reflecting the relationship between the targets and the outputs; the smaller the value, the higher the prediction efficiency.

Figure (8) illustrates the decrease in the network errors with increased training, as shown below:

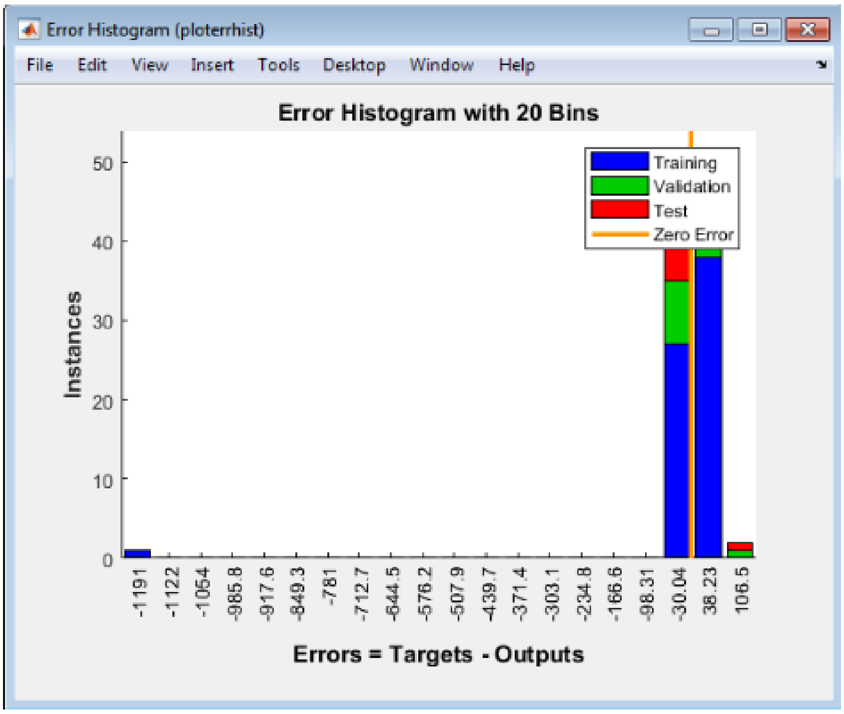


Figure (8): Network Errors After Training

The autocorrelation function lies within the confidence limits, indicating the efficiency of the network training. The value of the correlation coefficient (R = 6.51106e-1) is shown in the following figure:

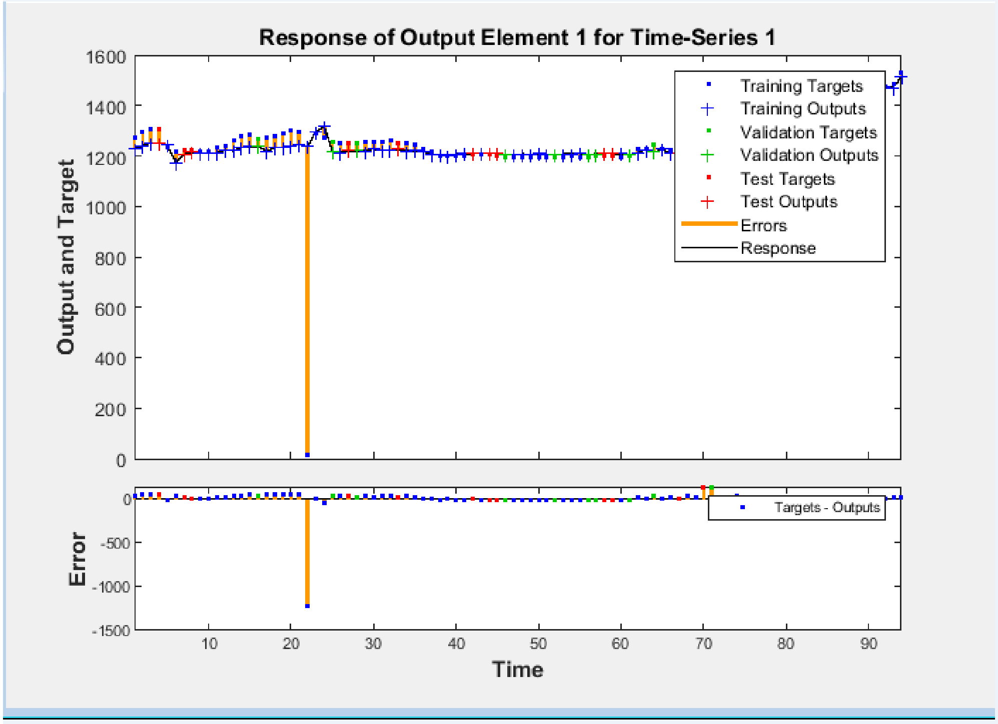


Figure (9): Autocorrelation Function

This figure illustrates the prediction of the exchange rate of the Iraqi dinar against the US dollar using neural networks, with the results presented in Table (7).

observation	Prediction	Observation	prediction	observation	Prediction
1592.09991	2023:01	1623.70582	2024:01	1655.61848	2025:01
1594.72182	2023:02	1626.35361	2024:02	1658.29144	2025:02
1597.34591	2023:03	1629.00353	2024:03	1660.96645	2025:03
1599.97218	2023:04	1631.65556	2024:04	1663.64353	2025:04
1602.60062	2023:05	1634.30972	2024:05	1666.32265	2025:05
1605.23122	2023:06	1636.96598	2024:06	1669.00382	2025:06
1607.86400	2023:07	1639.62435	2024:07	1671.68703	2025:07
1610.49893	2023:08	1642.28482	2024:08	1674.37227	2025:08
1613.13602	2023:09	1644.94738	2024:09	1677.05954	2025:09
1615.77525	2023:10	1647.61204	2024:10	1679.74882	2025:10
1618.41664	2023:11	1650.27877	2024:11	1682.44012	2025:11
1621.06016	2023:12	1652.94759	2024:12	1655.61848	2025:01

We observe an increase in the price for the years 2023 – 2025.

19. Conclusions:

1. Neural networks excel in handling data and deriving appropriate models for analysis, making them an effective tool for improving results. This is achieved by intensively training the network to develop a more accurate and suitable model for predicting various phenomena and patterns in the data.

2. There is an increase in the exchange rate of the Iraqi dinar using neural networks, which explained 70% of the observations (68) for training the network, 15% for validating the network, and 15% for testing the architecture of the network. Furthermore, increasing the training of the network leads to a reduction in the network errors, thereby providing greater accuracy in predictions compared to the Box-Jenkins method.

3. Neural networks become more effective as the size of the data increases, which leads to enhanced training, resulting in the best model and the lowest Mean Squared Error (MSE).

References:

[1] Ahmed, Ammar Shihab (2017), "Using Time Series Models to Forecast Iraq's Oil Exports for the Years (2016-2017)," Baghdad College of Economic Sciences Journal, Issue 52.

[2] Ahmed Mehaoua,Nabil Tabbane,& Sami Tabbane (2004),"Autoregressive,Moving Average and Mixed Autoregressive – Moving Average Processes For Forecasting Qos In Ad Hoc Networks For Real- Time service support",University of Versailles, France.

[3] Al-Awadi, Haider Abdullah (2015), "Using Some Time Series Methods to Forecast Births in Iraq," Higher Diploma in Applied Statistics, College of Administration and Economics, University of Baghdad.

[4] Al-Hadithi, Issam Maulood, 1993, "A Study on Forecasting the Production of Yellow Corn in Iraq," Master's Thesis in Statistics, University of Baghdad, College of Administration and Economics.

[5] Ali Basim Majid (2019), "Using Time Series to Forecast Consumer Price Indexes in Iraq for the Period (2018-2022)," Higher Diploma in Applied Statistics, College of Administration and Economics, University of Baghdad.

[6] Al-Suhail, Rabab Ali Abdul Rahman (2016), "Using Box-Jenkins Models to Forecast the Production Quantity of Some Plant and Animal Products in Baghdad Governorate," Higher Diploma in Applied Statistics, College of Administration and Economics, University of Baghdad.

[7] Bari, Adnan, and Majid Abdul Rahman, 2002, "Statistical Forecasting Methods, Part One," Department of Statistics and Operations Research, Volume One, King Saud University.

[8] Cavanaugh, J.E. , & Neath,A.A.(2019). The Akaike information criterion: Background, derivation, properties, application, interpretation, and refinements. Wiley Interdisciplinary Reviews: Computational Statistics,e1460.

[9] C.H.wu ,RD.soto ,p.p.valko ,A.M.Bubela ,(2000) , "Non- Parametric regression and Neural-Network infill drilling recovery models for carbonater reservoirs",computers Geosciences ,26,pp.975-987 .

[10] Darwish, Marwan (2018), "Using the Box-Jenkins Methodology to Forecast Cash Flows in Palestinian Banks: A Case Study of Palestine Bank," Journal of Al-Quds Open University for Administrative and Economic Research, Volume 3, Issue 9.

[11] Douglas C.Montgmery,chery L.Jennings,& Murat Kulahci(2008),"Introduction to Time Series Analysis and Forecasting',Library of Congress,united states of America.

[12] Da Silva, I.N.,Hernane Spatti,D.,Andrade Flauzino,R.,Liboni, L.H.B.,dos Reis Alves,S.F.(2017).Artificial Neural Network Arechitectures and Training Processes.In:Artificial Neural Networks.Springer,Cham.

[13] Hala Gabreel Mohamed (2017), "Using Time Series Analysis to Forecast Malaria Cases: A Case Study," Master's Thesis,

Department of Applied Statistics, University of Al-Jazira, Sudan.

[14] Ibrahim, Bassam Younis, 2004, "Forecasting Temperature in Khartoum State Using Box-Jenkins Models for Time Series," Sudan University, published in the Sudan Journal of Science and Culture, Volume 5, Issue 2.

[15] Imad Yagoub Hamid , (2011), "Use of Box – Jenkins and Artificial Neural Networks models in Time Series prediction for Sudanese Agricultural sector ", The Third International Conference for Arab Statisticians , Isso,No,501.

[16] Jumaa, Ahlam Ahmed and Hussein, Hala Fadel (2013), "Analysis of the Characteristics of Time Series Models for Oil Sector Data in Iraq for the Period (1958-2008)," Paper presented at the Eighth Scientific Conference, University of Karbala, College of Administration and Economics.

[17]Karlik,B.and Olgac,A.V.,2011.performance analysis of various activation functions in generalized MLP architectures of neural networks.International Journal of Artificial Intelligence and Expert Systems,1(4),pp.111-122.

[18] Majd Namaa,(2023)," Comparison Between Traditional Models and the Use of Artificial Intelligence Models (Neural Networks) to Predict Tobacco Production in Lattakia-Syria",Syrian Journal of Agricultural Research – SJAR 10(1) :, February ,p p. 142 – 155 .

WWW.alepposoft.com " Neural Network" 19 -

[20] Peter J.Brockwell, Richard A. Davis,Introduction to Time Series and Forecasting, Second Edition, springer-verlag,New York, USA,2002,p1.-18.10.4137/BECB.S31601.

[21] Pouliakis,A& Karakitsou,E&others.2016."Artificial Neural Networks as Decision Support Tools in Cytopathology:past present and future ".Biomed Eng Compute Biol.2016.1

[22] Sadia Abdul Kareem Ta'ma (2012), "Using Time Series to Forecast the Number of Malignant Tumor Cases in Anbar Governorate," Al-Anbar University Journal of Economic and Administrative Sciences, Volume 4, Issue 8.

[23] Saleh, Abbas Mahdi (2017), "Forecasting the Number of Live Births and Constructing Life Tables for Diyala Governorate," Higher Diploma in Applied Statistics, Department of Statistics, College of Administration and Economics, University of Baghdad.

[24] Sharma,S.and Athaiya, A.,2017.Activation functions in neural networks.Towards Data Sci,6(12),pp.310-316.

[25] Salioua Rahad and Matar Dhafir (2019)," Comparison of Prediction Performance Between Some Artificial Neural Networks and the Box-Jenkins Methodology", Iraqi Journal of Statistical Sciences, Issue 28, pp 51-76.

[26] Spyros Makridakis,& Michele Hibon(1997),"Arma Mosels and The Box-Jenkins Methodology",Insead,France.

[27] William W.Hsieh and Vancouver ,(2004),"Nonliner Multivariate and Time series Analysis By Neural Network methods "March.18, pp.1-25 . www.ocgy.ubc.ca/~william/Pubs/Rev.Geop.pdf

تطبيق الشبكات العصبية للتنبؤ بسعر الصرف للدينار العراقي مقابل الدولار الامريكي ومقارنتها مع طريقة بوكس جنكينز للسلسلة الزمنية ٢٠١٥-٢٠٢٢

م.د فاطمة عبد الحميد جواد البيرماني
جامعة بغداد/ كلية الادارة والاقتصاد
Fatimah.a@coadec.uobaghdad.edu.iq

المستخلص :
ظهرت في الاونة الاخيرة تغيرات في سعر صرف الدينار العراقي مقابل العملة الاجنبية والتي تعتبر من المؤشرات المالية المهمة التي تؤثر على سوق العمل وكذلك سوق تداول العملات ، ولغرض متابعة التغيرات الحاصلة في سعر الصرف للدينار العراقي مقابل الدولار الامريكي وللوقوف على المراحل القادمة وما يؤول اليه التغير الحاصل في سعر الصرف .
لذا فان الهدف من البحث هو التنبؤ بسعر الصرف للدينار العراقي مقابل الدولار الامريكي من خلال تطبيق منهجية بوكس جنكينز والشبكات العصبية للسنوات القادمة ، وللمفاضلة بين النماذج التنبؤية التقليدية ARIMA ونموذج الشبكة العصبونية والتي اظهرت دقة تنبئها من خلال تقليل متوسط مربعات الخطأ وذلك من خلال تدريب الشبكة واختيار النموذج الملائم وافضل معمارية لتمثيل السلسلة الزمنية .
فقد شملت الدراسة سلسلة زمنية تمثلت باسعار الصرف للدينار العراقي مقابل الدولار الامريكي من شهر كانون الثاني الى كانون الاول للسنوات من ٢٠١٥ – ٢٠٢٢ ،اخذت البيانات من المجموعة الاحصائية السنوية ٢٠٢١/٢٠٢٢ التي يصدرها الجهاز المركزي للاحصاء / العراق ومصدرها البنك المركزي العراقي ، ولايجاد النتائج تم تطبيق برنامج GRETL الاحصائي في تحليل السلسلة الزمنية وبرنامج MatlabR 2019b للتنبؤ عند استخدام الشبكات العصبية.

الكلمات المفتاحية : مجموعة التدريب ، مجموعة التحقق ، بنية الشبكات العصبية ، السلاسل الزمنية .