

## Prediction by using spectral analysis and artificial neural networks methodologies: Comparison Study

Emaan Yousif Abdoon  
Al -Qadisiyah University  
College of Computer Science &  
Information Technology  
Department of Mathematics  
emaan.yousif1991@gmail.com

Mohammed Habib Al- Sharoot  
Al –Qadisiyah University  
College of Administration &  
Economics  
Department of Statistics  
m.alsharood@gmail.com

Recived : 14\8\2017

Revised : //

Accepted : 11\9\2017

### Abstract

The variations in exchange rate, especially the sudden unexpected increases and decreases, have significant impact on the national economy of any country. Iraq is no exception; therefore, the accurate forecasting of exchange rate of Iraqi dinar to US dollar plays an important role in the planning and decision-making processes as well as the maintenance of a stable economy in Iraq. This research aims to compare spectral analysis methodology to artificial neural networks in terms of forecasting the exchange rate of Iraqi dinar to US dollar based on data provided by the Iraqi Central Bank for the period 30/01/2004 and 30/12/2014.

Based on the Mean Square Error (MSE), the Mean Absolute Error (MAE), and the Mean Absolute Percentage Error (MAPE) as criteria to compare the two methodologies, it was concluded that is artificial neural networks better than spectral analysis approach in forecasting.

**Keywords:** Time series analysis, spectral analysis, periodogram, Autoregressive Moving Average models, Artificial neural network, Backpropagation algorithm.

### 1- Spectral analysis methodology

The variations in time series are attributed to a specific phenomenon that is often analyzed in the time domain using autocorrelation function. However, there are many variations in time series that rely on the frequency making the analysis in the time domain not suitable anymore. Therefore, the series is analyzed using the spectral analysis methodology in the frequency domain. The time series spectrum represents the variance distribution of the series as frequency function.

### 2- Harmonic analysis [1, 2]

If the stationary time series  $W_t$  consists of one or more harmonic components, its harmonic model will be as follow:

$$W_t = A_i \cos(\lambda_i t) + B_i \sin(\lambda_i t) + u_t \quad :$$
$$i = 0,1,2, \dots, h \quad (1)$$

Where

$A_i = G_i \cos(\phi_i)$ ,  $B_i = -G_i \sin(\phi_i)$ ,  $G_i$ : Amplitude,  $\phi_i$ : Phase,  $\lambda_i$ : Frequency,  $u_t$ : random errors,  $h$ : number of harmonic components

Equation (1) above is called Fourier series while the frequencies  $\lambda_i$  are often referred to as Fourier frequencies or standard frequencies that are defined in the following formula:

$$\lambda_i = 2\pi \frac{i}{N} \quad (2)$$

Where  $N$  is a sample size

In addition to that,  $A_i$  and  $B_i$  in equation (1) are called Fourier coefficients.

The harmonic model formula relies on the number of observations of the stationary time series as shown below:

a- If the number of observations of the stationary time series is odd, the harmonic model will be:

$$W_t = A_0 + \sum_{j=1}^h [A_j \cos(\lambda_j t) + B_j \sin(\lambda_j t)] + u_t \quad (3)$$

Fourier coefficients can be estimated using the least squares method as shown below:

$$\hat{A}_0 = \frac{1}{N} \sum_{t=1}^N W_t \quad (4)$$

$$\hat{A}_j = \frac{2}{N} \sum_{t=1}^N W_t \cos(\lambda_j t) \quad (5)$$

$$\hat{B}_j = \frac{2}{N} \sum_{t=1}^N W_t \sin(\lambda_j t) \quad (6)$$

Where  $j = 1, 2, \dots, h$

b- If the number of observations of the stationary time series is even, the harmonic model will be:

$$W_t = A_0 + \sum_{j=1}^{h-1} [A_j \cos(\lambda_j t) + B_j \sin(\lambda_j t)] + A_h \cos(\lambda_h t) + u_t \quad (7)$$

And by using the least squares method, Fourier coefficient can be estimated using formula (4), (5), and (6) Where  $j = 1, 2, \dots, h - 1$ . for  $\hat{A}_h$  can be calculated as follow:

$$\hat{A}_h = \frac{2}{N} \sum_{t=1}^N W_t (-1)^t \quad (8)$$

And therefore, an estimation of amplitude of the  $j^{\text{th}}$  harmonic can be obtained using the following formula:

$$\hat{G}_j = \sqrt{\hat{A}_j^2 + \hat{B}_j^2} \quad (9)$$

and an estimation of phase of the  $j^{\text{th}}$  harmonic can be obtained using the following formula:

$$\hat{\Phi}_j = \tan^{-1} \left( \frac{-\hat{B}_j}{\hat{A}_j} \right) \quad (10)$$

### 3- The Periodogram [3, 4, 5]

The periodogram is a tool that is used to analyze time series that consist of sine and cosine waves with different frequencies. This tool is often referred to as  $I(\lambda_j)$  and can be defined in the following formula:

$$I(\lambda_j) = \frac{N}{2} (\hat{A}_j^2 + \hat{B}_j^2) \quad ; \quad j = 1, 2, \dots, h \quad (11)$$

This formula is used if the number of observations of the stationary time series is odd. However, if the number of observations of the stationary time series is even, the following formula can be used:

$$I(\lambda_j) = \frac{N}{2} (\hat{A}_j^2 + \hat{B}_j^2) + I(\lambda_h) \quad ;$$

$$j = 1, 2, \dots, h - 1 \quad (12)$$

Where  $I(\lambda_h) = N \hat{A}_h^2$

### 4- Spectral analysis stages

#### a. Determining the hidden periodicities [6, 7, 1]

In this stage, the hidden sinusoidal components (hidden periodicities) that exist in the time series are revealed through the application of Fisher test as follows :

$$T = \frac{I^{(1)}(\lambda_{(1)})}{\sum_{j=1}^m I(\lambda_j)} \quad (13)$$

Where

$$I^{(1)}(\lambda_{(1)}) = \max \{ I(\lambda_j) \} \quad ;$$

$$j = 1, 2, \dots, m$$

If the value of T statistic is less than the critical value g, we accept the null hypothesis and conclude that the data is random and does not contain hidden sine components.

If the critical value g is not available for  $m > 50$ , g can be calculated at specific significance level alpha as follow:

$$P(T > g_{\alpha}) \approx m(1 - g)^{m-1} \quad (14)$$

Where

$m=h$  when  $I(\lambda_j)$  is defined in formula (11) or  $m=h-1$  when  $I(\lambda_j)$  is defined in formula (12).

**b. Harmonic model selection [8, 1]**

In this stage, the number of significant hidden sine components that consist the harmonic model can be determined by testing the following hypothesis:

$$H_0 : \hat{A}_k = \hat{B}_k = 0$$

Against

$$H_1 : \hat{A}_k \neq 0 \text{ or } \hat{B}_k \neq 0$$

By using  $F_k$  statistic that is defined in the following formula:

$$F_k = \frac{(N-3)I(\lambda_k)}{2 \sum_{\substack{j=1 \\ j \neq k}}^m I(\lambda_j)} : k = 1, 2, \dots, m \quad (15)$$

Where  $F_k$  statistic follows the  $F$  distribution with 2 and  $(N - 3)$  degrees of freedom.

If  $F_k < F(2, N-3)$ , we accept the null hypothesis and conclude that the hidden sine component is not significant.

**c. Testing the suitability of the harmonic model [9,10]**

In this stage, randomization of the model residuals is tested through the two following tests:

1- **Ljung-Box test [9]**

2- **Cumulative periodogram test**

[10]

The statistic of cumulative periodogram test is defined in the following formula:

$$C(\lambda_l) = \frac{\sum_{j=1}^l I(\lambda_j)}{\sum_{j=1}^m I(\lambda_j)} : l = 1, 2, \dots, m \quad (16)$$

After calculating  $C(\lambda_l)$ , Kolmogorov smirnov test can be applied where Kolmogorov smirnov limits can be determined as follow:

**a)** The theoretical line: this line is drawn from  $(0, 0)$  to  $(\pi, 1)$

**b)** The upper limit: this line can be drawn using the following formula:

$$\text{Upper limit} = \text{theoretical line} + \frac{K_{\alpha}}{\sqrt{m}} \quad (17)$$

**c)** The lower limit: this line can be drawn using the following formula:

$$\text{Lower limit} = \text{theoretical line} - \frac{K_{\alpha}}{\sqrt{m}} \quad (18)$$

Where  $K_{\alpha}$  can be obtained from Table (8-5) in [8] at alpha significance level.

If the  $C(\lambda_l)$  value falls between the lower and upper limits, we can conclude that the series does not contain hidden sinusoidal components.

**3-Artificial neural networks**

[11, 12, 13]

Artificial neural network is a computational technique that simulates the way human brain uses to perform a specific task. It consists of several processing units defined as neurons or nodes that are organized in three levels (layers) including input level, hidden level that consists of one or more hidden layer, and the output level. the nodes of each level are associated with the next level through connection force called weights, which work on saving the acquired knowledge from the training of network.

The nodes at the input level are called the input nodes and so on for the other levels. In addition, there is a node called the bias  $b$  that has a positive value of one and has the same constant role in the regression model.

Each node of the hidden and output levels is provided with an activation function, linear or nonlinear, that works on processing the input signal and preventing the output of the processing node from reaching high value, which may stop the network leading to a failure of the training process.

Neural networks are divided to two types including single layer and multiple layer networks. The single layer network does not have the hidden level and contains one layer of weights that connects the input level to the output level. when applying the input signal  $v_j$  in this type of network, we can obtain the output signal  $z_i$  through the following formula:

$$z_i = g\left(\sum_{j=1}^R x_j v_j + b_i\right) : i = 1, 2, \dots, n \quad (19)$$

Where

$x_j$ : weights,  $g$ : activation function,  $R$ : number of input nodes,  $n$ : number of output nodes

As for multi-layer networks, it has the ability to solve more complex problems because it contains the hidden level with one hidden layer or more. The output signal is obtained in this type of networks as follows:

$$\mathbf{Z}^m = \mathbf{g}^m(\mathbf{X}^m \mathbf{Z}^{m-1} + \mathbf{b}^m) : m = 1, 2, \dots, M \quad (20)$$

Where:

M: number of layers

$\mathbf{Z}^{m-1}$ : output vector of the layer m-1

$\mathbf{g}^m$ : activation function of the layer m

$\mathbf{b}^m$ : bias vector of the layer m

$\mathbf{X}^m$ : weights matrix of the layer m

$\mathbf{Z}^0 = \mathbf{V}$ , and  $\mathbf{V}$  is the input vector.

### (3-1) Backpropagation algorithm [14, 15]

The backpropagation algorithm is a generalization of the least mean squares algorithm where it is used to train multiple layers networks. It is often referred to as BP and is considered one of the most used algorithms among the supervised learning algorithms. It consists of three stages:

#### A. Feed forward propagation stage [16, 11]

In this stage, inputs  $\mathbf{Z}^0 = \mathbf{V}$  are applied on the network and random initial weights  $\mathbf{X}^1$  are generated in small values. In addition, this stage includes determining the learning rate  $\eta$  and the momentum  $\gamma$  by small value that falls within (0, 1) and (1, 0] respectively. Data are processed starting from the input layer to the output layer throughout the hidden layers. The outputs of the network are obtained using the following formula:

$$\begin{aligned} z_j^m &= g^m(s_j^m) \quad m \\ &= 1, 2, \dots, M \quad (21) \\ &= g_j^m(\sum_{h=1}^{n^{m-1}} x_{j,h}^m z_h^{m-1} + b_j^m) \quad (22) \end{aligned}$$

Which can be written as a matrix

$$\begin{aligned} \mathbf{Z}^m &= \mathbf{g}^m(\mathbf{s}^m) \quad (23) \\ &= \mathbf{g}^m(\mathbf{X}^m \mathbf{Z}^{m-1} + \mathbf{b}^m) \quad (24) \end{aligned}$$

#### B. Backward Propagation Stage [17, 11]

In this stage, the sensitivities  $\delta$  is calculated from the last layer, which represents the output layer, to the first layer throughout the hidden layers.

Sensitivity  $\delta^M$  in the last layer M can be calculated using the following formula:

$$\delta^M = -2\hat{\mathbf{G}}^M(\mathbf{s}^M)(\mathbf{d} - \mathbf{z}) \quad (25)$$

Where

$\mathbf{d}$ : desired output

$\mathbf{z}$ : calculated output by network

$$\hat{\mathbf{G}}^M(\mathbf{s}^M) = \begin{bmatrix} \dot{g}_1^M(s_1^M) & 0 & \dots & 0 \\ 0 & \dot{g}_2^M(s_2^M) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \dot{g}_{n^M}^M(s_{n^M}^M) \end{bmatrix}$$

and

$$\dot{g}_j^M(s_j^M) = \frac{\partial}{\partial s_j^M} g_j^M(s_j^M) \quad j = 1, 2, \dots, n^M$$

While sensitivity  $\delta^m$  in the hidden layer m, where  $m= 1, 2, \dots, M-1$ , can be calculated using the following formula:

$$\delta^m = \hat{\mathbf{G}}^m(\mathbf{s}^m)(\mathbf{X}^{m+1})^T \delta^{m+1} \quad (26)$$

Where

$\mathbf{X}^{m+1}$ : weights matrix of the layer m+1

$$\hat{\mathbf{G}}^m(\mathbf{s}^m) = \begin{bmatrix} \dot{g}_1^m(s_1^m) & 0 & \dots & 0 \\ 0 & \dot{g}_2^m(s_2^m) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \dot{g}_{n^m}^m(s_{n^m}^m) \end{bmatrix}$$

and

$$\dot{g}_j^m(s_j^m) = \frac{\partial}{\partial s_j^m} g_j^m(s_j^m)$$

Where  $j=1, 2, \dots, n^m$

### C. Updating Weights Stage [17, 16]

After passing the feed forward propagation and the Backward Propagation stages, the stage of updating the weights and biases begins using the following formulas:

$$\mathbf{X}^m(k) = \mathbf{X}^m(k - 1) + \Delta\mathbf{X}^m(k - 1) \quad (27)$$

$$\mathbf{b}^m(k) = \mathbf{b}^m(k - 1) + \Delta\mathbf{b}^m(k - 1) \quad (28)$$

Where

3- In case of not using the momentum:

$$\Delta\mathbf{X}^m(k - 1) = -\eta\delta^m(\mathbf{z}^{m-1})^T$$

$$\Delta\mathbf{b}^m(k - 1) = -\eta\delta^m$$

4- In case of using the momentum:

$$\Delta\mathbf{X}^m(k - 1) = \gamma\Delta\mathbf{X}^m(k - 2) - (1 - \gamma)\eta\delta^m(\mathbf{z}^{m-1})^T$$

$$\Delta\mathbf{b}^m(k - 1) = \gamma\Delta\mathbf{b}^m(k - 2) - (1 - \gamma)\eta\delta^m$$

### 4- Data

The data used in this study is a time series of exchange rate of Iraqi dinar to the US dollar and was provided by the Iraqi Central Bank. The data consist of 132 monthly observations from 30/01/2004 to 30/12/2014 as show in table (1) below.

month	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014
1	1467	1453	1483	1323	1224	1179	1185	1185	1206	1226	1222
2	1409	1459	1480	1299	1225	1178	1185	1185	1236	1231	1222
3	1423	1469	1480	1290	1222	1178	1185	1185	1240	1255	1222
4	1443	1474	1481	1284	1216	1179	1185	1187	1263	1267	1218
5	1462	1473	1485	1275	1212	1187	1185	1196	1250	1270	1222
6	1460	1468	1485	1269	1205	1180	1185	1197	1241	1237	1213
7	1463	1476	1486	1261	1202	1184	1185	1197	1253	1218	1214
8	1463	1480	1488	1253	1196	1184	1185	1199	1248	1209	1213
9	1463	1481	1488	1249	1188	1183	1185	1200	1228	1211	1204
10	1463	1475	1486	1245	1185	1183	1185	1200	1200	1220	1207
11	1463	1477	1467	1240	1183	1183	1188	1200	1207	1218	1200
12	1462	1479	1394	1216	1180	1185	1195	1218	1222	1222	1205

### 5- spectral analysis methodology application

#### The first stage:

The time series of the exchange rate data was plotted as shown in figure (1) below

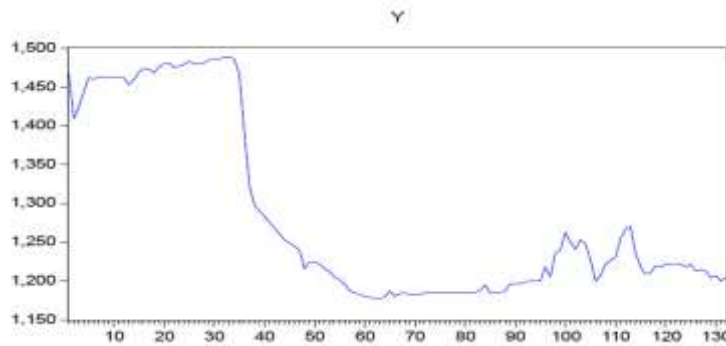


Figure 1: Exchange rate of Iraqi dinar to the US dollar

By looking at figure (1), we notice that the data  $Y_t$  does not fluctuate around constant level, and it takes a decreasing trend which indicates that the time series is not mean stationary and not variance stationary.

To check the accuracy of results about the stationary of the time series in the mean, autocorrelation function and partial autocorrelation function were plotted for the raw data as shown in figure (2) below.

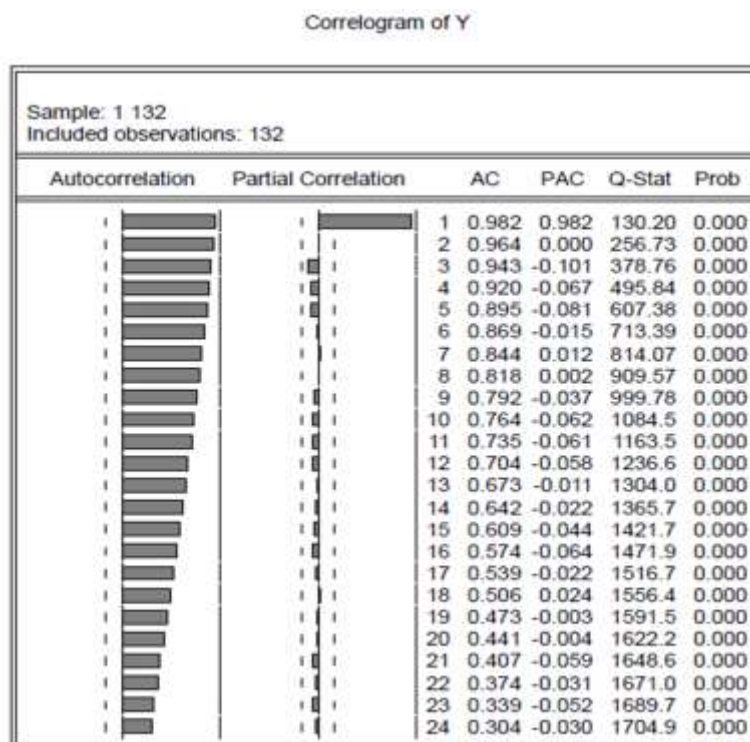


Figure 2: Correlogram of raw data  $Y_t$

By looking at figure (2), we notice that the autocorrelation function is slowly decreasing toward the zero and does not cut after the first and the second lag, which indicates that the raw data is not mean stationary.

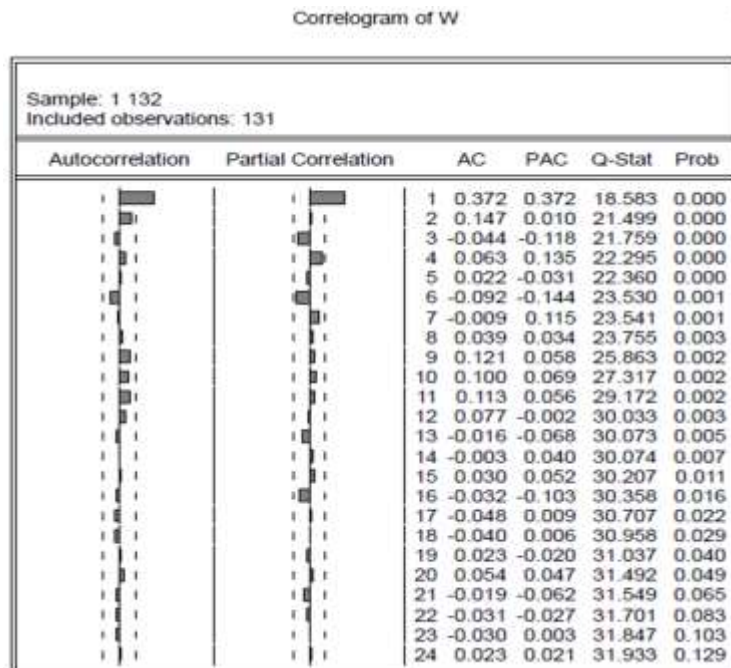
To increase accuracy in the results about the stationary of the time series in the variance, ADF, P.P, and KPSS tests were applied as shown in table (2) below.

Test	Model	Test Statistic	Critical Values	p-Value
ADF	without constant	-0.837279	1.943304-	0.3514
	With constant	-0.979861	2.883756-	0.7591
	With constant and time trend	-1.307575	3.444756-	0.8817
P.P	without constant	-1.397307	1.943304-	0.1504
	With constant	-1.46627	2.883756-	0.5477
	With constant and time trend	-1.200101	3.444756-	0.9059
KPSS	With constant	0.937165	0.643000	
	With constant and time trend	0.277682	0.146000	

By checking the P-value of each model of the estimated models for both tests ADF and P.P at 0.05 significance level (alpha), we accept the null hypothesis and concluded that the time series has unit root meaning that it is not variance stationary. In addition, through ADF test, we concluded that the time series needs to take differences.

By comparing the calculated value of KPSS test statistic to the critical value, we accept the alternative hypothesis, which means that the time series is not variance stationary.

From the results of plots and tests, we conclude that the time series is not mean stationary and not variance stationary. Therefore, the log transformation was applied then the first difference was taken to accomplish stationary in the series in the variance and the mean respectively as shown in figure (3) which shows that the autocorrelation function of the transformed data is cut after the first lag, which indicates that the time series is mean stationary.



**Figure 3: Correlogram of transformed data  $W_t$**

It becomes clear from the results of ADF, P.P, and KPSS tests of the transformed data, shown in table (3) below, that the time series is

variance stationary. In addition, the results of ADF indicate that the series does not need to take anymore differences, which indicates that the series is mean stationary.

**Table 3: The tests results of transformed data  $W_t$**

Test	Model	Test Statistic	Critical Values	p-Value
ADF	without constant	8.18208-	1.943304-	0.0000
	With constant	8.197312-	2.883756-	0.0000
	With constant and time trend	8.149873-	3.444756-	0.0000
P.P	without constant	8.344295-	1.943304-	0.0000
	With constant	-8.382012	2.883756-	0.0000
	With constant and time trend	8.343332-	3.444756-	0.0000
KPSS	With constant	0.214289	0.643000	
	With constant and time trend	0.098436	0.146000	

After checking the stationarity of the series, 65 frequencies were generated using formula (2). In addition, the

amplitude, the phase, and the periodogram were calculated as shown in table (4).

**4 : Periodogram analysis of  $W_t$  Table**

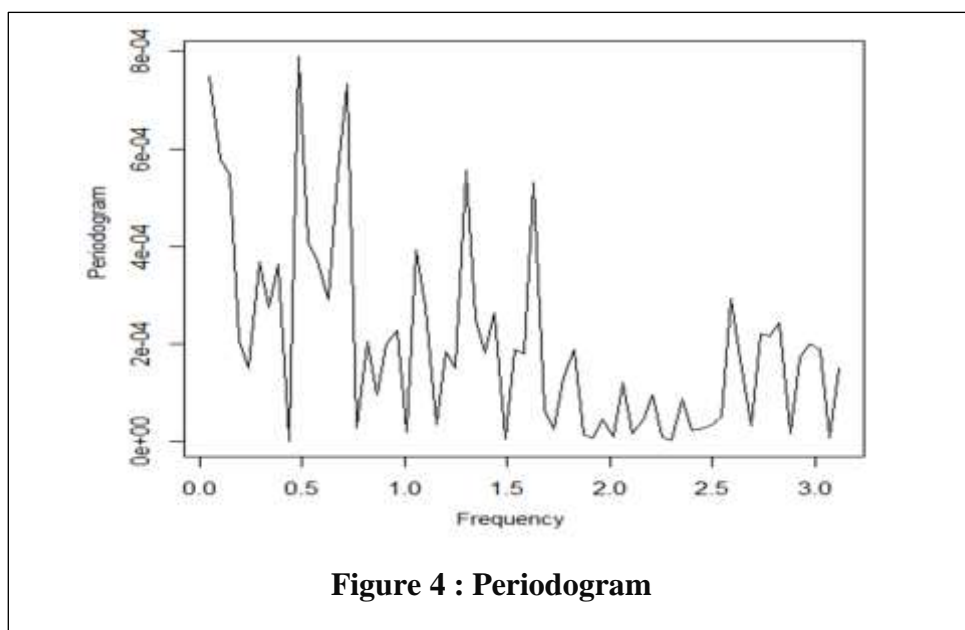
j	Frequency	period	$\hat{A}_j$	$\hat{B}_j$	Amplitude	Phase	periodogram
1	0.04796325	20.8492975	7.079433e-04	-3.307630e-03	0.0033825436	1.35994424	7.494249e-04
2	0.09592649	10.4246488	1.431931e-03	2.602334e-03	0.0029702810	-1.06776213	5.778783e-04
3	0.14388974	6.9497658	-1.660042e-03	2.361398e-03	0.0028865100	0.95806557	5.457421e-04
4	0.19185299	5.2123244	-1.057219e-03	-1.432414e-03	0.0017803154	-0.93497568	2.076037e-04
5	0.23981623	4.1698595	6.907379e-04	-1.340426e-03	0.0015079326	1.09497412	1.489379e-04
6	0.28777948	3.4748829	8.760026e-04	2.201576e-03	0.0023694553	-1.19210338	3.677379e-04
7	0.33574273	2.9784711	-1.694825e-03	1.152857e-03	0.0020497585	0.59732830	2.751989e-04
8	0.38370597	2.6061622	-7.703295e-04	-2.223622e-03	0.0023532748	-1.23730538	3.627326e-04
9	0.43166922	2.3165886	1.112947e-04	6.335595e-05	0.0001280644	-0.51751207	1.074232e-06
10	0.47963247	2.0849298	-7.756589e-04	3.382998e-03	0.0034707814	1.34541046	7.890342e-04
11	0.52759571	1.8953907	-1.071563e-03	2.255256e-03	0.0024968837	1.12723366	4.083551e-04
12	0.57555896	1.7374415	1.226869e-03	-2.036024e-03	0.0023770990	1.02848120	3.701143e-04
13	0.62352221	1.6037921	1.523896e-03	-1.452157e-03	0.0021049983	0.76129763	2.902317e-04
14	0.67148545	1.4892355	-2.849056e-03	5.581426e-04	0.0029032121	0.19345444	5.520760e-04
15	0.71944870	1.3899532	-3.324620e-03	-3.815690e-04	0.0033464453	-0.11427072	7.335146e-04
16	0.76741195	1.3030811	-6.058610e-05	-6.516940e-04	0.0006545042	-1.47809570	2.805862e-05
17	0.81537519	1.2264293	3.992304e-04	1.720010e-03	0.0017657352	-1.34272555	2.042173e-04
18	0.86333844	1.1582943	-5.347976e-04	1.077578e-03	0.0012029894	1.11011643	9.479051e-05
19	0.91130169	1.0973314	-3.786098e-04	-1.707657e-03	0.0017491249	-1.35261267	2.003932e-04
20	0.95926493	1.0424649	3.652009e-04	-1.827602e-03	0.0018637327	1.37356885	2.275142e-04
21	1.00722818	0.9928237	-2.190613e-04	5.006453e-04	0.0005464737	1.15833728	1.956050e-05
22	1.05519143	0.9476953	-2.440567e-03	-1.415041e-04	0.0024446655	-0.05791519	3.914535e-04
23	1.10315467	0.9064912	-1.399170e-03	-1.496164e-03	0.0020484589	-0.81888554	2.748501e-04
24	1.15111792	0.8687207	6.120694e-04	-3.864795e-04	0.0007238752	0.56321041	3.432169e-05
25	1.19908117	0.8339719	-1.507334e-03	7.399063e-04	0.0016791421	0.45631752	1.846784e-04
26	1.24704441	0.8018961	-1.252173e-03	-8.397729e-04	0.0015076987	-0.59075700	1.488917e-04
27	1.29500766	0.7721962	2.248463e-03	-1.856158e-03	0.0029156316	0.69011101	5.568095e-04
28	1.34297091	0.7446178	1.040600e-03	-1.651796e-03	0.0019522497	1.00862326	2.496388e-04
29	1.39093415	0.7189413	-1.153633e-03	-1.196109e-03	0.0016617904	-0.80347318	1.808813e-04
30	1.43889740	0.6949766	-2.000057e-03	-1.215087e-04	0.0020037445	-0.06067807	2.629820e-04
31	1.48686065	0.6725580	-1.684346e-04	2.189918e-04	0.0002762746	0.91516007	4.999460e-06



32	1.53482389	0.6515405	1.661626e-03	3.536203e-04	0.0016988374	-0.20968754	1.890362e-04
33	1.58278714	0.6317969	1.703586e-05	-1.653322e-03	0.0016534096	1.56049267	1.790615e-04
34	1.63075039	0.6132146	-1.036664e-03	-2.648826e-03	0.0028444593	-1.19775397	5.299571e-04
35	1.67871363	0.5956942	8.291315e-04	-5.393994e-04	0.0009891465	0.57676847	6.408590e-05
36	1.72667688	0.5791472	2.320018e-04	5.888433e-04	0.0006328991	-1.19547665	2.623676e-05
37	1.77464012	0.5634945	3.076818e-04	-1.367872e-03	0.0014020490	1.34954405	1.287561e-04
38	1.82260337	0.5486657	7.966094e-04	-1.493402e-03	0.0016925823	1.08077200	1.876467e-04
39	1.87056662	0.5345974	3.951427e-04	-2.576959e-04	0.0004717467	0.57789160	1.457669e-05
40	1.91852986	0.5212324	-3.668132e-05	-3.362969e-04	0.0003382915	-1.46215162	7.495895e-06
41	1.96649311	0.5085195	2.758703e-04	-7.867621e-04	0.0008337260	1.23355143	4.552899e-05
42	2.01445636	0.4964118	-3.849164e-05	-3.839125e-04	0.0003858373	-1.47086878	9.751014e-06
43	2.06241960	0.4848674	1.228237e-03	-5.733382e-04	0.0013554644	0.43673462	1.203421e-04
44	2.11038285	0.4738477	-4.790360e-04	-4.077087e-05	0.0004807679	-0.08490562	1.513953e-05
45	2.15834610	0.4633177	-5.107497e-04	-6.366272e-04	0.0008161859	-0.89467013	4.363344e-05
46	2.20630934	0.4532456	1.210339e-03	-7.011997e-05	0.0012123689	0.05786945	9.627441e-05
47	2.25427259	0.4436021	2.256535e-04	-3.054342e-04	0.0003797493	0.93450455	9.445724e-06
48	2.30223584	0.4343604	-1.495696e-04	-1.186235e-04	0.0001908994	-0.67051872	2.386990e-06
49	2.35019908	0.4254959	1.096871e-03	-3.403344e-04	0.0011484568	0.30085888	8.639142e-05
50	2.39816233	0.4169860	5.478249e-04	-2.295263e-04	0.0005939650	0.39675844	2.310803e-05
51	2.44612558	0.4088098	3.180891e-04	-5.428895e-04	0.0006292135	1.04079514	2.593208e-05
52	2.49408882	0.4009480	-3.528393e-04	6.147261e-04	0.0007087904	1.04973040	3.290614e-05
53	2.54205207	0.3933830	3.334777e-04	8.110359e-04	0.0008769188	-1.18069359	5.036862e-05
54	2.59001532	0.3860981	2.075319e-03	3.923407e-04	0.0021120793	-0.18684569	2.921876e-04
55	2.63797856	0.3790781	1.148417e-03	-1.121643e-03	0.0016052868	0.77360450	1.687899e-04
56	2.68594181	0.3723089	-5.626839e-04	-4.030420e-04	0.0006921387	-0.62157207	3.137817e-05
57	2.73390506	0.3657771	1.820756e-03	-1.943188e-04	0.0018310956	0.10632180	2.196157e-04
58	2.78186830	0.3594706	1.727725e-03	-5.538987e-04	0.0018143426	0.31024183	2.156155e-04
59	2.82983155	0.3533779	1.556185e-04	-1.921047e-03	0.0019273403	1.48996571	2.433090e-04
60	2.87779480	0.3474883	3.654281e-04	3.333941e-04	0.0004946608	-0.73959027	1.602715e-05
61	2.92575804	0.3417918	1.150607e-03	1.125550e-03	0.0016095833	-0.77439011	1.696947e-04
62	2.97372129	0.3362790	1.726059e-03	-2.961115e-04	0.0017512740	0.16989970	2.008859e-04
63	3.02168454	0.3309412	1.081432e-03	-1.309441e-03	0.0016982730	0.88047699	1.889106e-04
64	3.06964778	0.3257703	-3.246632e-04	8.699341e-05	0.0003361161	0.26179988	7.399799e-06
65	3.11761103	0.3207584	1.507766e-03	-1.383658e-04	0.0015141012	0.09151244	1.501589e-04

By examining Table (4) and the periodogram plot shown in Figure (4) below, we can conclude that there are three distinctive peaks

at  $\lambda_1, \lambda_{10}, \lambda_{15}$  frequencies, which indicate the existence of hidden sinusoidal components in the data.



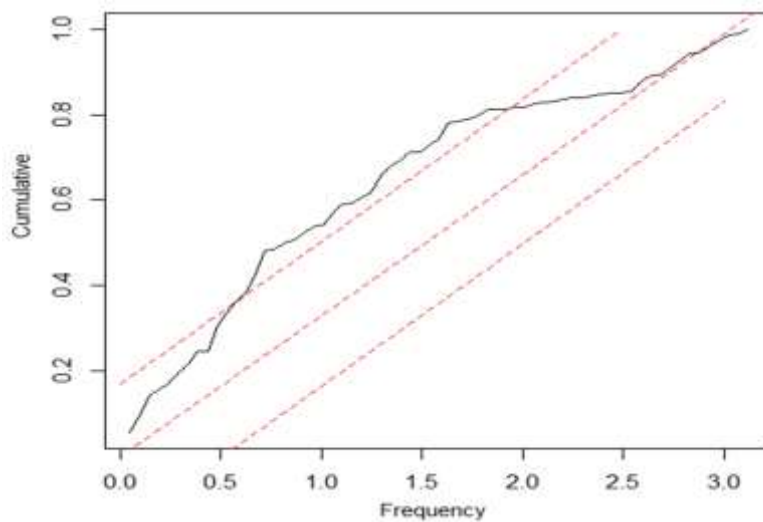
**Figure 4 : Periodogram**

To increase the results accuracy, the cumulative periodogram test was performed on the stationary data where  $C(\lambda_t)$  was calculated using formula (16) as shown in Table (5)

below. In addition, the upper and lower limits were drawn using (17) and (18) formulas respectively as shown in figure (5).

**Table 5 : Cumulative periodogram of  $W_t$**

I	Frequency	Cumulative Periodogram	I	Frequency	Cumulative Periodogram
1	0.04796325	0.05662134	34	1.63075039	0.78135100
2	0.09592649	0.10028182	35	1.67871363	0.78619289
3	0.14388974	0.14151430	36	1.72667688	0.78817516
4	0.19185299	0.15719940	37	1.77464012	0.79790307
5	0.23981623	0.16845211	38	1.82260337	0.81208035
6	0.28777948	0.19623583	39	1.87056662	0.81318167
7	0.33574273	0.21702796	40	1.91852986	0.81374800
8	0.38370597	0.24443351	41	1.96649311	0.81718786
9	0.43166922	0.24451467	42	2.01445636	0.81792458
10	0.47963247	0.30412862	43	2.06241960	0.82701679
11	0.52759571	0.33498110	44	2.11038285	0.82816063
12	0.57555896	0.36294436	45	2.15834610	0.83145727
13	0.62352221	0.38487226	46	2.20630934	0.83873110
14	0.67148545	0.42658329	47	2.25427259	0.83944475
15	0.71944870	0.48200256	48	2.30223584	0.83962509
16	0.76741195	0.48412247	49	2.35019908	0.84615223
17	0.81537519	0.49955171	50	2.39816233	0.84789811
18	0.86333844	0.50671342	51	2.44612558	0.84985736
19	0.91130169	0.52185374	52	2.49408882	0.85234352
20	0.95926493	0.53904314	53	2.54205207	0.85614902
21	1.00722818	0.54052100	54	2.59001532	0.87822469
22	1.05519143	0.57009651	55	2.63797856	0.89097729
23	1.10315467	0.59086227	56	2.68594181	0.89334800
24	1.15111792	0.59345538	57	2.73390506	0.90994064
25	1.19908117	0.60740840	58	2.78186830	0.92623105
26	1.24704441	0.61865762	59	2.82983155	0.94461378
27	1.29500766	0.66072628	60	2.87779480	0.94582468
28	1.34297091	0.67958725	61	2.92575804	0.95864564
29	1.39093415	0.69325339	62	2.97372129	0.97382318
30	1.43889740	0.71312249	63	3.02168454	0.98809596
31	1.48686065	0.71350021	64	3.06964778	0.98865504
32	1.53482389	0.72778247	65	3.11761103	1.00000000
33	1.58278714	0.74131112			



**Figure 5 : Cumulative periodogram of  $W_t$**

By examining figure (5), we notice that  $C(\lambda_j)$  values fall out of the upper limit, which indicates the existence of hidden sine components in the stationary data. This result emphasizes that the data is not completely random as shown in figure (4).

**The second stage**

Based on the values of  $I(\lambda_j)$  periodogram shown in table (4),  $F_k$  statistic, defined in formula (15), was applied as shown in table (6) below. periodogram shown in table (4),  $F_k$

**Table 6 :  $F_k$  Statistic application results**

k	Frequency	F-Statistic	F(2,128)	K	Frequency	F-Statistic	F(2,128)
1	0.04796325	3.841263532	3.00	34	1.63075039	2.669436583	3.00
2	0.09592649	2.921839197	3.00	35	1.67871363	0.311388428	3.00
3	0.14388974	2.752365939	3.00	36	1.72667688	0.127117126	3.00
4	0.19185299	1.019842640	3.00	37	1.77464012	0.628702526	3.00
5	0.23981623	0.728369720	3.00	38	1.82260337	0.920394664	3.00
6	0.28777948	1.828973830	3.00	39	1.87056662	0.070561791	3.00
7	0.33574273	1.358951182	3.00	40	1.91852986	0.036266160	3.00
8	0.38370597	1.803378278	3.00	41	1.96649311	0.220910593	3.00
9	0.43166922	0.005194758	3.00	42	2.01445636	0.047184775	3.00
10	0.47963247	4.057155804	3.00	43	2.06241960	0.587240968	3.00
11	0.52759571	2.037417789	3.00	44	2.11038285	0.073289431	3.00
12	0.57555896	1.841133116	3.00	45	2.15834610	0.211682804	3.00
13	0.62352221	1.434848268	3.00	46	2.20630934	0.468935871	3.00
14	0.67148545	2.785700301	3.00	47	2.25427259	0.045706437	3.00
15	0.71944870	3.754928909	3.00	48	2.30223584	0.011544125	3.00
16	0.76741195	0.135962744	3.00	49	2.35019908	0.420481252	3.00
17	0.81537519	1.002945982	3.00	50	2.39816233	0.111931911	3.00
18	0.86333844	0.461655906	3.00	51	2.44612558	0.125638041	3.00
19	0.91130169	0.983876631	3.00	52	2.49408882	0.159510775	3.00
20	0.95926493	1.119362529	3.00	53	2.54205207	0.244482597	3.00
21	1.00722818	0.094722730	3.00	54	2.59001532	1.444736099	3.00
22	1.05519143	1.950520308	3.00	55	2.63797856	0.826708920	3.00
23	1.10315467	1.357191961	3.00	56	2.68594181	0.152086419	3.00
24	1.15111792	0.166390435	3.00	57	2.73390506	1.079846238	3.00
25	1.19908117	0.905629647	3.00	58	2.78186830	1.059851496	3.00
26	1.24704441	0.728141247	3.00	59	2.82983155	1.198527339	3.00
27	1.29500766	2.810633769	3.00	60	2.87779480	0.077591585	3.00
28	1.34297091	1.230307102	3.00	61	2.92575804	0.831197742	3.00
29	1.39093415	0.886751384	3.00	62	2.97372129	0.986333087	3.00
30	1.43889740	1.297400132	3.00	63	3.02168454	0.926683895	3.00
31	1.48686065	0.024183504	3.00	64	3.06964778	0.035800975	3.00
32	1.53482389	0.927308872	3.00	65	3.11761103	0.734409593	3.00
33	1.58278714	0.877707418	3.00				

By examining table (6), we conclude the existence of three significant hidden sinusoidal components at  $\lambda_1, \lambda_{10}, \lambda_{15}$  frequencies because the calculated F value at these frequencies is higher than the critical value  $F(2, 128)$ , which indicate to the unacceptability of the null hypothesis at this frequencies. Therefore, the harmonic model of the data can be written in the following formula:

$$\hat{W}_t = \sum_{l=1,10,15} [\hat{A}_l \cos(\lambda_l t) + \hat{B}_l \sin(\lambda_l t)] + u_t$$

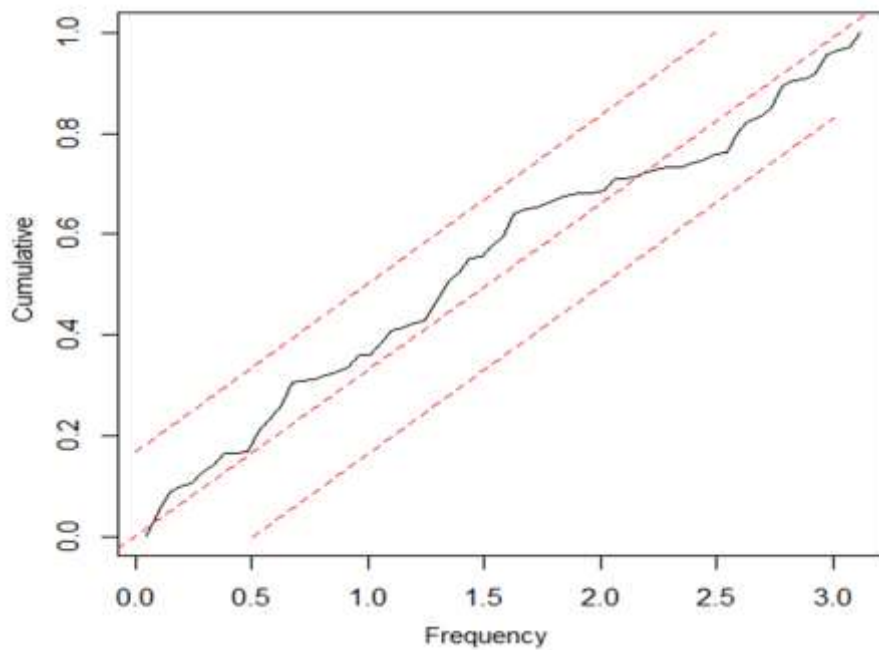
Where  $t = 1, 2, \dots, N$

### The third stage

In this stage, residuals of the selected harmonic model from the second stage were tested through the application of Ljung- Box Test and the cumulative periodogram test as shown in table (7) and figure (6) below.

**Table 7 : Ljung- Box Test**

$(Q_{LB})$	$\chi^2(df = 18, alpha = 0.05)$	p-value
25.244	28.869	0.3926



**Figure 6 : Cumulative periodogram of residuals**

By examining table (7), we conclude that the residuals are random and represent the white noise. By examining figure (6), we notice that the cumulative periodogram values fall between the lower and upper limits, which indicates that the residuals do not contain hidden sinusoidal components.

## 6- Artificial neural network methodology application

The main step in designing the neural network model of a specific time series is determining the number of input variables. By applying Box-Jenkins methodology to the series of exchange rate, we find the best model to represent the data is ARIMA (1,1,0) . This mean that the input variables include only the variable  $Y_{t-1}$  . therefore, the number of input nodes is equal to one. Therefore, the number of input nodes equal to one. Because the goal

here is to predict one-step-ahead, one output node was set in the output layer, which include

one variable  $Y_t$ . In addition, one hidden layer was determined for the hidden level. By choosing backpropagation algorithm to train the network, a 0.5 learning speed and 0.9 momentum were selected, and we include 100% of the data for the training due to the small sample size. Because there is no constant rule to select the activation functions in both the hidden and output layers, 5 models including ANN(1), ANN(2), ANN(3), ANN(4), and ANN(5) were built with different activation functions as shown in table (\*). Based on the activation function in the output layer of each model, the processing formula was determined. The number of hidden nodes in each models were also determined based on the try and error approach and the following formulas:

**Table 8 : Artificial neural networks models**

First Model ANN( 1 )	Activation function for hidden layer				Hyperbolic tangent			
	Activation function for output layer				sigmoid			
	Data preprocessing formula				normalized			
	Number of Hidden Nodes	MSE	MAE	MAPE	Number of Hidden Nodes	MSE	MAE	MAPE
	1	355.3003	13.8103	1.0669	6	244.4384	10.8591	0.8442
	2	239.2133	10.7915	0.8421	7	240.6842	10.4926	0.8174
3	314.3509	12.3766	0.96	8	320.4959	12.2783	0.9625	
4	363.1231	14.0481	1.0874	9	268.5468	11.936	0.9279	
5	299.6979	12.6116	0.9751	10	276.1918	11.8453	0.9259	
Second Model ANN( 2 )	Activation function for hidden layer				sigmoid			
	Activation function for output layer				sigmoid			
	Data preprocessing formula				normalized			
	Number of Hidden Nodes	MSE	MAE	MAPE	Number of Hidden Nodes	MSE	MAE	MAPE
	1	308.7995	12.6732	0.98	6	400.6688	13.8814	1.0779
	2	399.0621	14.5312	1.0978	7	366.3151	14.5746	1.1027
3	337.6484	12.7363	1.0029	8	272.8888	9.5703	0.7261	
4	246.7104	10.5956	0.8088	9	297.0384	12.5679	0.9652	
5	256.438	11.5061	0.8838	10	272.4871	11.658	0.9101	
Third Model ANN( 3 )	Activation function for hidden layer				Hyperbolic tangent			
	Activation function for output layer				Identity			
	Data preprocessing formula				normalized			
	Number of Hidden Nodes	MSE	MAE	MAPE	Number of Hidden Nodes	MSE	MAE	MAPE
	1	177.339	7.6599	0.5865	6	184.5194	8.0842	0.6196
	2	173.3501	7.6115	0.5843	7	196.4128	8.754	0.6751
3	198.175	8.6334	0.6629	8	179.8835	7.9118	0.6056	
4	183.9683	8.0917	0.6195	9	182.1894	7.9497	0.6101	
5	183.2493	7.9365	0.6103	10	184.5148	8.1362	0.6227	
Forth Model ANN( 4 )	Activation function for hidden layer				sigmoid			
	Activation function for output layer				identity			
	Data preprocessing formula				normalized			
	Number of Hidden Nodes	MSE	MAE	MAPE	Number of Hidden Nodes	MSE	MAE	MAPE
	1	211.7407	9.3136	0.7118	6	182.8886	8.0308	0.6238
	2	222.3407	9.6718	0.7421	7	198.721	8.6888	0.6715
3	185.2495	8.2581	0.6293	8	181.8466	7.9717	0.6123	
4	195.651	8.7934	0.6728	9	220.4262	9.1451	0.7009	
5	191.5193	8.686	0.6657	10	185.0319	8.1435	0.6259	
Fifth Model ANN( 5 )	Activation function for hidden layer				Hyperbolic tangent			
	Activation function for output layer				Hyperbolic tangent			
	Data preprocessing formula				Adjusted normalized			
	Number of Hidden Nodes	MSE	MAE	MAPE	Number of Hidden Nodes	MSE	MAE	MAPE
	1	298.7304	12.7281	0.9971	6	307.0816	12.5883	0.992
	2	183.2822	8.6717	0.6709	7	284.4868	11.752	0.9224
3	175.6466	8.3228	0.6452	8	267.8135	11.5292	0.899	
4	274.034	11.7873	0.9162	9	281.4044	11.748	0.923	
5	283.5786	11.9748	0.9335	10	304.3154	12.8228	1.0107	

Regardless the number of the hidden nodes, by examining table (8) and based on MSE, MAE, and MAPE criteria, the best model among the designed models is ANN (3). Therefore, we conclude that the best activation function for the hidden layer is the bipolar function and the linear function for the output layer. In addition, by testing the third model, it is apparent that the best number of nodes for the hidden layer is 2. Therefore, we conclude that the best

formula to determine the number of hidden nodes is  $2R$ .

Based on the third model ANN (3) with 2 nodes and constant requirements of the other network except the sample size, data was divided to two sets including the training and the testing with specific portions as shown in table(9). The network was retrained again and the results are shown in table (9).

**Table 9 : Values of the criteria when data partitioning**

Training Data	Testing Data	MSE	MAE	MAPE
100	0	173.3501	7.6115	0.5843
90	10	184.357	8.2431	0.6318
80	20	183.1978	7.2871	0.5619
70	30	180.9971	7.763	0.5966
60	40	188.8886	7.575	0.5855
50	50	179.1047	7.9445	0.6062
40	60	198.5476	9.2811	0.7166
30	70	189.0676	8.8814	0.674
20	80	204.1265	8.046	0.6266
10	90	185.3909	7.3814	0.573

By examining table (9), we conclude that the inclusion of all the data in the training leads to the lowest potential error. This is clear through the MSE criterion that reach its lowest value when 100% of the data is included in the training, which indicates a safe primary selection of the data size.

Based on the third model with 2 hidden nodes and constant network requirements except the momentum value, different models were designed with different momentum values as shown in table (10) below.

**Table 10 : Values of the criteria during momentum change and fixed learning rate**

Learning Rate	Momentum	Time	MSE	MAE	MAPE
0.5	0.9	0:00:00.09	173.3501	7.6115	0.5843
0.5	0.8	0:00:00.08	184.3754	8.1804	.6289
0.5	0.7	0:00:00.05	171.2072	7.4626	.5748
0.5	0.6	0:00:00.05	170.5924	7.5583	.5800
0.5	0.5	0:00:00.17	188.2508	8.3374	.6406
0.5	0.4	0:00:00.37	180.5767	7.9174	.6067
0.5	0.3	0:00:00.08	173.5609	7.5861	.5816
0.5	0.2	0:00:00.03	171.6374	7.3010	.5630
0.5	0.1	0:00:00.09	197.1893	8.9972	.6925

By examining table (10), we conclude that the momentum value affects the training time and the error calculated by the network. We also conclude that the best value of momentum that gives the lowest possible error at appropriate time with 0.5 learning rate is 0.6.

Based on the last modifications of the third model of the network, a comparison was conducted between the network training without data processing and the network training with data processing by using normalized formula . The results are shown in table(11).

**Table 11: Values of criteria for final model without processing**

Final Model	MSE	MAE	MAPE
Without Processing	13345.0652	99.8602	7.5583
With Processing	170.5924	7.5583	.5800

By examining table (11), we conclude that data processing is a crucial step before providing the network with data. This is apparent through the values of MSE, MAE, and MAPE criteria as shown in table (11).

Therefore, the best model of the neural network that can be used in the estimation is the third model with 2 hidden nodes and 0.6 momentum at 0.5 learning rate.

methodology	MSE	MAE	MAPE
Spectral analysis	394.7326	11.2156	0.81900
artificial neural networks	170.5924	7.5583	.5800

### 7- Conclusions:

- 1- The series of exchange rate of Iraqi dinar to the US dollar is non-stationary in the mean and the variance.
- 2- The best model of the artificial neural network to forecast the exchange rate using backpropagation algorithm is the network designed with one variable ( $Y_t$ ), hyperbolic activation function in the hidden layer and linear activation function in the output layer, learning rate of (0.5), (0.6) momentum, and two hidden nodes in one hidden level.
- 3- The stationary exchange rate series contain hidden periodicities
- 4- Based on the MSE, MAE and MAPE criterion, it is apparent that artificial neural networks methodology is better than the spectral analysis in forecasting the exchange rate of Iraqi dinar to the US dollar.

### 8- Recommendations:

- 1- Compare the backpropagation networks and the Jordan or Elman networks in predicting the exchange rate.
- 2- Conducting a research on choosing the truncation point in the spectral windows and the lag windows
- 3- Apply the hybrid methodology to predict the exchange rate.

By comparing the calculated values of MSE, MAE, and MAPE for both methodologies, we conclude that the best methodology in forecasting the exchange rate is artificial neural network methodology as shown below.

### References

1. Wei,W., Time Series Analysis :Univariate and Multivariate Methods,United states of America. Pearson Addison Wesley, (2006).
2. Al-douri, A., The spectral analysis and harmonic analysis of the time series with application from the earthquakes in sulaimaniya governorate. Master thesis, University of Baghdad, Department of statistics, Iraq, (1991).
3. Chatfield, C., The analysis of time series: An introduction. NewYork, Chapman & Hall/CRC, (2004).
4. Jentsch, C. and Pauly, M., A note on using periodogram-based distances for comparing spectral densities. Elsevier, 82(1), PP.158 – 164, (201)
5. Halliday, D., Rosenberg, J., Rigas, A. and Conway,B., A periodogram –based test for weak stationarity and consistency between sections in time series. Elsevier, 180(1), PP.138-146, (2009)
6. Nowroozi, A., Table for Fisher’s Test of Significance in Harmonic Analysis. Geophysical journal international, 12(5), PP.517-520, (1967)
7. Novick, S., Analysis of Fisher’s Test for Hidden Periodicities. Master Thesis, Leigh University, Department of Applied Mathematics, (1995)
8. Fuller, W., Introduction to statistical time series. John Wiley & Sons, New York, (1996)



9. Hanke, J.E. & Wichern,D.W., Business Forecasting. PHI Learning, U.S.A., (2009)
10. Box,G., Jenkins,G., and Reinsel,G., Time Series Analysis: Forecasting and Control. John Wiley & Sons, New Jersey, (2008).
11. Hagan, M. and Demuth, H.B., Neural Network Design, Amazon, 2nd Edition. (2008).
12. Kulkarni, M., Patil, S., Rama, G. and Sen, P., Wind speed prediction using statistical regression and neural network. Journal of Earth System Science, 117(4)P.P: 457–463, (2008)
13. Meng, X., Jia,M. and Wang,T., Neural network prediction of biodiesel kinematic viscosity at 313 K. Elsevier, 121, P.P: 133-140, (2014).
14. Khashei,M. and Bijari,M., An artificial neural network (p,d,q) model for time series forecasting. Elsevier, 37(1), P.P: 479 – 489, (2010).
15. Tarsauliya, A., Kant, S. and Kala,R., Analysis of Artificial Neural Network for Financial Time Series Forecasting. International Journal of Computer Applications, 9(5), (2010).
16. Suzuki,K., Artificial Neural Networks- Methodological Advances and Biomedical Applications. InTech, India, (2011).
17. Zurada, J.M., Introduction to Artificial Neural Systems. west publishing company, United States of America, (1992).

### التنبؤ باستخدام التحليل الطيفي والشبكات العصبية الاصطناعية: دراسة مقارنة

محمد حبيب الشاروط  
جامعة القادسية  
كلية الادارة واقتصاد  
قسم الاحصاء  
m.alsharood@gmail.com

ايمان يوسف  
جامعة القادسية  
كلية علوم الحاسوب وتكنولوجيا المعلومات  
قسم الرياضيات  
emaan.yousif1991@gmail.com

#### المستخلص :

ان التقلبات التي تحدث في سعر الصرف ولا سيما الارتفاعات والانخفاضات المفاجئة وغير المتوقعة لها تأثير كبير على الاقتصاد القومي لأي دولة ومنها العراق. لذا فإن التنبؤ الدقيق بسعر صرف الدينار العراقي مقابل الدولار الامريكي له اثر كبير في عملية التخطيط واتخاذ القرار بالإضافة الى المحافظة على التوازن والاستقرار الاقتصادي للعراق .

يهدف هذا البحث الى المقارنة بين اسلوبي التحليل الطيفي و الشبكات العصبية الاصطناعية في التنبؤ بسعر صرف الدينار العراقي مقابل الدولار الامريكي بالاعتماد على البيانات المأخوذة من البنك المركزي العراقي للفترة من 30/01/2004 الى 30/12/2014 .

وبالاعتماد على متوسط مربع الخطأ MSE ،متوسط مطلق الخطأ MAE ومتوسط مطلق الخطأ النسبي MAPE كمعايير احصائية للمفاضلة بين الاسلوبين تم التوصل الى ان اسلوب الشبكة العصبية أفضل من اسلوب التحليل الطيفي في التنبؤ.