

# **A proposed Method for the Analysis of Multiple Regression using Artificial Intelligence**

**Dr**

**Mohamed Abdel Salam Agamy**  
Assistant Professor of Statistics

**Mahmoud El-Sayed Wahba**  
Demonstrator at Department of  
Statistics

**Faculty of Commerce - Al-Azhar University**



## **Abstract**

This research presents a proposed method to overcome difficulties which we face when we use traditional regression analysis method and provide solutions in an automated manner that does not require experience or a large field study. Some of these difficulties are mathematical difficulties, some of which are the need for considerable experience to choose the appropriate data model, especially when the traditional regression analysis method is not available.

The new approach raises forecast accuracy with powerful Neural Network MATLAB code. This approach automatically trains neural networks and applies them for regression analysis, thus getting accurate, business or stock market predictions doesn't require much effort or time. This means faster and more precise results than ever before.

### **This research aims to:**

1. Introducing a new method for the analysis of multiple regression using artificial neural networks, which represent one of the most important areas of artificial intelligence.
2. Comparing the suggested method of artificial neural networks with the traditional method of multiple regression analysis to determine which is better for prediction.
- 3- Conducting an applied study using realistic data.
- 4 - Formulation of a computer program using MATLAB software packages to compare the two methods.
5. Formulate a program using MATLAB software packages to predict regression models using both neural networks and conventional methods

### **It has become clear through the applied study that:**

- a. Proposed method can be used to model data and predict its unknown values.
- b - The proposed method using neural networks is better than traditional in the case of nonlinear models.
- c. Performance of the proposed method is better than traditional in the case of small samples.
- d - Convergence of the results of the two methods of increasing sample sizes.

### (1) Introduction

Regression analysis is one of the most widely used methods in all fields of knowledge and science. It can be said that regression analysis is one of the most widely used methods in all fields of knowledge and science where the value or value of a dependent variable can be predicted by knowing the values of independent variables. We may want to determine the demand for a commodity (dependent variable) Is the price of the item (independent variable), the size of the consumer's input (another independent variable) or their tastes ... etc. In other words, the most independent variables have an effect on the dependent variable. The traditional regression analysis performed through four stages: Model Identification, Parameters Estimation, Diagnostic Checking and Forecasting.

The problem of defining the model used in the forecasting or estimation process using regression models is one of the most important problems faced by researchers. Many factors and variables overlap in determining the choice of the appropriate model for analysis. Therefore, regression models can be classified according to these factors and the variables affecting them (Fahmi, 2005)

With the development of computer science - which has become a major factor in almost all activities of life - researchers have reached a kind of programs depends on the method of building a technique that simulates the human in the way he takes his decisions daily, called artificial intelligence.

Artificial Intelligence AI is one of the modern computer sciences, it is based on modern generation of computers, which aims to make computer able to simulate intelligence operations of human mind, so that the computer will have the ability to solve problems and make logical and uncluttered decisions to emulates the human brain. Its languages are: Lisp: List Processing Language, and Prolog: Programming in Logic, and these languages allow the programmer great potential in writing programs. Artificial intelligence includes several branches, including expert systems, data mining, genetic algorithms, and neural networks.

This research presents a proposed method for the analysis of multiple regression using artificial neural networks.

### **(1-1) Research problem:**

The problem of research is as follows:

- 1 - There is no clear way to analyze multiple regression using artificial intelligence, which will be presented in this research.
- 2 - Traditional methods face many difficulties in formulating the appropriate model for certain data. Some of these difficulties are mathematical difficulties, some of which require extensive experience in selecting the appropriate data model.
3. Lack of software packages ready to predict regression models appropriately using neural networks.

### **1-2 Importance of research:**

Importance of research is reflected in the use of neural networks to analyze multiple regression, where neural networks are working on:

- 1- Automatic prediction to be applied in the case of multiple linear and nonlinear regression models.
- 2- Get good predictions faster than the traditional method of multiple regression analysis.
- 3 - The possibility of using relatively smaller samples than those selected by conventional methods in any model.
- 4 - The possibility of dealing with the model and prediction of unknown values, even if not achieved any of the terms of the traditional multi-regression model.
5. To achieve maximum benefit of the available data for regression analysis and prediction of dependent variable values using artificial intelligence methods.

### **1-3 Research Objectives:**

The main objectives of this thesis can be summarized in the following:

1. Introduction of a new method for the analysis of multiple regression using artificial neural networks, which represent one of the most important areas of artificial intelligence.

- 3- Conducting an applied study using realistic data.
- 4 - Formulation of a computer program using MATLAB software packages to compare the two methods.
5. Formulate a program using MATLAB software packages to predict regression models using the two methods

#### **1-4 Previous literature**

There are many studies in the field of artificial intelligence in general and neural networks, the following are the most important:

In 1996, Arminger presented one of the most important researches on the relationship between statistical models and artificial neural networks, which he explained that statistical models can be expressed using neural networks and that the estimates of the parameters of statistical models can be represented in the weights of the neural network.

In the study of Rashad, Zaki (2010) presented multiple regression analyzes using neural networks and compared them with the method of stepwise regression analysis. Neural networks were found to be more accurate than the linear regression method in detecting variables with insignificant differences.

In the study of Alhakak, Aljarrah(2011) proposed the application of hybrid models consisting of exponential boot models and neural networks with another model of neural networks, the ARIMA model, and another composite model of exponential boot, neural networks and ARIMA using multiple regression to calculate integration weights. Then he applied these models on the daily exchange rate of the US dollar against the Iraqi dinar in the Iraqi market in the period 2011-2012. The ANARIMA model was chosen as the best model according to the MSE comparison scale.

In the study of Aljammal, Saber (2012) used a hybrid method between artificial neural networks and linear regression model. The mechanism of this hybrid method was described and compared with the neural network models and the linear regression model to determine the efficiency of this technique.

In the study of the Almorad (2012), where the regression was analyzed using the classical methods and the techniques of intelligence and the comparison between them and it was concluded that the neural networks are better in their results.

Ali(2013) presented a new way of determining the relationship between economic variables that explain their behavior with the aim of analysis or forecasting, or both, such as studying the impact of investment on GDP and its impact on both national income and average growth rate The individual in this model is multidimensional.

The neural network method was also predicted by identifying the study problem, then designing an appropriate neural network and designing an automated algorithm that mechanically trains and trains the network to minimize the error factor by adjusting the value of the weights to estimate the model's structure. To assess the accuracy of the model, to measure the quality of the system as well as the use of some statistical hypotheses for each equation and conduct the test accuracy of predictions and their ability to reduce the size of error, especially in nonlinear models where the neural networks proved fully capable in such models.

The study of the Alsalmy, Afaf (2017) aimed to introduce the applications of artificial intelligence to support the retrieval of information, and explained the characteristics of artificial intelligence that distinguish it from natural intelligence, and highlighted the use of computer applications of artificial intelligence in the promotion of information retrieval services and development, Analytical approach through reference to the methods of modern technology in the Internet. The present study aims at providing a new and effective way to predict the use of Artificial Neural Network (ANN) technology, comparing the proposed method with the traditional method of multiple regression.

## **(2) Artificial Neural Networks:**

Artificial Neural Network (ANN) is a data processing system that simulates the human brain to achieve performance similar to human performance in solving some problems. Artificial neural networks are a way of artificial intelligence that offers alternative solutions as well as their ability to process data without preconditions, mimicking the structure and functioning of the nervous system in humans, and entered the global business circle by Mc-Cultch & Pilts (1943) ), Sometimes called Deep Learning.

Deep Learning is just a new name for an old technology. Neural networks are used in many fields and applications, including signal processing, control, pattern recognition, medicine, speech production, speech recognition [Allam Zaki, 2000, pp. 15-19].

The idea of the proposed method is to design an appropriate artificial network that simulates the multiple regression model and then train this network using input and output data available to reach standard weights of the model. This is called supervisory training.

The network performance is then tested through several limited criteria. After successfully passing these criteria, the network is used to predict the unknown values of the dependent variable in the multiple regression.

## **(3) Stages of Building Model in Neural Networks**

The stages of multiple regression analysis using artificial neural networks or deep learning method are:

- (1) Specification of network architecture or neural network design.
- (2) Determination of network weights.
- (3) Network performance testing.
- (4) Forecasting or prediction.

The following diagram shows the four stages: Anders, 1999



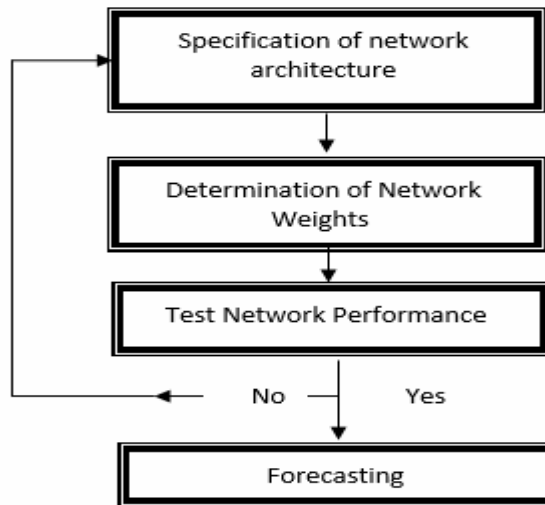


Figure (3-1) stages of multiple regression analysis using neural network

### (3-1) Network Design:

Network design or structural characterization is intended to design a suitable network for multiple regression analysis. This network includes Input Layer, Output Layer, Hidden Layers, Network Function, Connectors between Inputs, Hidden Layer and Output. The values of the independent variables are  $X_1, X_2, X_3, X_4, \dots, X_K$  and output are the values of the  $Y$  variable.

This network uses the Supervised Learning method because it contains input and output. The appropriate function for multi-regression analysis in MATLAB is "trainlm" which is a Levenberg-Marquardt function to train the neural network and update their weights and the amount of bias.

It is operating according to the examples for "Levenberg-Marquardt backpropagation" the fastest training technology deployment back function and is the first choice recommended in the case of supervisory training and do not need more than the other methods additional memory. The number of hidden layers was chosen to be 2 layers. This stage includes the collection and processing of input and output data for network training as well as designing an appropriate network for multiple regression analysis.

a) **Data collection and processing:** Data is collected and processed for both the dependent and the independent variable. Data processing including **normalization** stage, **standardization** stage, **transforming** stage and addressing problems of **missing values and Outliers**.

**Normalization** of data meaning that making statistical data more usable in the programming and analysis of the normalization of data than in its raw state before normalizing it, in some cases we need to convert input values into standard values using a normal distribution function, and this normative value may range from 0 to 1 or -1, so as to give optimum results and avoid the effect of extreme values (Fortin et al, 1997). However, Burke and Lgnizio (1992) stated that it is not necessary to know the probability distribution of input data.

**Scaling:** A basic process for output data, because output data must match the range of conversion functions used in the output layer (for example, values from -0.1 0.1 for tanh transfer and 0: 1 for sigmoid transfer. That is, the data entered into the network must be calibrated, that is, make it standard by subtracting each value from its average and then dividing the output by its standard deviation.

**Transformation:** The process of conversion in the process of network manipulation means converting input data into certain known forms (linear, logarithmic, exponential, etc.) may be useful for improving the performance of the neural network (Masters 2008).

There are many types of networks that can be used in multiple regression analysis in the suggested way. For example, in MATLAB there are 13 networks that can be used in this field:

**traincgf – traincgp – traingd – traingda – traingdm – traingdx – trainlm – trainoss – trainrp – trainscg – trainb – trainc – trainr .**

The following table shows the objective of each network and its effectiveness and impact in training the network:

| <b>Training Functions</b> |  |
|---------------------------|--|
| traincgf                  | Fletcher-Powell conjugate gradient backpropagation.          |
| traincgp                  | Polak-Ribiere conjugate gradient backpropagation.            |
| traingd                   | Gradient descent backpropagation.                            |
| traingda                  | Gradient descent with adaptive lr backpropagation.           |
| traingdm                  | Gradient descent with momentum backpropagation.              |
| traingdx                  | Gradient descent with momentum and adaptive lr backprop.     |
| trainlm                   | Levenberg-Marquardt backpropagation.                         |
| trainoss                  | One-step secant backpropagation.                             |
| trainrp                   | Resilient backpropagation (Rprop).                           |
| trainscg                  | Scaled conjugate gradient backpropagation.                   |
| trainb                    | Batch training with weight and bias learning rules.          |
| trainc                    | Cyclical order incremental training with learning functions. |
| trainr                    | Random order incremental training with learning functions.   |

**Table (3-1) Types of functions used in the design and training of the network**

For example, the "trainlm" function, we find that it acts as a preliminary training function for the network and displays and changes the weights of that network optimization of data and change the values of parameters through the following codes:

**net.trainFcn = 'trainlm' ... .. (3 – 1)**

**where:** trainFcn: A function that specifies the default parameters.

trainlm: Network training function

As well as the rest of the functions used in the training of the network has been shown using experimental samples that the best is the network "trainlm" where it gave highly efficient results. This network uses a method called general manner Levenberg-Marquardt and we will use the same network in the proposed method for the analysis of multiple regression. The following diagram shows multiple regression artificial neural network model. Where X represents the independent variables included in the program "Input", and Z represents the hidden layers that the network "Hidden Layers" in which training has

been, and Y represents the dependent variable which outputs the data are ready "Output".

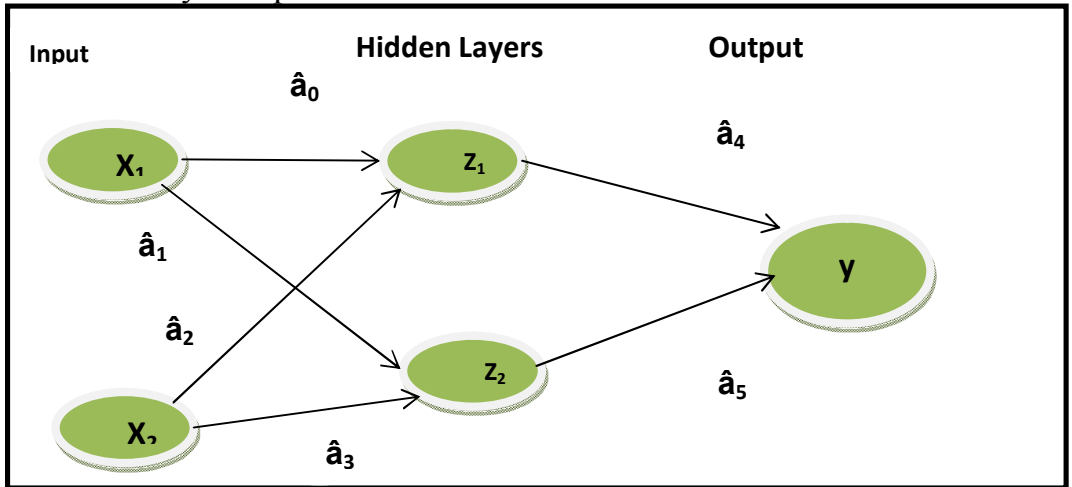


Figure (3-2) Multiple regression model as artificial neural network

The neural network model, therefore, can be in the following form:

$$Y = g(X, W) + \varepsilon \quad \dots\dots\dots (3-2)$$

Where:

- X: represents the input of the network
- Y: represents the output of the network
- W: represent the weights of the grid.
- $\varepsilon$ : Random error limit.

That is, the linear form can be written as follows:

$$Y = Z \cdot W_2 + \varepsilon_2 \quad \dots\dots\dots (3-3)$$

Where:

$$W_2 = \begin{bmatrix} \beta_4 \\ \beta_5 \end{bmatrix}$$

$$Z = X \cdot W_1 + \varepsilon_1 \quad \dots\dots\dots (3-4)$$

Where:

$$\beta_1 = [\beta_0 \ \beta_1 \ \beta_2 \ \beta_3] \quad , \quad X = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix}$$

$\beta_1, \beta_2$ : Model parameters (network weights).

Substituting Z in the Y function we get:

$$Y = (X W_1 + \varepsilon_1) W_2 + \varepsilon_2$$

$$Y = X W_2 W_1 + \varepsilon_1 W_2 + \varepsilon_2 \quad \dots \quad (3-5)$$

by placing:  $\beta = W_2 W_1$  ,  $\varepsilon = \varepsilon_1 W_2 + \varepsilon_2$

We get :

$$Y = X \beta + \varepsilon \quad \dots\dots\dots (3-6)$$

Y: A dimension vector ( $n \times 1$ ) represents the views of the dependent variable sample.

X: The matrix of its dimensions ( $n \times k$ ) represents the observations of the sample of the independent variables and their number( $k-1$ ).

$\beta$  : A column vector representing the unknown community parameters of  $k$ .

$\varepsilon$ : A dimension column vector ( $n \times 1$ ) represents random errors

Thus, we have obtained the traditional model of multiple regression. It is possible to add more than one hidden layer and we will get a model like the previous model (3 – 6). The structure of the network used in this study is the Multi-Layer Feedforward Neural Network (MFNN). The multilayered front-end network can solve many complex problems that the single-layer network cannot resolve. The structure of this network is one of the most used neural network structures. In this type of network, the structure of the model is based on two basic bases: Mc-Cultch & Pilts (1943) ).

**First: Selecting the network input.**

**Second: Selecting internal network architecture.**

The network input and output ( $x, y$ ) and the 2 hidden layers are selected. This network is suitable for regression analysis and other details of network design are selected automatically in MATLAB or any other program including Artificial intelligence, and allows modifications if necessary.

The structure of this network has many advantages, including:

- 1- Natural extension of the known linear model structures.
- 2 - The internal construction enables the gradual expansion of a high degree of flexibility in accordance with the model.
3. The required structural decisions shall be within the reasonable limits.

To set up a front-end feedback network, we'll use the following command in MATLAB:

**net = fitnet(2,'trainlm') ... (3 - 7)**

Where 'fitnet' is the function of neural network design, and the number of hidden layers used in this network is 2, which is enough to train a neural network for regression analysis, as demonstrated by the experimental sample study.

"Trainlm" represents the name of the function used in a subsequent network training, and by examining experimental samples showing that this function is the best function for multiple regression analysis.

### **3-2) Determination of network weights:**

Estimating the network weights (or parameters of the model) means that we use artificial neural networks in estimating the appropriate value for each weight in the network design as well as estimating the appropriate bias for each. This is done through several criteria, the most important of which is to get the average of the MSE to the lowest possible value. The network weight estimation is obtained by training the network using the training data set, which represents the initial information that the network will learn. Therefore, the weights must be updated during the training phase, and for this update several different algorithms are used (within the MATLAB packages) according to network type. The most important of these algorithms is the reverse scatter diagram algorithm used in the training of multi-layer neural networks with front and nonlinear feeders.

Backscatter diagram algorithm: The most commonly used training method in neural network training. The main objective of using this method of training is to calculate the error between the target output and the actual output in the output layer and then return it again to each hidden layer in order to adjust the communication weights until Input layer, and this adjustment continues in weights until the network achieves less total square errors during training (Dhar and Stein, 1996).

During training, network weights change and are reallocated in order to reach the minimum value of the cost function or the so-called performance function. The default

performance of the front feed networks is Mean Square Error (MSE). In most neural network applications, these weights are determined by reducing the Sum of Squared Errors (SSE). The network training is done within MATLAB program using following command:

$$\mathbf{net1} = \mathbf{train}(\mathbf{net}, \mathbf{x}, \mathbf{y}) \quad \dots \dots \quad (3 - 8)$$

Where :

train: a special command to start the network training process.

net: a network that is designed at the network design stage.

X: matrix of independent observed variable.

Y: vector of dependent observed variable.

net1: is the resulting network after estimating the weights.

### 3-3) Test network performance

This stage is one of the basic stages in the analysis of multiple regression using artificial neural networks. According to this stage it is determined either to continue the analysis process and to achieve what is intended from the analysis model or to return to the starting point of identification, estimation and then testing. In order to test the validity of the neural network model, the data is divided into two groups of training groups and a test group, an important process of modeling using neural networks, where the outputs of each model are given the expected results. The quality of the neural network performance is tested through several criteria similar to those used in the traditional regression model quality test. It is based mainly on the principle of attaining the lowest value of MSE or any measure that represents a function of errors, these criteria include:

#### 3.3.1 Mean Absolute Deviation (MAD):

$$\text{MAD} = \frac{1}{n} \sum_{i=1}^n \frac{|Y_i - \hat{Y}_i|}{Y_i} \quad (3-9)$$

Where:  $Y_i$ : Target observation.

$\hat{Y}_i$ : Predicted value.

$n$  : observations or sample size.

**3-3-2 Mean Squared Error (MSE):**

$$MSE = \frac{1}{n} \sum_{i=1}^n \left( \frac{Y_i - \hat{Y}_i}{Y_i} \right)^2 \quad (3-10)$$

Where:  $Y_i$ ,  $\hat{Y}_i$  and  $n$  as defined by equation (3 - 9).

**3.3.3 Mean Absolute Error (Ratio):**

$$MAPE = \left( \frac{1}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \right) \times 100 \quad (3-11)$$

Where:  $Y_i$ ,  $\hat{Y}_i$  and  $n$  as defined by equation (3 - 9).

**3.3.4 Information criteria:** It is the relationship between the residual square and the number of the model parameters. Its purpose is to weigh the model error against the number of parameters, since additional parameters should be entered only if the value of the information criterion is small and the selected is the model with the smallest value of the information standard.

The most important information standards are:

**3.3.4.1 Information Standard (AIC):** This standard is known as the Akaike information criterion (AIC), which is calculated by the following relationship (Akaike, 1973):

$$AIC = \ln \left( \frac{\varepsilon \varepsilon'}{n} \right) + \frac{2k}{n} \quad \dots\dots\dots (3 - 12)$$

**Where:**  $\varepsilon$  : Error limit.

$k$  : Number of free form parameters.

$n$  : observations.

**3.3.4.2 Network Information Standard (NIC):**

In order to avoid the defect of the (AIC) which is not valid for models with misconfiguration, Murata / Yoshizawa / Amari (1994) proposed an information standard called the Network Information Criterion (NIC) which calculated from MSE and the penalty term as following:

$$NIC = MSE + \frac{1}{n} . tr[BA^{-1}] \quad \dots\dots\dots (3 - 13)$$

**Where :**

**A:** The result of multiplying square slope of weights by square error of observations  $SE_n$



**B:** the expectation of the result of multiplying the vector slope in its transpose by the square error.

Since the network must be able to draw the exact function, which means that the network has no error in characterization, and therefore can achieve the approximate relationship:

$$B = 2\sigma^2 A \dots\dots\dots (3-14)$$

This leads to:

$$tr[BA^{-1}] = 2\sigma^2 tr[I] = 2\sigma^2 k \dots\dots (3-15)$$

Thus, the NIC criterion can be reduced to the AIC formula presented by Amemiya (1980):

$$AIC = MSE + \sigma^2 \frac{2k}{n} \dots\dots\dots (3-16)$$

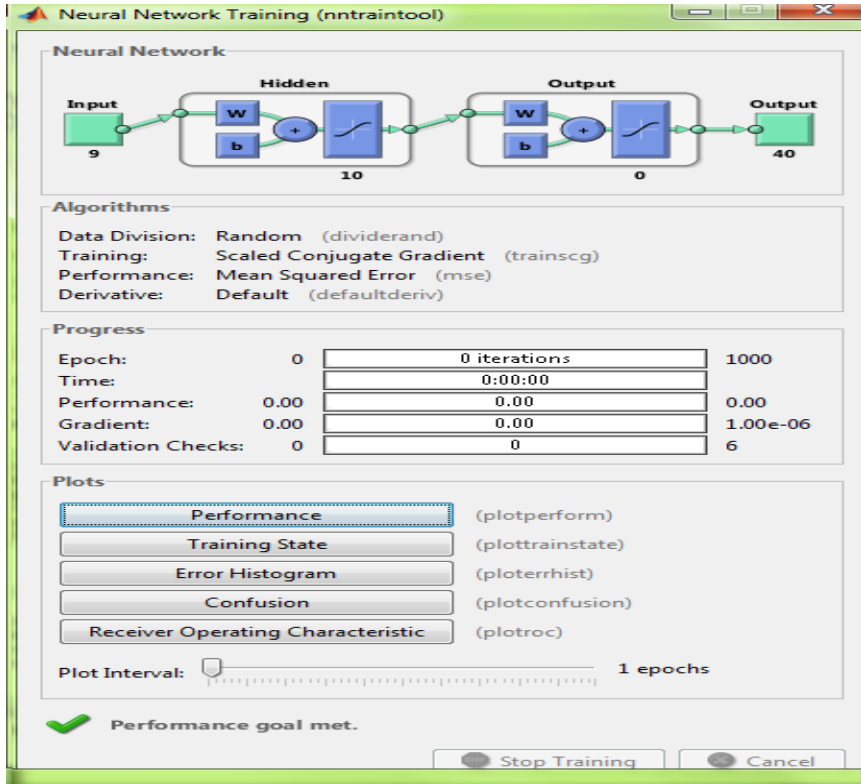
The NIC can be used to determine the weights that should remain in the network model. The NIC also gives an indication of whether or not the network contains hidden units that are not required. In the case of hidden units that are not required, the information criterion takes a high value.

**3.3.4.3 Cross Validation Method:**

This method uses the Validation Set to determine when neural network training is stopped. This method can be used in the case of non-formal models and neural networks (Hardle W., 1990). Neural network training continues as long as it is reduced the error resulting from the neural network. Using this method, it is possible to calculate the predicted error of unknown data. Among the many models, the model that gives the least error is selected. In order to estimate this error, the observations are divided into M equal groups then selecting one sub-group as a validation group and the rest (M-1) groups used to train the network and estimate the parameters, then the error is calculated from the sub-group, the process is done with all sub-groups. Here is the general forecasting error calculated by cross-validation method, which is the average of all MSE for all different validity groups, which can be written as follows (Anders, 1999):

$$CV = \frac{1}{M} \sum_{m=1}^M MSPE_m \quad \dots\dots (3-16)$$

One of the main drawbacks to cross validation in the validity test is that it is a very cumbersome criterion, because the network training to estimate parameters in this way must repeat M times.



The previous figure shows the most important criteria used for network training and weight estimation (model parameters estimation) as well as validation of the model in the MATLAB program.

### 3-4) Forecasting:

After confirming the quality of the design and performance of the network, it is used to predict the unknown observations of the phenomenon. The trained network is fed with the values of the independent variables  $x$  and we ask them to estimate the unknown values of the dependent variable  $y$ .

This is done through an appropriate computer program that has been corrected in this message for this purpose.

To predict the values of the dependent variable in multiple regression using artificial neural networks, the following function is used:

$$Y = \text{sim}(\text{net}, \text{xx}) . \quad \dots \dots \quad (3-17)$$

Where :

- Y: the predicted values that are estimated for the dependent variable y using the artificial neural network method.
- net: represents the network that is designed and trained.
- XX: represents the array of independent variables and their dimensions (n.k).

#### **)4) Applied study**

##### **(4 – 1) Introduction:**

In this section, the proposed method using the neural network or deep learning method, one of the branches of artificial intelligence, will be applied to multiple regression analysis and compare this method in the traditional way. The study relies on data published in the audited financial statements of 60 companies listed on the Egyptian Stock Exchange These lists and financial reports were provided through EGID through the annual publication issued by the Central Agency for Public Mobilization and Statistics (CAPMAS) for 2015. The study variables include a variable and five independent variables:

- X<sub>1</sub> : Investment policy in working capital.
- X<sub>2</sub> : Working Capital Financing Policy.
- X<sub>3</sub> : CashRatio.
- X<sub>4</sub> : Ratio of fixed financial assets to total assets.
- X<sub>5</sub> : Growth opportunities in the company's sales.

Y variable represents the return on total assets (market yield).

The studied population is composed of Egyptian shareholding companies listed on the Egyptian Stock Exchange for easy access to their data from the Egyptian Stock Exchange, in addition to the credibility and reliability of the data and financial reports published by these companies, due to the degree of scrutiny and control subject to it from the General

Authority for Financial Supervision and sovereign entities, and in the appendix (A) a table that specifies the values of the independent variables and the dependent variables of the 60 companies.

In this research we examine the multiple regression analysis of the Y dependent variable on 5 independent variables  $X_1, X_2, X_3, X_4$  and  $X_5$  using the proposed by method and compare it with the traditional method. In order to accomplish this, one of the MATLAB software packages was used. A suitable computer program was designed using MATLAB packages (attached in Annex B) The results obtained from the proposed method presented in subsection (4-2) while the results of the traditional regression analysis of the applied study data will be presented in (4-3). In the last section (4-4), the proposed methods using the neural networks and the traditional method of multiple regression will be compared.

#### **(4.2) Analysis of the proposed method:**

The stages of multiple regression analysis using neural networks are:

- (1) designing suitable network.
- (2) Determination of network weights.
- (3) Testing the network performance.
- (4) Forecasting.

Below we review what was done during each stage in the proposed method for multiple regression analysis using the neural network method through the data of the applied study.

##### **4.2.1 Network Design:**

Appropriate network includes an input layer with independent variable values and a output layer with dependent variable values. A front-end network has been designed to reverse-feed the error using the following command in Matlab:

$$\mathbf{net} = \mathbf{fitnet}(2, \mathbf{trainlm}') \quad \dots \dots \dots \quad (4-1)$$

The 'fitnet' function is called the 'function fitting neural' " network', which is called net and includes 2 hidden layers.

The "trainlm" function is used in training this network.

The remaining details of the network design are automatically determined in the MATLAB program, with the possibility to modify these details if necessary. The "trainlm" function has been used because it has proven efficient as it serves as a preliminary training function for the network and displays and changes the weights of that network well in the case of multiple regression models it changes the values of network weights according to the nature of the data.

This network uses a general method called Levenberg-Marquardt backpropagation, where X represents independent variables that are entered into the input grid. Z represents the hidden layers. Y represents the dependent variable that outputs output data. We have taken 2 hidden layers.

#### **4.2.2 Determination of network weights:**

The weights of the network are obtained by training the network using the training data set. These weights represent the initial information that the network will learn, and then the weights are updated during the training phase. For this update, several different algorithms (within MATLAB packages) were used according to network type. One of the most important algorithms is the (Back Scatter diagramAlgorithm), which is used in the training of multi-layer neural networks with frontal feed. During the training, the network weights and the amount of bias are changed repeatedly until the minimum value of the cost or so-called performance function is reached. The default performance of the front feed networks is the (MSE).

#### **Backscatter diagramalgorithm:**

This technique was first introduced by Paul Werbos (1974) and was developed by (Rumelhart Hinton, Williams, 1986) by constructing its first neural network with more than a hidden layer. The main objective of using this method of training is to calculate the error between the target output and the actual output in the output layer and then return it again to each hidden layer in order to adjust the weights of the connection until the input layer is reached. This adjustment is done in weights until the network achieves the lowest MSE in training (Dhar and Stein, 1996).

This was done using the MATLAB data using the following command:

$$\text{net1} = \text{train}(\text{net}, \text{X}, \text{Y}) \dots \dots \dots (4-2)$$

The network is evaluated and validated using the X inputs that represent the independent variable values matrix and the Y output vector representing the dependent variable values for the trained network “net1”.

#### 4.2.3 Test network performance:

Network performance has been tested through several criteria, the most important of which is to obtain the lowest MSE where it calculated as follows:

$$\text{MSE} = \frac{\sum [(Y_i) - \hat{Y}_i]^2}{n} \quad (4-3)$$

In addition to other criteria including the number of times the network weight adjustment (parameters of the model) called Epoch, and in order to test the performance of the neural network, the available data is divided into two groups:

Training group: This group usually contains 70% of sample views and is used to obtain the appropriate form.

Test group: This group contains 30% of the sample observations. The group used in the training is not used in the test.

Some of the statistical tools used to test the performance of the neural network are some of the appropriate measures for multiple regression analysis have been reviewed previously. The measures to be used in the study are MAD, MSE, MAEP, for their common use in multiple regression analysis

#### 4.2.4 Forecasting:

After confirming the validity of the model, it is used to predict unknown observations of the phenomenon. The efficiency of statistical models is measured by their ability to represent reality and their predictability of unknown values, and the less correct models are the best models.

The comparison of the good models is done by using the estimated total prediction error using these models.

When predicting multiple regression values using artificial neural networks using MATLAB, the following command was used:

$$Y0 = \text{sim}(\text{net}, X) \quad \dots \dots \quad (4-4)$$

Where:

Y0: represent the predictive values of three unknown values of the dependent variable (the return on total assets) using the artificial neural network method

net: the network that has been trained using a supervisory training method.

X: The independent variables values matrix.

#### 6\4.2.5 Proposed method analysis results:

After calculating the predictive values of the three unknown observations of the Y variable using the neural network method, the accuracy of these predictions was calculated by statistical measures based on the difference between the predictive value and the actual value of the multiple regression (residuals)

##### • Mean deviations MAD:

The MAD scale for artificial neural networks is calculated as follows:

$$MAD(x) = \frac{1}{n} \sum |y_x - y_{ANN}| \quad \dots \dots \dots \quad (4-5)$$

Where:  $y_x$ : represent the real observations.

$y_{ANN}$ : the prediction of neural networks method.

##### • Mean Squared Error MSE:

The MSE standard for artificial neural networks is calculated as follows:

$$MSE(x) = \frac{1}{n} \sum (y_x - y_x(ANN))^2 \quad \dots \quad (4-6)$$

The following is a summary of the results of the applied study using the suggested method of artificial intelligence:

**First: MAD results for artificial neural network predictions:**

After obtaining the values of the MAD through the MATLAB program for prediction of neural networks, for the variable Y is summarized in table (4-1):

Table (4-1) MAD values to predict the three unknown observations:

| <b>MAD</b>    | <b>Pred. horizon</b> |
|---------------|----------------------|
| 0.0515        | $Z_1$                |
| 0.0117        | $Z_2$                |
| 0.0128        | $Z_3$                |
| <b>0.0253</b> | <b>Average</b>       |

From table (4-1), the MAD values for the three unknown observations of the dependent variable are relatively close, and the lowest  $Z_2$  was 0.0117 and the highest  $Z_1$  was 0.0515.

### **Second: MSE results for artificial neural network predictions:**

Table (4.1) shows the amount of MSE through the MATLAB program for prediction of neural networks distributed by the three unknown observations.

Table (4-2) MSE to predict the three unknown observations

| <b>MSE</b>   | <b>Pred. horizon</b> |
|--------------|----------------------|
| 0.0027       | $Z_1$                |
| 0.0001       | $Z_2$                |
| 0.0002       | $Z_3$                |
| <b>0.001</b> | <b>Average</b>       |

From the table (4-2), the MSE values for the three unknown observations of the dependent variable are relatively close. The lowest of them is  $Z_2$ , where it was 0.0001 and the highest  $Z_1$  was 0.0027.

### **4.3 Analysis with the traditional method:**

Multiple regression analysis involves four stages: model identification, characterization, validation, and use of the model in prediction. Each of these four stages is reviewed in some detail next:

#### **4.3.1 Model Identification:**

To identify a multiple regression model in the traditional method of applied study data, the scatter matrix was plotted as shown in Fig. 6-1. The correlation matrix of the dependent



variable and the five independent variables were calculated. Multiple regression was defined in the traditional way in MATLAB using the following command:

$$[aa \ b \ cc2] = \text{regress} (y, x) \quad \dots \dots \dots \quad (4-7)$$

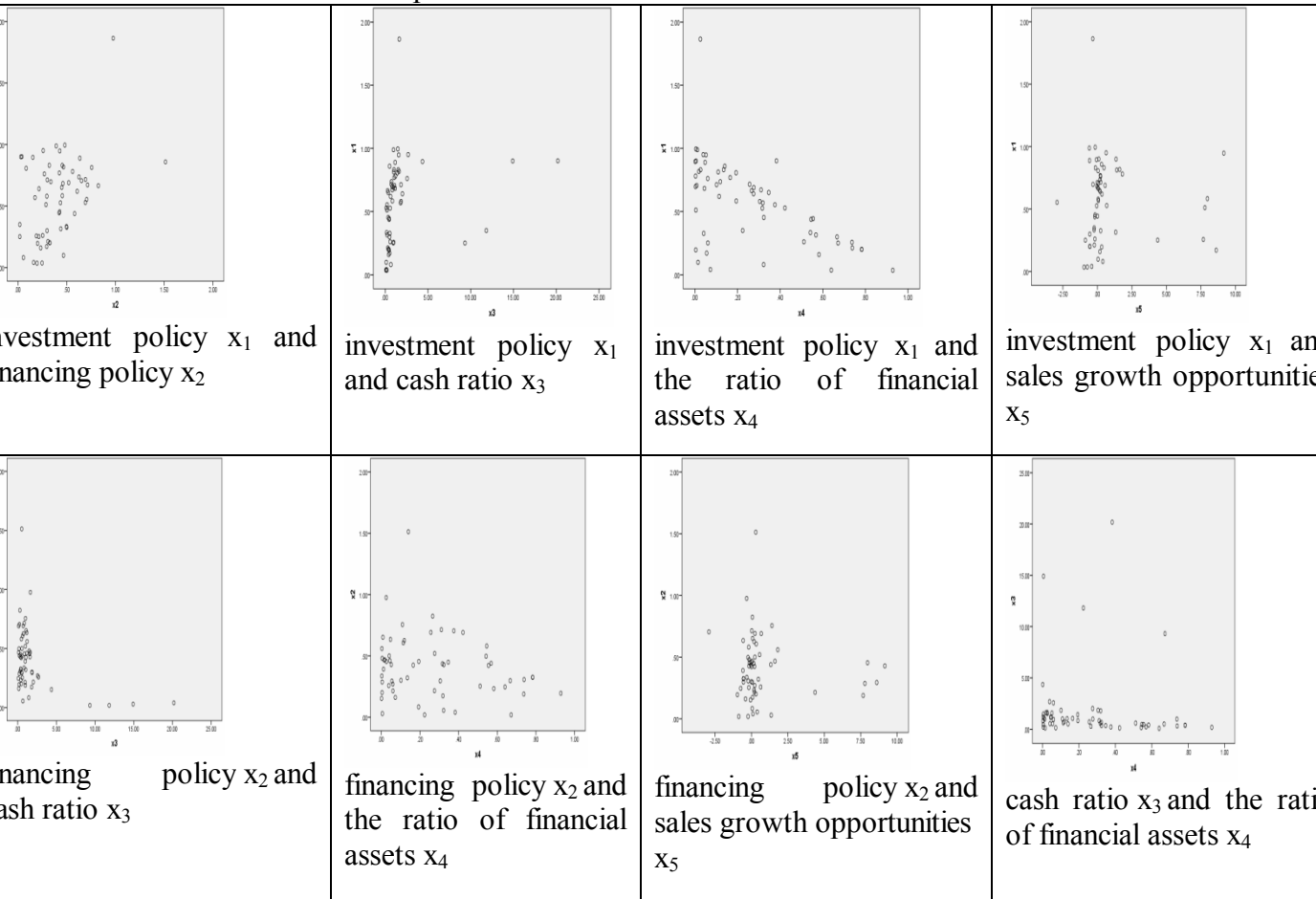
Where:

Y: represents the vector of dependent variable values.

X: represents the matrix of the five independent variables values.

regression: Performs a multiple regression analysis of the dependent variable Y on five independent variables X.

cc2: Represents the form residuals.



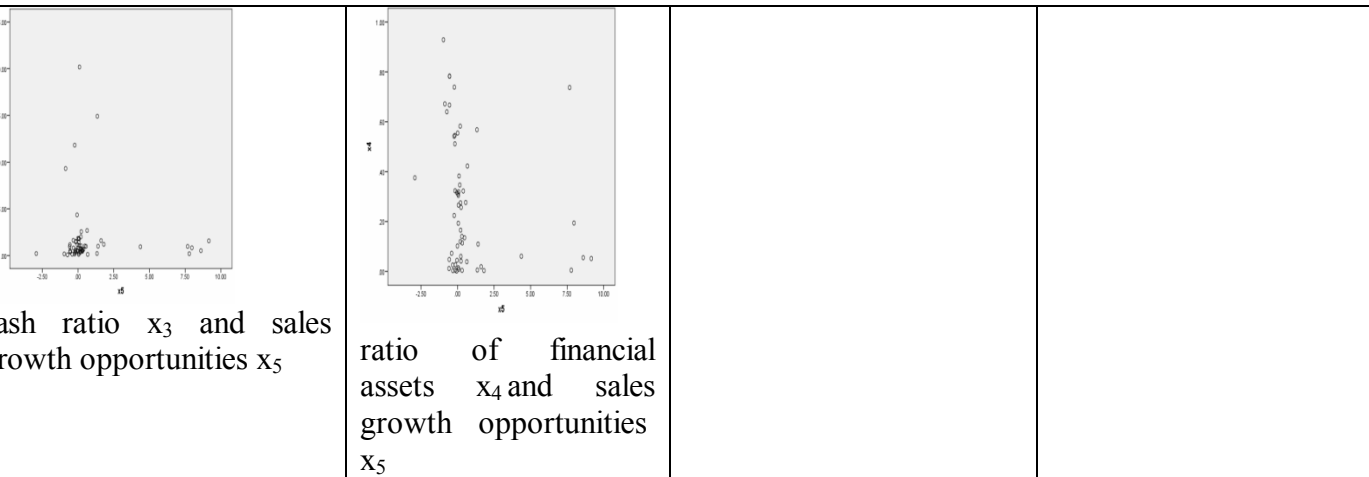


figure (4-1) The scatter diagram of the Relationship between each independent variable and the other

It is clear from previous forms that independent variables do not have a self-correlation.

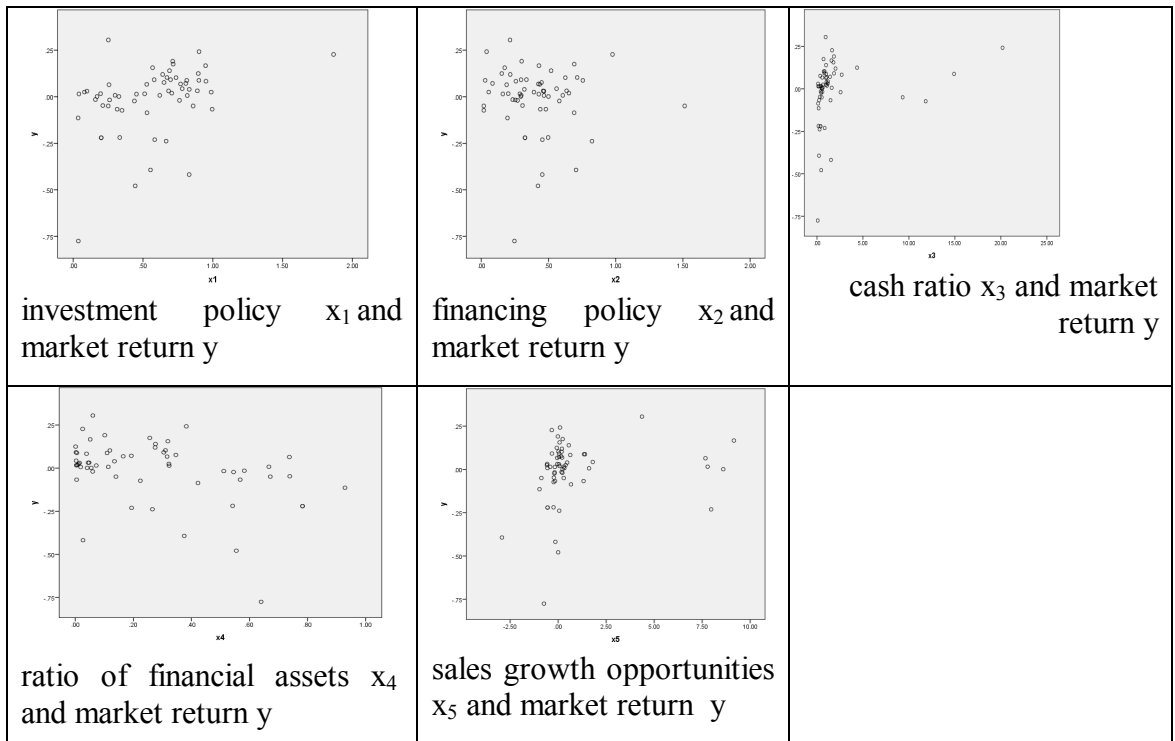


Figure (4.2): The scatter diagram of the relationship between each independent variable  $x_i$  and the dependent variable  $y$

The scatter diagram matrix shows a linear relationship between the variable  $y$  and each independent variable  $x_i$ . This is confirmed by the linear correlation matrix showing a linear correlation between the dependent variable and each independent variable. The proposed model for representing the relationship between  $y$  and independent variables is:

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \beta_5 x_5 + \epsilon \dots \dots \dots (4-8)$$

Where:  $Y$ : Return on total assets (market return).

$X_1$ : Investment policy in working capital.

$X_2$ : Working Capital Financing Policy.

$X_3$ : CashRatio.

$X_4$ : Ratio of fixed financial assets to total assets.

$X_5$ : Growth opportunities in the company's sales.

$\beta_0, \beta_1, \beta_2, \beta_3, \beta_4, \beta_5$  : Regression coefficients to be estimated.

#### 4.3.2 Estimation of Model parameters:

Estimating the parameters of the model is meant to estimate the values of the coefficients  $\beta_0, \beta_1, \beta_2, \beta_3, \beta_4, \beta_5$ .

$$[a \ b \ c] = \text{regress} (y, x) \dots \dots \dots (4-9)$$

Where:

$Y$ : The values of the dependent variable.

$X$ : Matrix of independent variables values.

$C$ : Residual values.

#### 4.3.3 Diagnostic Checking of the Model:

Three of the other criteria used to judge the validity of the model are the three criteria:  $R^2$ , F test, and series of t tests, and these have shown the quality of the model.

#### 4.3.4 Use the model in prediction:

After diagnostic checking of the proposed regression model of applied study data in the traditional method, it was used to predict future values. What concerns us here is the amount of model residuals that will be compared with the model residuals estimated using the proposed method using neural networks. Model residuals represent the vector  $c$  in the following command of MATLAB program (Appendix A) used in multiple regression analysis.

$$[a \ b \ c] = \text{regress} (y, x) .$$

#### 4.3.5 Traditional analysis method results:

After calculating the predictive values of the three unknown observations of the Y variable using the traditional method of multiple regression, the accuracy of these predictions was calculated using the statistical measures already mentioned in the previous section, which depend on the difference between the predictive value and the actual value of the multiple regression.

#### First: MAD results of prediction for traditional regression methods:

The next table (4-3) shows the MAD values calculated using the MATLAB program for conventional regression prediction...

Table (4-3) MAD values for the traditional method of multiple regression

| MAD           | Pred. horizon  |
|---------------|----------------|
| 0.0817        | Z <sub>1</sub> |
| 0.0419        | Z <sub>2</sub> |
| 0.0174        | Z <sub>3</sub> |
| <b>0.0470</b> | <b>Average</b> |

From the above table, the lowest value for MAD is corresponding to the unknown view of Z<sub>3</sub> where it was 0.0174 and the highest value for MAD is 0.0817 which corresponds to the view of Z<sub>1</sub>.

#### Second: MSE results for the traditional regression method:

Table (4-4) shows the amount of MSE using the MATLAB program for traditional method predictions distributed by the three anonymous observations.

| MSE           | Pred. horizon  |
|---------------|----------------|
| <b>0.0067</b> | <b>Z1</b>      |
| <b>0.0018</b> | <b>Z2</b>      |
| <b>0.0003</b> | <b>Z3</b>      |
| <b>0.0029</b> | <b>Average</b> |

Table (4-4) MSE values for the traditional method of multiple regression predictions

From the previous table, the lowest value for MSE is corresponding to the  $Z_3$  observation and its value is 0.0003 and the maximum value is 0.0067 which corresponds to the  $Z_1$ .

#### **4 - 4 Comparison of simulation results in the two methods:**

The aim of this research is to try to achieve a new and effective method to analyze the multiple regression using artificial neural networks ANN. In order to verify the quality of the new method, it is necessary to compare it in another method that proved its efficiency, which is the traditional method of multiple regression. In order to compare the new method with the traditional method.

To compare the new method with the traditional method, data were collected from 60 companies listed on the Egyptian stock exchange. The comparison between observations predictions and their true values is used to measure the accuracy of predictions on the one hand while the performance of the suggested method of artificial neural networks is compared with the performance of the traditional method on the other hand to measure the accuracy of the proposed method of prediction according to three famous criteria for accuracy of prediction (MSE, MAD, MAEP).

The results of these measures for both methods will be presented to measure the quality of the proposed method by the quality of their predictions compared to the real data on the one hand and comparing the predictions of the traditional regression analysis method predictions on the other. As following:

##### **4.4.1 Comparison of MAD Results:**

Table (4-5) contains the results of the MAD scale for both the traditional regression prediction method and the artificial neural network predictions for the three unknown observations of the dependent variable:

Table (5-6) MAD values for traditional method and neural network predictions

| Reg    | NN     | Pred. horizon |
|--------|--------|---------------|
| 0.0817 | 0.0515 | Z1            |
| 0.0419 | 0.0117 | Z2            |
| 0.0174 | 0.0128 | Z3            |
| 0.0470 | 0.0253 | Average       |

From the previous table, we note that the value of MAD for the neural network predictions of the three unknown observations of the dependent variable is relatively small while the values in the traditional method are large compared to the proposed method. Thus, the performance of the neural network method is better than the traditional multiple regression method. This is clear by draw by drawing the double columns of the two methods in the figure (6-3) the following:

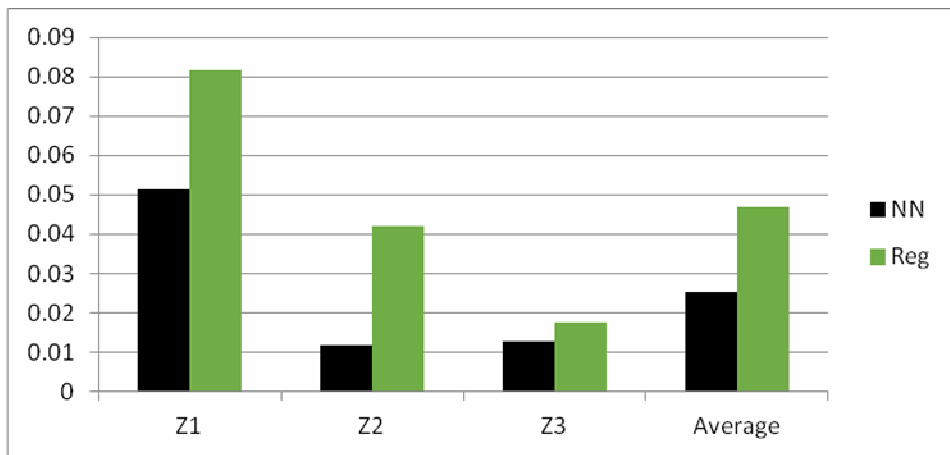


Figure (4-3) MAD values for traditional method and neural network predictions

#### 4.4.2 Comparison of results of the MSE scale:

Table (4-6) contains the results of the MSE scale for both the traditional regression prediction method and the artificial neural network predictions for the three unknown observations of the dependent variable:

Table (4-6) MSE values for traditional method and neural network predictions

| Reg    | NN     | Pred. horizon |
|--------|--------|---------------|
| 0.0067 | 0.0027 | Z1            |
| 0.0018 | 0.0001 | Z2            |
| 0.0003 | 0.0002 | Z3            |
| 0.0029 | 0.001  | Average       |

From the previous table we note that the value of MSE for the neural network predictions of the three unknown observations of the dependent variable is relatively close and that the values in the traditional method are large compared to the proposed method. Thus, the performance of the neural network method is better than the traditional multiple regression method, This is clear by draw the double columns of the two methods in the Figure (6-4) :

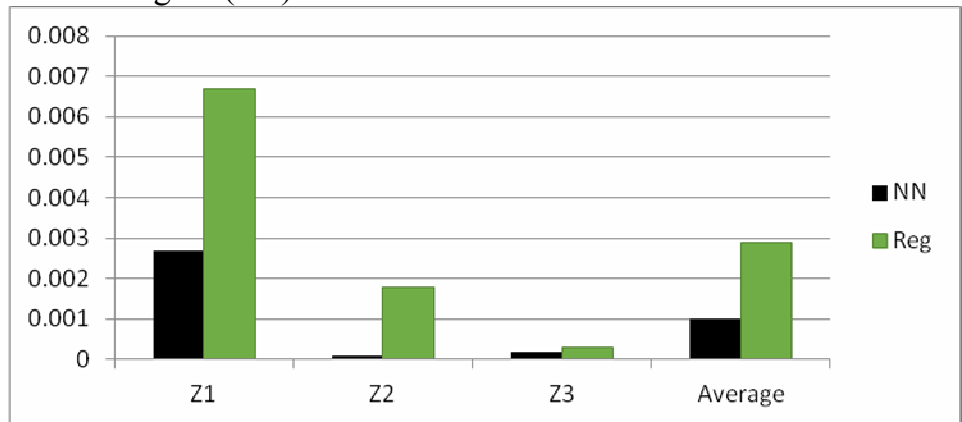


Figure (4-4) MSE values for traditional method and neural network predictions

The results of the applied study suggest that the proposed method of multiple regression analysis using neural networks is applicable in applied studies and yields better results than the traditional method of multiple regression. A complete regression analysis and prediction of the unknown values of the dependent variable can be obtained using the proposed method of regression analysis using artificial intelligence through the MATLAB program, which is designed for this purpose and attached in Bendix A .

## (5) Results

The main objectives of this research can be summarized in the following:

1. Introduction of a new method for the analysis of multiple regression using artificial neural networks, which represent one of the most important areas of artificial intelligence.

3- Conducting an applied study using realistic data.

4 - Formulation of a computer program using MATLAB software packages to compare the two methods.

5. Formulate a program using MATLAB software packages to predict regression models using the two methods

**It has become clear through the applied study that:**

a. Proposed method can be used to model data and predict its unknown values.

b - The proposed method using neural networks is better than traditional in the case of nonlinear models.

c. Performance of the proposed method is better than traditional in the case of small samples.

d - Convergence of the results of the two methods of increasing sample sizes.

## (6) Recommendations:

Some suggestions and ideas can be included for future studies as follows:

1 - Apply the proposed method of multiple regression analysis using artificial intelligence in different application fields, and this can be accomplished with the help of the MATLAB program which was designed for this purpose and attached in Appendix B to the last letter.

2 - Attention to artificial intelligence and its areas and work on its application in all areas of life.

3 - Mainstream the method of artificial neural networks on generalized regression.

4 - Circulation of the method of artificial neural networks on the double-linear regression.

5. Mainstreaming the method of artificial neural networks on nonlinear regression.



## References

- 1) Akaike .H(1973): A New Look at the Statistical Model Identification, IEEE Transactions on Automatic Control.
- 2) Abudl –Azeem, Ali (2013) “Use Nets of Artificial Neural Networks to Predict of Macroeconomic Multidimensional model in Iraq 1996 to 2007”, Journal of the Gulf economy
- 3) Al shawadfi , Gamal A.(1994) ."Bayesian Inference of ARMAX Models" , Scientific Magazine, Faculty of Commerce, Al-Azhar University ,Cairo, Egypt, Vol. 20 July 1994 .
- 4) Alhakak, Nada, AL jarrah, Nawal (2013) , Using hybrid methods in forecasting the exchange rate To the US Dollar against the Iraqi Dinar, Journal of the College of Baghdad for Economic Sciences University Issue No. thirty
- 5) Aljammal, Zakaria Yahya and Saber, Umar Wafiq, (2012), Comparison of Prediction of the Use of Neural Neural Network using Neural Networks and Regression Analysis, Iraqi Journal of Statistical Sciences.
- 6) Allam Zaki Issa, Essam Zaki Issa(2000); For Neural Networks: Engineering Architecture, Algorithms, Applications / Translation and Preparation Aleppo, Syria: Dar Shuaa for Publishing, Distribution and Science
- 7) Almorad, Aeeda Younis (2012), Comparison between classical regression and artificial neural networks in predicting the levels of research results of students of the Faculty of Physical Education. Iraqi Journal of Statistical Sciences, v21
- 8) Alsalmay, Afaf (2017) Artificial intelligence applications to retrieve information in Google, information study magazine, v19
- 9) Anders, U. (1996). Statistical model building for neural networks. In 6th AFIR Colloquium.
- 10) Arminger, G., & Enache, D. (1996). Statistical models and artificial neural networks. In Data Analysis and

- Information Systems (pp. 243-260). Springer, Berlin, Heidelberg.
- 11) Brigham, E. F., & Ehrhardt, M. C. (2013). Financial management: Theory & practice. Cengage Learning.
  - 12) Bruce D. Baker , 1998 , " A Comparison of Linear Regression and Neural Network Methods for Forecasting Educational Spending " , Department of Teaching and Leadership , 202 Bailey Hall , University of Kansas .
  - 13) Burdakov, O., Grimvall, A., Hussian, M.(2004).A Generalized PAV algorithm for monotonic regression in several variables. J.Comput.Statistics,p:23-27,prague.
  - 14) Burdakov, O., Grimvall, A., Sysoev, O. (2006). Data preordering in generalized PAV algorithm for monotonic regression. J. Comput. Math. Vol:24,p:771–790.
  - 15) Burke. L and Lgnizio. James P (1992:) “Neural networks and operations research: An overview”, Computers & Operations Research Volume 19, Issue
  - 16) Canotoni,E.;and Ronchetti,E.:(2001): “Robust inference for Generalized Linear Models “; JASA, Vol.96, No. 455, pp. 1022-1030.
  - 17) Cochran, D & Orcutt, G.H. (1949), “ Application Of Least Squares Regressions To Relationships Containing Autocorrelated Error Terms” JASA, Vol. 44.
  - 18) Demuth, H., & Beale, M. (1992). Neural network toolbox. For Use with MATLAB. The MathWorks Inc, 2000.
  - 19) Dette, H., Scheder, R. (2006). Strictly monotone and smooth nonparametric regression for two or more variables. Can. J. Stat. 34:535–561.
  - 20) Dhar, Vine. and Stein, R. (1996)."Intelligent Decision Support Methods." , the Science of Knowledge Work (Ed) , Prentice Hall Business Publishing, Nj, USA .
  - 21) Elman J.L. (1993) " Learning and development in neural networks" , the importance of starting small , Cognition ,

- International Journal of Forecasting , 48 , pp.71-99 .  
Evaluation .
- 22) Fahmi, Muhammad, (2005), Statistics Without Suffering, Concepts with Applications using SPSS, Part II, Institute of Public Administration, Riyadh, Saudi Arabia.
  - 23) Fortin, N. M., & Lemieux, T. (1997). Institutional changes and rising wage inequality: is there a linkage?. Journal of Economic Perspectives, 11(2), 75-96.
  - 24) Hardle, W. (1990)." Applied Nonparametric Regression". Econometric Society Monograph Series 19, Cambridge University Press. ISBN 0-521-42950-1 (333 p).
  - 25) Maier, H. R., & Dandy, G. C. (2000). Neural networks for the prediction and forecasting of water resources variables: a review of modelling issues and applications. Environmental modelling & software, 15(1), 101-124.
  - 26) Masters, T., (1993) "Practical neural network recipes in c++", Academic Press, San Diego, California.
  - 27) McCulloch, W., S., and Pitts, W., (1943) "A logical calculus of the ideas immanent in nervous activity", Bulletin of Mathematical Biophysics, 5, 115-137.
  - 28) Mostafa, M. M. (2004). Forecasting the Suez Canal traffic: a neural network analysis. Maritime Policy & Management, 31(2), 139-156.
  - 29) Negnevitsky, M. (2005). Artificial intelligence: a guide to intelligent systems. Pearson Education.
  - 30) Rashad, Nadwa Khazal, Zaki, . Azza Hazem (2010): Use artificial neural networks in analysis. Multiple linear regression. Rashad, Journal of the development of Rafidain No. 99 Volume 32
  - 31) Rumelhart, D. E., Hinton, G. E., and Williams, R. J. (1986). "Learning representations by back-propagating errors". Nature, 323, 533--536.
  - 32) Rutkowski, L. , ( 2004 ) , " Generalized Regression Neural Networks in Time-Varying Environment " IEEE , Transactions on Neural Networks , Vol .15 , No3 , pp.576-596 .

**Appendix (A)**  
**MATLAB program for applied study using artificial intelligence**  
**and comparing it with the traditional method of multiple**  
**regression**

% Here a feedforward network is trained and regression performed  
on its targets and outputs.ch6

```
x= [0.438444128 0.665499966 0.948187239 0.170759121 0.583731637  
0.041002866 0.349561836 0.650865937 0.712767403 0.580205261 0.568834554  
0.683338832 0.036740893 0.901301436 0.159195605 0.76917527 0.327059315  
1.864607956 0.895849862 0.858721725 0.817339249 0.736617573 0.250853464  
0.511748172 0.780954909 0.900262957 0.098413173 0.995457295 0.196371346  
0.697708531 0.70625111 0.988557829 0.035204727 0.212188105 0.299535125  
0.200606851 0.200606851 0.831275598 0.67164998 0.813050579 0.528550715  
0.080369688 0.761454122 0.716039181 0.640381813 0.256619818 0.251733566  
0.553823595 0.453169061 0.444030087 0.333377265 0.260711953 0.689817609  
0.314971852 0.830501159 0.950905208 0.526803616 0.806950894 0.889189959  
0.619970046;  
0.580675712 0.823783041 0.428871018 0.294101969 0.454395922  
0.160740528 0.018827119 0.448871759 0.300511377 0.295824677 0.173980465  
0.46054705 0.246774214 0.03981235 0.233266403 0.424496792 0.499836263  
0.976131837 0.152633956 1.512813511 0.466825289 0.625569529 0.018720928  
0.286876545 0.559895291 0.02924919 0.4665979 0.480221609 0.200690522  
0.336082701 0.651912985 0.391549379 0.195075439 0.306780422 0.297784901  
0.32528635 0.32528635 0.320431611 0.71345323 0.755330085 0.690835184  
0.055582373 0.269988598 0.691641883 0.215682795 0.188542822 0.214251839  
0.704685844 0.427854005 0.42211779 0.497112965 0.253112627 0.519802002  
0.439230504 0.453729741 0.257793459 0.435788104 0.083653521 0.634611329  
0.606461299;  
0.466316665 0.284232577 1.573652109 0.526526933 0.835742195  
0.142722968 11.82034055 0.351828665 1.84097958 1.850567406 1.795337483  
1.161546685 0.081420587 20.18042219 0.405586363 1.051026047 0.539163705  
1.636876219 4.356403158 0.514957761 1.602538339 0.729178723 9.319611884  
0.204256984 1.204772454 14.91044231 0.122786397 1.470345713 0.454922953  
0.812428689 1.082594031 0.967877282 0.180467244 0.312794828 0.506889829  
0.385944424 0.385944424 1.027144993 0.825338752 1.003235838 0.126311933  
0.663213894 2.56767641 0.710938641 2.011612526 0.98002539 0.94279231  
0.219304328 0.351013998 0.457998039 0.165287801 0.590099054 0.988279512  
0.225484947 1.523069899 2.686089909 0.574067733 1.421465104 1.162921711  
0.627203696;  
0.545195689 0.26572505 0.051487798 0.054925866 0.193936492 0.07233428  
0.22406315 0.346454561 0.101475784 0.304735958 0.318809733 0.044428827  
0.63998292 0.382344065 0.582164817 0.165213839 0.040989668 0.025476716  
0.001220368 0.139957519 0.018398014 0.119125212 0.67167666 0.004414585  
0.002999963 0.005319595 0.01455258 0.004542705 0.003935637 0.002616703  
0.008107607 0.011442171 0.928811917 0.738724762 0.666779783 0.782641935  
0.782641935 0.134853807 0.310613117 0.109338144 0.422526196 0.322395207  
0.059924833 0.256511897 0.275036044 0.737283877 0.06040775 0.375295696
```

```

0.323270719 0.554536388 0.541893824 0.511400205 0.276161794 0.567844204
0.026505044 0.038757511 0.316271014 0.193049106 0.047515648 0.113660696;
-0.188883178 0.062143278 9.153817779 8.600508793 7.967968776 -
0.408257132 -0.233046956 0.152021132 -0.011846792 0.054854606
0.075378649 -0.038708476 -0.734515283 0.100282289 0.180205696
0.209815138 0.232880304 -0.330481349 -0.067319739 0.294995544
1.612011252 0.18531765 -0.871163917 7.784591744 1.805036483 1.350406596
0.047313177 -0.160923337 0.306474741 -0.319522516 0.076979923 -
0.583923523 -0.96568688 -0.222131872 -0.552912425 -0.548392603 -
0.548392603 0.472326299 0.017594855 1.402181608 0.673882755 0.389948087
0.224330829 0.239999608 0.19968434 7.671125873 4.365826752 -2.923599112
-0.170401585 0.005145853 -0.237590797 -0.182703437 0.558124797
1.326198298 -0.142837043 0.633746366 -0.050955159 0.049713775 -
0.568587376 0.323151726];
y = [-0.022846228 -0.2384455 0.166912887 0.001299114 -0.230477166
0.015023309 -0.073024784 0.076502189 0.190723701 0.091241932
0.155889269 0.031341764 -0.774903459 0.242169325 -0.015124253
0.068555061 0.001266267 0.227235832 0.12472979 -0.049588568 0.006452626
0.102433709 -0.049419449 0.015879576 0.043329588 0.088021689
0.030074525 -0.066888326 0.016636563 0.091872806 0.019174683
0.024884002 -0.113911234 -0.047096367 0.007766586 -0.220237081 -
0.220237081 0.039597357 0.103191055 0.087801425 -0.0861986 0.025007276
-0.019459641 0.175138767 0.119568624 0.064491558 0.304888938 -
0.393081806 0.014050309 -0.478736286 -0.218826458 -0.017130881
0.139745145 -0.066942781 -0.418057358 0.083422982 0.06667938
0.071690883 0.031844366 0.007347079];
n=size(y'); h= 3 ;
% 3... .. comparison and testing phase ... ..
% Compute regression coefficients for a linear model with an interaction term:
net = fitnet(1,'trainingd');
net.trainParam.showWindow =0 ;
[net,tr]= train(net, x,y);
y0=sim(net ,x);
y1=y0(n-h+1:n);
[aa b cc2] = regress(y' , x');
ytrue = y(n-h+1:n);
c1 = y1 - ytrue;
c2 = cc2(n-h+1:n);
s01 = abs(c1); s02 = (s01).^2;
r01 = abs(c2') ; r02 = (r01).^2;
s01 < r01 ; ss =ans
% 3... .. Results ... ..
disp('..... RESULTS..... ')
disp('MAE RESULTS ')
MAE = [ s01 ;r01]
disp('MSE RESULTS ')
MSE = [s02 ;r02]
disp('MAP RATIOS RESULTS ')

```

MAEP = sum(ss)\*100

### Appendix C

Applied study data representing the financial statements of 60 companies listed on the Egyptian Stock Exchange

| م  | X <sub>1</sub> | X <sub>2</sub> | X <sub>3</sub> | X <sub>4</sub> | X <sub>5</sub> | Y      |
|----|----------------|----------------|----------------|----------------|----------------|--------|
| 1  | 0.438444       | 0.580676       | 0.4663         | 0.545          | -0.18888       | -0.023 |
| 2  | 0.6655         | 0.823783       | 0.2842         | 0.266          | 0.062143       | -0.238 |
| 3  | 0.948187       | 0.428871       | 1.5737         | 0.051          | 9.154          | 0.167  |
| 4  | 0.170759       | 0.294102       | 0.5265         | 0.055          | 8.601          | 0.001  |
| 5  | 0.583732       | 0.454396       | 0.8357         | 0.194          | 7.968          | -0.23  |
| 6  | 0.041003       | 0.160741       | 0.1427         | 0.072          | -0.40826       | 0.015  |
| 7  | 0.349562       | 0.018827       | 11.82          | 0.224          | -0.23305       | -0.073 |
| 8  | 0.650866       | 0.448872       | 0.3518         | 0.346          | 0.152021       | 0.077  |
| 9  | 0.712767       | 0.300511       | 1.841          | 0.101          | -0.01185       | 0.191  |
| 10 | 0.580205       | 0.295825       | 1.8506         | 0.305          | 0.054855       | 0.091  |
| 11 | 0.568835       | 0.17398        | 1.7953         | 0.319          | 0.075379       | 0.156  |
| 12 | 0.683339       | 0.460547       | 1.1615         | 0.044          | -0.03871       | 0.031  |
| 13 | 0.036741       | 0.246774       | 0.0814         | 0.64           | -0.73452       | -0.775 |
| 14 | 0.901301       | 0.039812       | 20.18          | 0.382          | 0.100282       | 0.242  |
| 15 | 0.159196       | 0.233266       | 0.4056         | 0.582          | 0.180206       | -0.015 |
| 16 | 0.769175       | 0.424497       | 1.051          | 0.165          | 0.209815       | 0.069  |
| 17 | 0.327059       | 0.499836       | 0.5392         | 0.041          | 0.23288        | 0.001  |
| 18 | 1.864608       | 0.976132       | 1.6369         | 0.025          | -0.33048       | 0.227  |
| 19 | 0.89585        | 0.152634       | 4.3564         | 0.001          | -0.06732       | 0.125  |
| 20 | 0.858722       | 1.512814       | 0.515          | 0.14           | 0.294996       | -0.05  |
| 21 | 0.817339       | 0.466825       | 1.6025         | 0.018          | 1.612011       | 0.006  |
| 22 | 0.736618       | 0.62557        | 0.7292         | 0.119          | 0.185318       | 0.102  |
| 23 | 0.250853       | 0.018721       | 9.3196         | 0.672          | -0.87116       | -0.049 |
| 24 | 0.511748       | 0.286877       | 0.2043         | 0.004          | 7.784592       | 0.016  |
| 25 | 0.780955       | 0.559895       | 1.2048         | 0.003          | 1.805036       | 0.043  |
| 26 | 0.900263       | 0.029249       | 14.91          | 0.005          | 1.350407       | 0.088  |
| 27 | 0.098413       | 0.466598       | 0.1228         | 0.015          | 0.047313       | 0.03   |
| 28 | 0.995457       | 0.480222       | 1.4703         | 0.005          | -0.16092       | -0.067 |
| 29 | 0.196371       | 0.200691       | 0.4549         | 0.004          | 0.306475       | 0.017  |
| 30 | 0.697709       | 0.336083       | 0.8124         | 0.003          | -0.31952       | 0.092  |
| 31 | 0.706251       | 0.651913       | 1.0826         | 0.008          | 0.07698        | 0.019  |
| 32 | 0.988558       | 0.391549       | 0.9679         | 0.011          | -0.58392       | 0.025  |
| 33 | 0.035205       | 0.195075       | 0.1805         | 0.929          | -0.96569       | -0.114 |
| 34 | 0.212188       | 0.30678        | 0.3128         | 0.739          | -0.22213       | -0.047 |
| 35 | 0.299535       | 0.297785       | 0.5069         | 0.667          | -0.55291       | 0.008  |
| 36 | 0.200607       | 0.325286       | 0.3859         | 0.783          | -0.54839       | -0.22  |

| م  | X <sub>1</sub> | X <sub>2</sub> | X <sub>3</sub> | X <sub>4</sub> | X <sub>5</sub> | Y      |
|----|----------------|----------------|----------------|----------------|----------------|--------|
| 37 | 0.200607       | 0.325286       | 0.3859         | 0.783          | -0.54839       | -0.22  |
| 38 | 0.831276       | 0.320432       | 1.0271         | 0.135          | 0.472326       | 0.04   |
| 39 | 0.67165        | 0.713453       | 0.8253         | 0.311          | 0.017595       | 0.103  |
| 40 | 0.813051       | 0.75533        | 1.0032         | 0.109          | 1.402182       | 0.088  |
| 41 | 0.528551       | 0.690835       | 0.1263         | 0.423          | 0.673883       | -0.086 |
| 42 | 0.08037        | 0.055582       | 0.6632         | 0.322          | 0.389948       | 0.025  |
| 43 | 0.761454       | 0.269989       | 2.5677         | 0.06           | 0.224331       | -0.019 |
| 44 | 0.716039       | 0.691642       | 0.7109         | 0.257          | 0.24           | 0.175  |
| 45 | 0.640382       | 0.215683       | 2.0116         | 0.275          | 0.199684       | 0.12   |
| 46 | 0.25662        | 0.188543       | 0.98           | 0.737          | 7.671          | 0.064  |
| 47 | 0.251734       | 0.214252       | 0.9428         | 0.06           | 4.365827       | 0.305  |
| 48 | 0.553824       | 0.704686       | 0.2193         | 0.375          | -2.9236        | -0.393 |
| 49 | 0.453169       | 0.427854       | 0.351          | 0.323          | -0.1704        | 0.014  |
| 50 | 0.44403        | 0.422118       | 0.458          | 0.555          | 0.005146       | -0.479 |
| 51 | 0.3333773      | 0.497113       | 0.1653         | 0.5419         | -0.237591      | -0.219 |
| 52 | 0.260712       | 0.2531126      | 0.5901         | 0.5114         | -0.182703      | -0.017 |
| 53 | 0.6898176      | 0.519802       | 0.9883         | 0.2762         | 0.5581248      | 0.1397 |
| 54 | 0.3149719      | 0.4392305      | 0.2255         | 0.5678         | 1.3261983      | -0.067 |
| 55 | 0.8305012      | 0.4537297      | 1.5231         | 0.0265         | -0.142837      | -0.418 |
| 56 | 0.9509052      | 0.2577935      | 2.6861         | 0.0388         | 0.6337464      | 0.0834 |
| 57 | 0.5268036      | 0.4357881      | 0.5741         | 0.3163         | -0.050955      | 0.0667 |
| 58 | 0.8069509      | 0.0836535      | 1.4215         | 0.193          | 0.0497138      | 0.0717 |
| 59 | 0.88919        | 0.6346113      | 1.1629         | 0.0475         | -0.568587      | 0.0318 |
| 60 | 0.61997        | 0.6064613      | 0.6272         | 0.1137         | 0.3231517      | 0.0073 |