International Journal of Environmental Sciences
ISSN: 2229-7359
Vol. 11 No. 12s,2025
https://theaspd.com/index.php

Forecasting Future Water Demand In Baghdad: A Comparison Between The Two Methods Multiple Regression And MLP-Neural Network Method

[1]**Hassan Saad Abbas**, [2]**Ahmed Shaker Mahmoud**, [3]**Baraq Sobhi Kamel**
[1,2,3]Department of Investment Management and Business Economics, College of Business Economics, Al-Nahrain University, Baghdad, Iraq
[1]cdkl2d1x58@gmail.com, [2]ahmedshakerma@gmail.com, [3]dr.barraq@nahrainuniv.edu.iq

Abstract
The research aims to predict the quantities of water needed in the city of Baghdad for a period ofThe next 24 months were forecasted using two models: Multiple Regression and MLP-Neural Network. This model relies on the city's water consumption data obtained from the Baghdad Municipality, specifically the Baghdad Water Department, for the period from January 2014 to May 2024, a total of 125 months. It also takes into account some influencing factors such as population, maximum temperature, and monthly humidity, to estimate the quantities Baghdad will need in the future. This results in the best and most accurate model for predicting water consumption in Baghdad.
The research concluded that the best model suitable for predicting water consumption in the city of Baghdad among the models used is:MLP(3,3,1) in terms of accuracy measures and the average absolute error rate (MAPE=5%), while the average absolute error rate for the Multiple Regression model (MAPE=6%). The research also found that there is a correlation and influence between some influencing factors (population and maximum temperature) and water consumption, as these factors were able to explain 0.703% of the changes in water consumption in Baghdad. Finally, the research recommends relying on artificial intelligence models to support water planning decisions, and periodically updating water consumption data to enhance the accuracy of future predaictions. The research also recommends the necessity of taking the necessary measures to limit water consumption in the city, through pricing, awareness, education, intermittent supplies, and other methods that preserve water resources, address some influencing factors, and achieve water sustainability.
Keywords: *Water demand forecasting,Multiple Regression, MLP-Neural Network.*

1. INTRODUCTION

Freshwater is a vital resource that is scarce and threatened with decline, and is essential for the preservation and sustainability of life. Achieving balance between the ecosystem and other systems, and protecting it from depletion, is considered one of the basic goals that sustainable development seeks to achieve.and aI,. 2024:1). (Saleh) Water is one of the most depleted natural resources. Since Iraq is one of the Middle East's most water-scarce countries, and is described as a semi-arid region, Iraq is vulnerable to a major future water crisis. The causes of this crisis are climate change, population growth, and the declining share of Iraq's water imports due to the policies of neighboring countries and others. Since Baghdad is the largest city in Iraq in terms of population, with an estimated population of 8,640,484 in 2023, we also find that water demand in the city is extremely high, based on data from the Iraqi Ministry of Water Resources. Therefore, the need to forecast water demand in the city is increasing. Forecasting future water demand in Baghdad plays an important role in determining the city's future water needs. This is achieved through the use of various forecasting models such as Multiple Regression and MLP-Neural Networks, which help decision-makers determine the city's future water needs and make decisions and take various actions that contribute to reducing water consumption in Baghdad and achieving water resource sustainability.

2. RESEARCH METHODOLOGY

2.1 The Research Problem

In light of the increasing demand for water in the city of Baghdad with the increase in the city's population and other influencing factors such as weather factors, versus a decrease in Iraq's share of water resources controlled by neighboring countries, this situation requires an accurate forecast of monthly water consumption based on some influencing factors such as (population, maximum temperature, and humidity), and accordingly the research question can be formulated as follows: How can each ofMultiple Regression and MLP-NN models predict water consumption in Baghdad for the next 24 months, based on several influencing factors. Which of the two models provides higher prediction accuracy?

2.2 Research Objectives

There are several objectives that the research seeks to achieve, including:

1. Forecasting water consumption in Baghdad city for 24 months using Multiple Regression and MLP-NN models.
2. Analyze the relationship between some influencing factors (population, maximum temperature, and humidity) and water consumption in the city of Baghdad, and explain the extent of their influence.

.3 Comparing the results of prediction models and identifying the most accurate model in predicting the future water needs in the city. Accuracy measures, namely Mean Squared Error (MSE), Mean Absolute Error (MAD), and Mean Absolute Percentage Error (MAPE), were used to determine the best prediction model.

4. Providing recommendations based on the results of prediction models to support decision-making in managing future water consumption in the city of Baghdad.

3. The Theoretical Aspect

3.1 Forecasting Water Consumption

Forecasting is defined as: the art and science of predicting future events.(Al-Zuhairi, 2023: 53). Forecasting is a primary goal of analyzing time series data for various phenomena. There are numerous statistical methods and models used to analyze data over a specific period of time, analyze the causes of these changes, understand their nature, and then predict the future of that phenomenon. These statistical models vary in their predictive ability, as the accuracy of their results depends on several factors, such as the quality of the data upon which the study is based and the changes that affect this data (Hamid & Abushaba, 2020: 42). Forecasting water consumption is also essential in any country, especially those that lack water resources and where consumption is increasing. Forecasting helps relevant authorities understand future patterns of consumption and develop methods that help reduce excessive consumption and achieve water sustainability (Younis, 2016: 1). Water consumption forecasting is defined as: The process of estimating future water consumption based on historical consumption patterns, using forecasting models that help specialists plan and manage water. (SantAna, 2021: 2) Many urban areas face water shortages, so these areas need accurate forecasting of future water consumption due to its effective role in managing water economies and determining future needs. This is done through the use of various forecasting models (Chen, et al, 2017: 2).

3.2 The importance of Forecasting Water Consumption

The importance of forecasting water consumption for the relevant authorities is highlighted in the following:(Lima, et al,. 2023:2) (Zheng, et al,. 2022:1) (Gil-Gamboa, et al,. 2024:1):

1. Forecasting plays a key role in planning, organizing, and controlling water consumption, which can be short, medium, or long term.
2. It helps water authorities improve water distribution and achieve a balance between water supply and demand.

.3 Determine future water needs.

.4 Assessing the factors affecting water consumption and implementing rapid water management measures.

.5 Avoid future water crises and achieve water sustainability.

6. Accurate forecasting helps determine the future trend of consumption, which helps relevant authorities take measures to reduce consumption.

3.3 Forecasting methods Forecasting Models

Forecasting methods are divided into two types:(Montgomery, et al,.2015:3-4) (Saeed & Aqla, 2018:300):

1.3.3 Qualitative approaches or methods

These methods rely on experience and estimates rather than on quantitative data for the phenomenon, and are usually used when there is no historical data for the phenomenon under investigation. These methods include intuition and experience, the Delphi model, the expert panel, and others.

2.3.3 Quantitative methods or methods

Quantitative methods are divided into two types:

- Time series models: such as:ARIMA,
- Causal relationship models: such as:**Multiple Regression**

In this research, we focus on causal models to predict water consumption quantities, and analyze the influence of external factors such as (population, weather factors).

Multiple regression modelMultiple Regression

Regression analysis is based on two basic models: simple and multiple. The simple model is based on the relationship between the dependent variable and the(Y) and one independent variable (X), while the multiple model is based on the relationship between the dependent variable (Y) and more than one independent variable (), and the purpose of both models is prediction (Hamad, 2023:191).$X_1, X_2 \dots . X_n$

The multiple regression equation is as follows:

$$\boldsymbol{Y_t = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + .. + \beta_n X_n + e_t} \quad \dots (1)$$

Whereas:

Y_t:Represents the phenomenon being studied (dependent variable).

$\beta_1, \beta_2, \dots, \beta_n$:Represents the model parameters.

$X_1, X_2, .., X_n$The independent variables represent the variables that affect the dependent variable.

e_t:Values of the random variable.

2 Artificial Neural Networks Artificial Neural Networks (ANNs)

Artificial neural networks are defined as a computational technique that mimics the way the human brain performs a specific task. Neural networks consist of several processing units, defined as neurons or nodes, organized into three levels (layers): the input level, the hidden level consisting of one or more hidden layers, and the output level. Nodes at each level are connected to the next level through connection strengths called weights, which store the knowledge gained from training the network.(Al-Sharoot & Abdon, 2017: 3) (Khalil & Hamad, 2023: 4266). Artificial neural networks (ANNs) have several advantages, such as capturing correlations or detecting consistency within a set of patterns and being used as an approximation of a random function. ANNs also take samples of data rather than the entire data to obtain solutions, which saves money and time (Younis, 2016: 6).

1.2 Components of Artificial Neural Networks (ANNs)

An artificial neural network consists of a group of processing units called neurons, which are very similar to the neurons in the human brain. These units are connected to each other in the form of connections called a neural network. The general form of these neural networks consists of the following components: (Ubah, et al, 2021:2) (Awad & Alkelani, 2019:43):

.1 Input Level: It represents the layer through which the network is fed with data from the outside.

.2 Hidden Level: It lies between the input layer and the output layer, and processes the data before passing it to the output.

.3 Output Level: It represents the final results produced by the network after processing.

4. Weights: These are connections between the layers of the neural network, and they transfer weighted data.

2.14.1.2 Multilayer Perceptual Neural Network (MLP—NN)

It is one of the most widely used neural networks in prediction, which uses a feed-forward network with one or more hidden layers. The structure consists of (MLP-NN) consists of three layers: input, hidden, and output. There may be more than one hidden layer connected by non-cyclic links. The nodes in the different layers are also known as processing elements. The three layers are linked by a force called a weight. There are two sets of weights:

.1. Weights of the hidden input layer (wj,i).

.2 hidden layer output weights (wk,j).

These weights give the neural network great flexibility to adapt freely to the data. Multi-layer neural networks have the potential to solve more complex problems because they contain one or more hidden layers at each hidden level.(Al-Sharoot & Abdoon, 2017:3). Diagram (1) shows the structure of multi-layer neural networks.

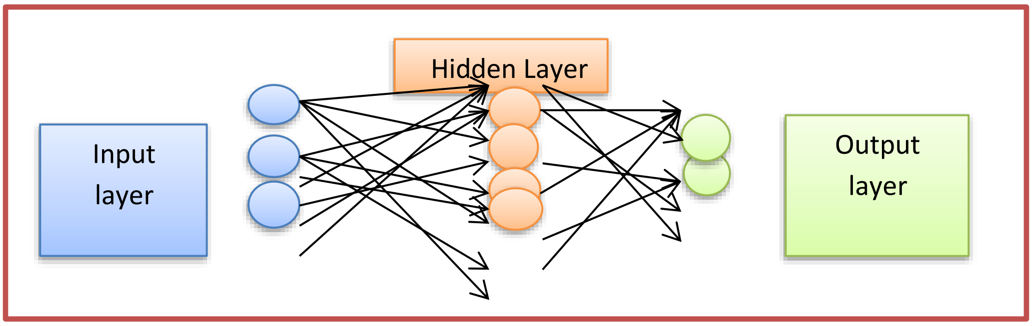


Figure (1): Structure of the Multilayer Perceptron Neural Network (MLP-NN).

Source: Prepared by the researcher based on(Agrawal & Adhikari, 2013:26).

•The model outputs are usually calculated using the following mathematical formula:(Kassem, et al, 2020:33):

(2) ..$Y_t = F_0\left[\sum_{j=1}^{q}(W_{k,j} * F_n(\sum_{i=1}^{p}(W_{j,i} + X_i) + b_j)\ ) + b_k\ \right]$

Whereas:

Y_t :Represents the output variable.

X_i :Represents the input variable.

p,q: represent the number of input and hidden nodes respectively.

W_{ji} :Represents the weights of the hidden inputs.

W_{kj} :Hidden output weights.

b_j :It is the bias of the hidden layer.

b_k It is the bias of the output layer.

$F_n F_0$ Activation functions for the hidden layer and the output layer.

3.14.1.2 Steps for forecasting using artificial neural networks**Forecasting Steps Using (ANNs)**

A prediction model can be built using artificial neural networks through a set of steps, which are:(Salaheddine, et al, 2024:95) (Al-Badiri & Al-Kaabi, 2024:98-99) (Awad & Alkelani, 2019:43):

The first step: It consists of identifying the variables (inputs), which are the dependent variable and the independent variables.Y_i(..), the number of independent variables does not necessarily have to be large in neural networks because too many variables can confuse the operation of the neural network. Source$X_1, X_2 X_n$

Step two: Data fragmentation

The data is divided into two groups (training and testing). Training helps the network learn, while testing measures the accuracy of prediction.

Step three: Determine the structure of the network model (neural network architecture)

The definition of neural network architecture includes the following:

- Number of input nodes, usually equal to the number of inputs.
- The number of hidden layers depends on the error value in the network.
- The number of hidden ganglia depends on the experience.
- The output layer is usually equal to one.

International Journal of Environmental Sciences
ISSN: 2229-7359
Vol. 11 No. 12s,2025
https://theaspd.com/index.php

Step Four: Network training
The network is trained by selecting the appropriate training algorithm. Studies indicate that the best algorithm for training the network is the backpropagation algorithm in the field of prediction.
Step five: Evaluation criteria
There must be a standard to measure the accuracy of prediction, and studies indicate the use of the relative error standard.(Relative Error) which expresses the error as a percentage of the actual values.
Step 6: Implementation
It is one of the most important steps in prediction using neural networks, as it tests the network's adaptive capacity, with the possibility of retraining and achieving the least square error when the data changes.
4. The Applied Aspect
4.1 Research Data
Data on monthly water consumption in Baghdad were obtained from the Baghdad Municipality specifically (Baghdad Water Department). Data on the population of Baghdad were also obtained from the Central Statistical Organization, and data on maximum temperatures and humidity were obtained from the General Authority for Meteorology for the period from January2014 AD until May 2024 AD, for a period of (125) months, as shown in Table (1).

2014					**2015**				
Month	Water Consumption (Thousands)m^3	Population of Baghdad City	Maximum Temperatures ()C°	Relative Humidity (%)	Month	Water Consumption (Thousands)m^3	Population of Baghdad City	Maximum Temperatures ()C°	Relati Humi (%)
1	66611	6524783	16.4	81	1	77084	6715656	17.4	
2	59966	6540481	19.1	58	2	70760	6731812	20.6	53
3	70290	6556212	25.6	52	3	78308	6748008	25.4	43
4	71007	6571974	31.8	39	4	77325	6764242	30.8	31
5	78732	6587794	37.9	26	5	80130	6780515	38.3	24
6	82156	6603643	41.7	20	6	86372	6716827	42.1	21
7	82476	6619530	44	22	7	88155	6813178	46.1	18
8	83258	6635455	25.1	22	8	88210	6829569	45.8	23
9	82869	6651418	31.6	28	9	88528	6,846,000	42.6	26
10	82205	6667420	32.9	41	10	86213	6862469	34.3	43
11	77159	6683460	22.8	57	11	78536	6878961	25.3	62
12	81534	6699539	19.6	65	12	76155	6895529	17.1	69
2016					**2017**				
Month	Water Consumption (Thousands)m^3	Population of Baghdad City	Maximum Temperatures ()C°	Relative Humidity (%)	Month	Water Consumption (Thousands)m^3	Population of Baghdad City	Maximum Temperatures ()C°	Relati Humi (%)
1	75650	6914146	16.2	65	1	82671.5	7134312	16.1	70
2	73477	6929588	21.5	62	2	78442	7150237	18	50
3	79515	6945064	25.1	50	3	82681.5	7166262	23.8	56
4	79736	6960575	32.3	41	4	87125	7182087	30.4	41
5	88859	6976120	37.2	32	5	92999.3	7198012	38.1	26
6	93180	6991700	42.4	24	6	95149	7213938	42.7	21

7	95251	7007314	45.5	21	7	102458.5	7229863	47.5	16
8	95005	7022165	46.7	23	8	110904	7245788	47.4	18
9	94644	7038639	40.1	29	9	109678.8	7261712	43.5	22
10	99883	7054787	35.3	36	10	107606.8	7277637	33.7	33
11	90349	7070113	24.7	38	11	101172.5	7293562	25.4	48
12	87321	7086139	15.8	66	12	90573.5	7309369	21	53
2018					**2019**				
Month	Water Consumption (Thousands) m^3	Population of Baghdad City	Maximum Temperatures ()C°	Relative Humidity (%)	Month	Water Consumption (Thousands) m^3	Population of Baghdad City	Maximum Temperatures ()C°	Relati Humi (%)
1	87127.5	7341707	18.9	53	1	91617	7553373	16.9	70
2	81047.9	7357892	20.6	64	2	81627.6	7569908	19.1	68
3	98032.1	7374076	28	41	3	94962.7	7586444	22	58
4	99705.2	7390261	29.6	48	4	94359	7602980	28	49
5	102463.5	7406446	35.7	38	5	104508.1	7619515	38.7	27
6	103014.2	7422631	42	23	6	104246.2	7636050	44.8	21
7	104457.7	7438816	44.1	21	7	111080	7652586	44.1	21
8	101661.8	7,455,000	43.6	25	8	110278	7669121	45.3	24
9	102454.2	7471185	42.1	27	9	110278	7685656	41.4	28
10	102115.4	7487370	34.2	41	10	105929	7702191	35	46
11	96250.1	7503555	22	78	11	100098	7718727	26	49
12	95883.7	7519741	17.8	82	12	98406	7735263	19.1	73
2020					**2021**				
Month	Water Consumption (Thousands) m^3	Population of Baghdad City	Maximum Temperatures ()C°	Relative Humidity (%)	Month	Water Consumption (Thousands) m^3	Population of Baghdad City	Maximum Temperatures ()C°	Relati Humi (%)
1	99243	7770427	17.4	69	1	102974	7985164	19.7	55
2	89282	7787497	19	58	2	96057	8002602	21.2	59
3	104361	7804568	24.8	55	3	104329	8020039	25.4	41
4	99505	7821638	30.8	40	4	109237	8037477	33.7	31
5	108349	7838708	38.2	27	5	112917	8054914	40.7	24
6	108084	7855778	42.4	23	6	111963	8072352	43.1	21
7	111427	7872848	47.4	21	7	115947	8089790	46.6	21
8	113312	7889921	43.8	24	8	117222	8107218	46	24
9	111981	7906989	43.9	27	9	114169	8129634	40.3	29
10	109968	7924060	36.1	34	10	113894	8147082	35	37
11	104255	7941130	25.1	60	11	108226	8164530	26.8	50
12	102086	7955928	18.9	69	12	97231	8179210	19.4	59
2023					**2022**				
Month	Water Consumption (Thousands) m^3	Population of Baghdad City	Maximum Temperatures ()C°	Relative Humidity (%)	Month	Water Consumption (Thousands) m^3	Population of Baghdad City	Maximum Temperatures ()C°	Relati Humi (%)
1	97369	8214202	17.1	62	1	93147	8443726	16.3	79
2	88058	8231575	21.9	50	2	85682	8460227	19.6	50
3	100949	8248985	22.9	38	3	95617	8479157	25.3	57

4	101548	8266430	32.4	31	4	100512	8496927	30.1	43
5	111820	8283876	35.9	26	5	112828	8514735	36.6	28
6	109295	8301398	43.7	20	6	114204	8532581	41.9	35
7	112976	8318955	45.1	18	7	117217	8550464	45.8	22
8	113626	8336493	45.8	20	8	122579	8568384	46.4	24
9	110480	8354181	42	26	9	119703	8586342	42.6	26
10	111023	8371849	36.6	31	10	86664	8604338	34.2	39
11	101915	8389556	25.6	60	11	75715	8622371	26	58
12	110700	8407444	19.2	77	12	70953	8640484	20.8	71
4									
Month	Water Consumption (Thousands) m^3	Population of Baghdad City	Maximum Temperatures () C°	Relative Humidity (%)	Month	Water Consumption (Thousands) m^3	Population of Baghdad City	Maximum Temperatures () C°	Relati Humi (%)
1	105178	8678250	20.2	61	4	111851	8732917	33	35
2	102714	8696448	20.7	65	5	115709	8751220	36.5	32
3	110299	8714674	24.6	48					

table(1) Monthly water consumption values in the city of Baghdad calculated (Thousand) with data on population, maximum temperatures and humidity for the period from January 2014 AD to May 2024 AD.m^3

2.4 Forecasting future water demand in Baghdad

Monthly water consumption in Baghdad is predicted using some factors affecting consumption, namely (Baghdad population, maximum temperatures, and humidity). In the first stage, Pearson's correlation coefficient is used.(Pearson) to measure the correlation of each of the influencing factors with monthly water consumption in the city of Baghdad. In the second stage, the influencing factors with correlation will be predicted to be used as inputs later in building models to predict water consumption in the city of Baghdad, according to the time series characteristics of each factor. In the final stage, water consumption will be predicted using a multiple regression model.(Multiple Regression), multi-layer neural networkMLP- Neural Network.

1.2.4 Testing the association between some influencing factors and water consumption

At this stage, Pearson's correlation coefficient determines the relationship between the variables. We test the correlation between the dependent variable, which represents monthly water consumption in Baghdad, and the independent variables (Baghdad's population, maximum temperatures, and monthly humidity). We test the main hypothesis, which states that there is no statistically significant relationship between some influencing factors and water consumption in Baghdad. We also test its sub-hypotheses, which are:

1. There is no statistically significant relationship between the population factor of Baghdad and water consumption in Baghdad.
2. There is no statistically significant relationship between the maximum temperature factor and water consumption in Baghdad.
3. There is no statistically significant relationship between the monthly humidity factor and water consumption in Baghdad.

Table (2) Correlation coefficients between the dependent variable (monthly water consumption) and the independent variables (population, maximum temperatures, and humidity).

The worker	Water Consumption (Thousands) m^3		Correlation type	Correlation strength
Population of Baghdad City	P	0.684**	My package	Medium to strong
	Sig	0.000		
	P	0.518**	My package	Medium

Maximum Temperatures ()C°	Sig	0.000		
Relative Humidity (%)	P	-0.438**	reverse	Medium
	Sig	0.000		

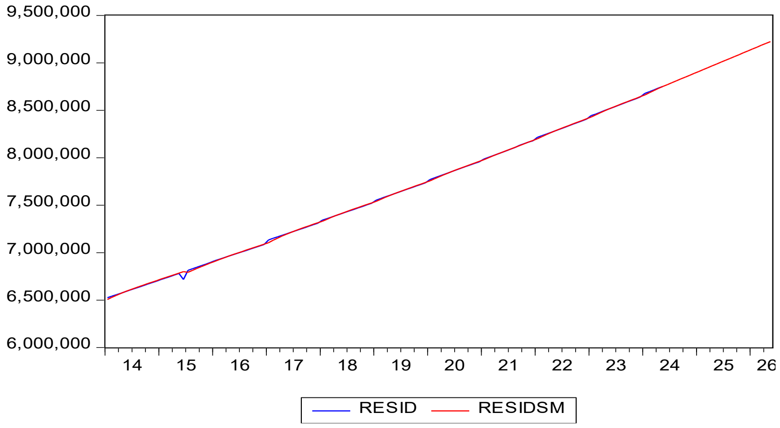
Note: **Significant correlation at a significance level of 0.01. *Significant correlation at a significance level of 0.05. Source: Prepared by the researcher based on the outputs of SPSS Version 26

Through the results of the table (2) We answer the correlation hypotheses as follows:

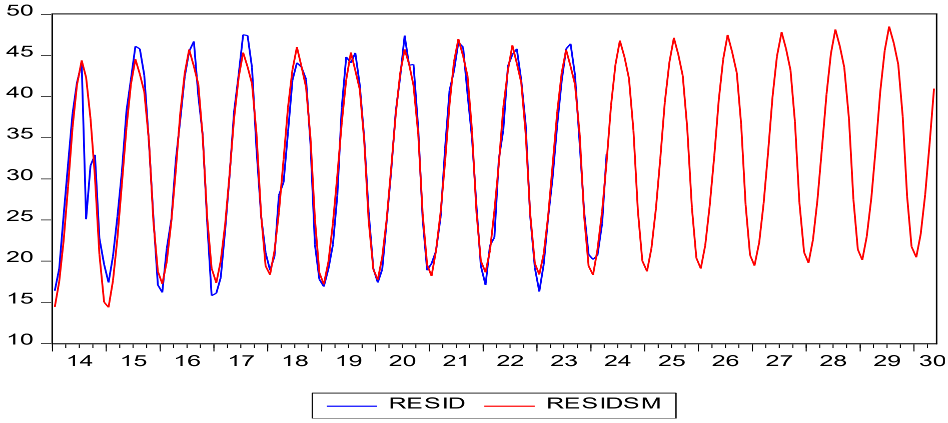
1. There is a direct positive correlation between the population of Baghdad and water consumption, where the correlation coefficient reached (0.684), and the p-value was (0.000), which indicates that the relationship is statistically significant at the (0.01) level. Therefore, the alternative hypothesis is accepted and the null hypothesis, which states that(There is no statistically significant relationship between the population factor of Baghdad and water consumption. In the city of Baghdad.
2. There is a direct positive correlation between the maximum temperature factor and water consumption in the city of Baghdad, where the correlation coefficient reached (0.518), and the p-value was (0.000), which indicates that the relationship is statistically significant at the level (0.01). Therefore, the alternative hypothesis is accepted and the null hypothesis, which states that(There is no statistically significant relationship between the maximum temperature factor and water consumption. In the city of Baghdad.
3. There is an inverse correlation between the humidity factor and water consumption in the city of Baghdad, assuming the other factors are constant, as the correlation coefficient reached (-.438), and the p-value was (0.000), which indicates that the relationship is statistically significant at the (0.01) level. Therefore, the alternative hypothesis is accepted and the null hypothesis, which states that(There is no statistically significant relationship between the monthly humidity factor and water consumption. In the city of Baghdad.The main hypothesis is verified, as we notice the presence of a statistically significant correlation between some influencing factors and water consumption in the city of Baghdad, and all of them were statistically significant at the significance level.(0.01), so the alternative hypothesis is accepted and the null hypothesis is rejected, which states that (there is no statistically significant relationship between some influencing factors and water consumption in the city of Baghdad).

2.2.4 Predicting factors affecting water consumption in Baghdad Forecasting Influencing FactorsAt this stage, the influencing factors are predicted, namely (the population of Baghdad, maximum temperatures, and monthly humidity), in order to use them later as inputs in building models to predict water consumption in Baghdad. By drawing the time series of the factors used, it was shown that the time series of the population factor of Baghdad contains an upward trend, so the model was used.Double Exponential Smoothing)) which is consistent with the upward trend property of the series, and the time series of the maximum temperature factor and the monthly humidity percentage had the constant seasonal property, so they were predicted using the (Holt-Winter—Additive) model, and the predictive values of the three factors were as in Figure (2), (3) and (4).

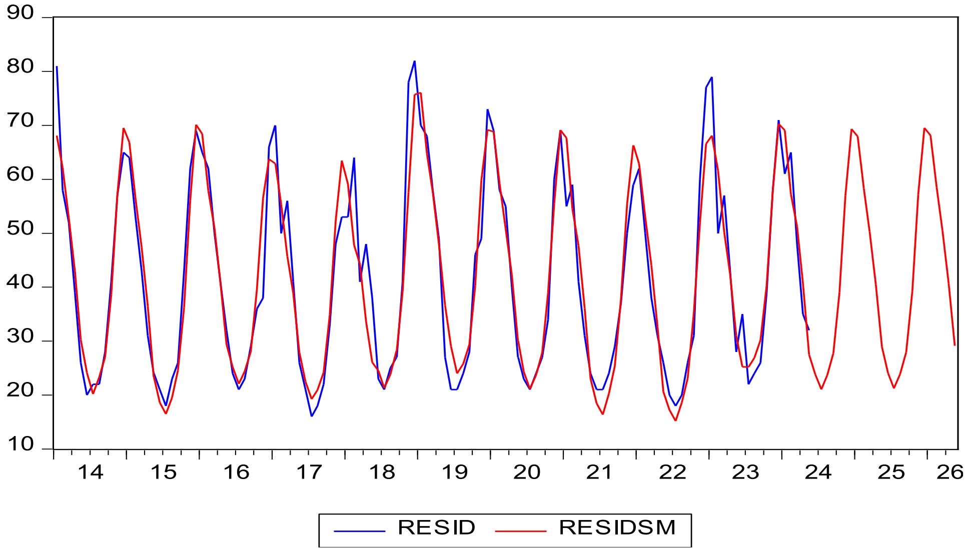
International Journal of Environmental Sciences
ISSN: 2229-7359
Vol. 11 No. 12s,2025
https://theaspd.com/index.php



Shape (2) Predictive values of the population of Baghdad city during the forecast period (24 months) using the Double Exponential Smoothing model.



Shape (3) Predictive values of the maximum temperature in Baghdad city during the forecast period (24 months) using the Holt-Winter—Additive model.

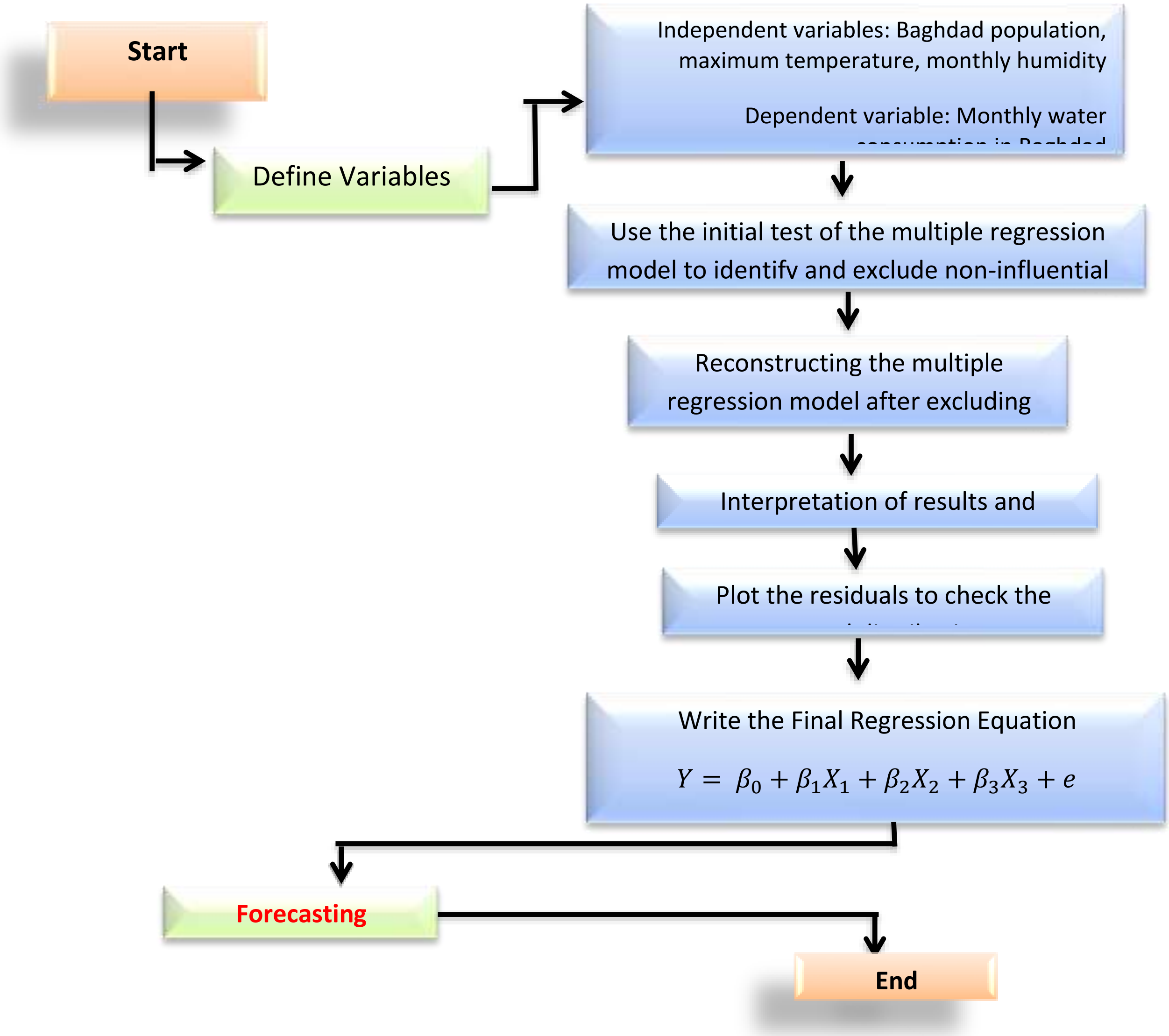


Shape (4) Predictive values of humidity in Baghdad city during the forecast period (24 months) using the Holt-Winter—Additive model.

3.2.4 Forecasting Using Multiple Regression Model

Monthly water consumption in Baghdad city is predicted using multiple regression model.Multiple Regression, based on the results of the prediction of the factors used in the research (the population of Baghdad, the maximum temperature, and the humidity percentage), and the diagram (1) shows the stages of applying the Multiple Regression model in predicting water consumption.

International Journal of Environmental Sciences
ISSN: 2229-7359
Vol. 11 No. 12s,2025
https://theaspd.com/index.php



The plan (1) Sequence of stages of applying the Multiple Regression model in predicting water consumption.

• **Definition of variables:**The variables are initially defined based on the multiple regression model equation.Multiple Regression:

$$\boldsymbol{Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + e} \quad .. \ (42)$$

Whereas:

:Y represents the monthly water consumption in Baghdad city.
The variable represents the population of Baghdad.X_1
:Represents the maximum temperature variable.X_2
X_3:Represents the monthly humidity variable.
β_0:Represents the constant limit parameter.
β_1 :Represents the population parameter.
β_2:Represents the maximum temperature parameter.
β_3 :Represents the humidity parameter.

• **Initial test of the model:**

At this stage the model is tested.Multiple Regression is the initial method to determine the factors that do not affect water consumption in the city of Baghdad. Table (3) shows the results of the initial test of the Multiple Regression model.

Table (3) Results of the initial test of the Multiple Regression model.

Multiple Regression									
R^2	$A\ R^2$	F	Sig	M	β	Beta	t	Sig	VIF
0.708	0.701	97,874	0.000b	C	-30837.114			0.007	
				X_1	0.014	0.663	13,329	0.000	1.027
				X_2	665,763	0.495	3.828	0.000	6,936
				X_3	4,022	0.005	0.04	0.968	6,929

Source: Prepared by the researcher based on the outputs of SPSS Version 26

Whereas:

Variance Inflation Factor (VIF): It represents the variance inflation factor, and measures the problem of multicollinearity. It is acceptable if its value (VIF < 5) means that there is no multicollinearity problem, and it is unacceptable if its value (VIF > 5) means that there is a multicollinearity problem and it must be addressed (Al-Kayyal et al., 2024: 242) (Bassiouni and Abdel Khaleq, 2023: 1187).

It is clear from the results of the table (3) The monthly humidity factor is not statistically significant and its value is greater than 0.05, and the VIF value of the factor is greater than 5, which means the presence of a multicollinearity problem.$X_2 X_3$

• **Reconstruct the model:**

We exclude the monthly humidity factor (), to increase the accuracy of the prediction, and repeat the test again to obtain the results of the table ($X_3$4) Verifying the main hypothesis, which is: There is no statistically significant moral effect of some factors affecting water consumption in the city of Baghdad, as well as its sub-hypotheses, which are:

1. There is no statistically significant effect of population on water consumption in Baghdad.
2. There is no statistically significant effect of maximum temperature on water consumption in Baghdad city.
3. There is no statistically significant effect of humidity on water consumption in Baghdad.

Table (4) Results of the multiple regression model after removing the monthly humidity factor.X_3

Multiple Regression									
R^2	AR^2	F	Sig	M	β	Beta	t	Sig	VIF
0.708	0.703	148,022	0.000b	C	-30513.083		-3.797	0.000	
				X_1	.014	0.664	13,556	0.000	1.002
				X_2	659,209	0.490	10,015	0.000	1.002

Source: Prepared by the researcher based on the outputs of SPSS Version 26

• Interpretation of model results and hypothesis testing:

The results of the table showed (4) The value of the coefficient of determination reached (0.708), and the value of the corrected coefficient of determination () reached (0.703) at the significance level (0.000), which indicates that the independent variables (the population of Baghdad city, and the maximum temperatures) explained 70.3% of the changes that occur in the monthly water consumption in Baghdad city, while there is a percentage of 29.7% explained by other variables that were not included in the model, and the model is statistically acceptable according to the value of F)) which is (148.022) and at a significance level ((0.000, and the value of VIF for the factors () and () is small, and this indicates the significance of the Multiple Regression model and it can be relied upon to predict the monthly water consumption in Baghdad city through the independent variables. Therefore, we answer the main hypothesis and its sub-hypotheses based on the results of Table (4) as follows:$A\ R^2 X_1 X_2$

1. It was found that there is a significant effect of the population factor on water consumption in the city of Baghdad, as the effect ratio reached (0.664), and the p-value was (0.000), which

International Journal of Environmental Sciences
ISSN: 2229-7359
Vol. 11 No. 12s,2025
https://theaspd.com/index.php

indicates the significance of the relationship at the significance level (0.05), so the alternative hypothesis is accepted and the null hypothesis is rejected, which states that(There is no statistically significant effect of the population factor on water consumption in the city of Baghdad).

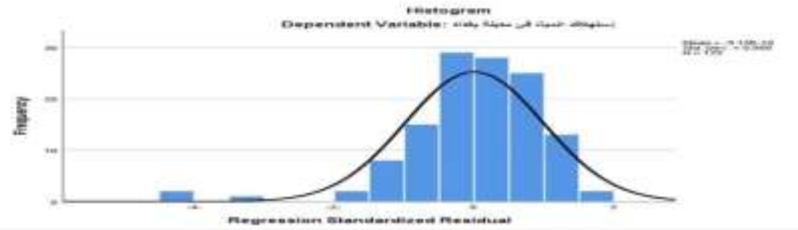
2. There is a significant effect of the maximum temperature factor on water consumption in the city of Baghdad, where the effect ratio reached (0.490), and the p-value was (0.000), which indicates the significance of the relationship at the significance level (0.05), so the alternative hypothesis is accepted and the null hypothesis is rejected, which states that(There is no statistically significant effect of the maximum temperature factor on water consumption in the city of Baghdad).

3. The results showed that the monthly humidity factor does not significantly affect water consumption in the city of Baghdad, so the null hypothesis is accepted, which states that:(There is no statistically significant effect of the monthly humidity factor on water consumption in the city of Baghdad).

.4 It was found that there was a significant effect of some of the factors used, as the value wasThe p-value for the population and temperature factor is (0.000), which indicates the significance of the relationship at the significance level (0.05), so the null hypothesis is rejected and the alternative hypothesis is accepted, which states that:(There is no statistically significant effect of some factors affecting water consumption in the city of Baghdad).

- **Drawing the remainders:**

The figure shows (5) The normality of the distribution of the residuals. It was confirmed that the data follow the normal distribution by conducting the (Kolmogorov) test, where it was shown that the value (Sig=0.053) is greater than 5%, and this indicates that the data follow the normal distribution, as Figure (6) shows the collection of data around the straight line. Therefore, the residuals follow the normal distribution, which is one of the conditions for the validity of the regression analysis procedure.



Shape ((5) It shows the normality of the distribution of residuals.

Shape (6) Distribution of monthly water consumption data around a straight line.

- **Create a model equationMultiple Regression:**

The equation of the multiple regression model is written based on the results of the table (17-3) As follows:

$$\boldsymbol{Y = -30513.083 + 0.014X_1 + 659.209X_2 \quad ..(43)}$$

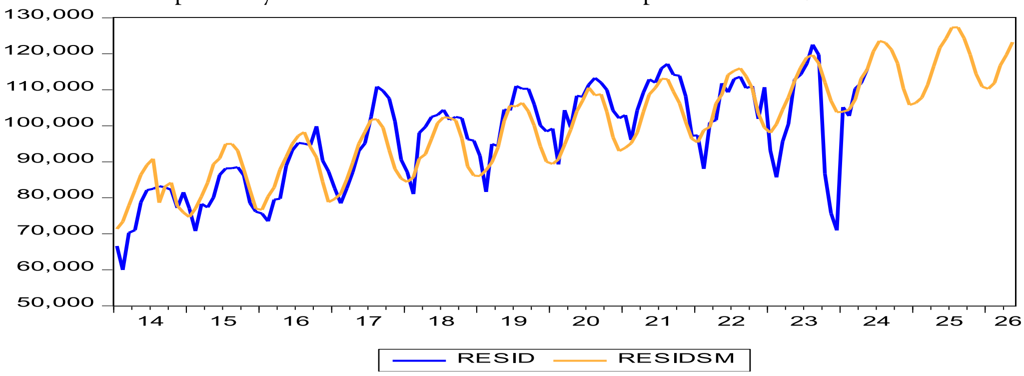
- Predicting water consumption in Baghdad:

After creating the multiple regression model equation, we can predict the monthly water consumption in the city of Baghdad, and by substituting the value of each of the following:)) which represents the predicted values of the monthly population factor of Baghdad city, and () which represents the predicted values of the monthly maximum temperature factor that were predicted. The results of the predicted values of monthly water consumption in Baghdad city using the multiple regression model equation were as in Table (5), and Figure (7) shows the actual values and predicted values of water consumption in Baghdad city.X_1X_2

Table (5) Predictive values for monthly water consumption in the city of Baghdad (Thousand) for the period from June 2024 to May 2026 using the Multiple Regression model.m^3

Predicted Values Using Multiple Regression					
June 2024	120607	February 2025	107780	October 2025	119739
July 2024	123585	March 2025	111351	November 2025	114411
August 2024	123069	April 2025	116571	December 2025	110863
September 2024	121169	May 2025	121658	January 2026	110281
October 2024	117358	June 2025	124241	February 2026	111875
November 2024	110382	July 2025	127285	March 2026	116765
December 2024	105779	August 2025	127362	April 2026	119677
January 2025	106318	September 2025	124342	May 2026	123248

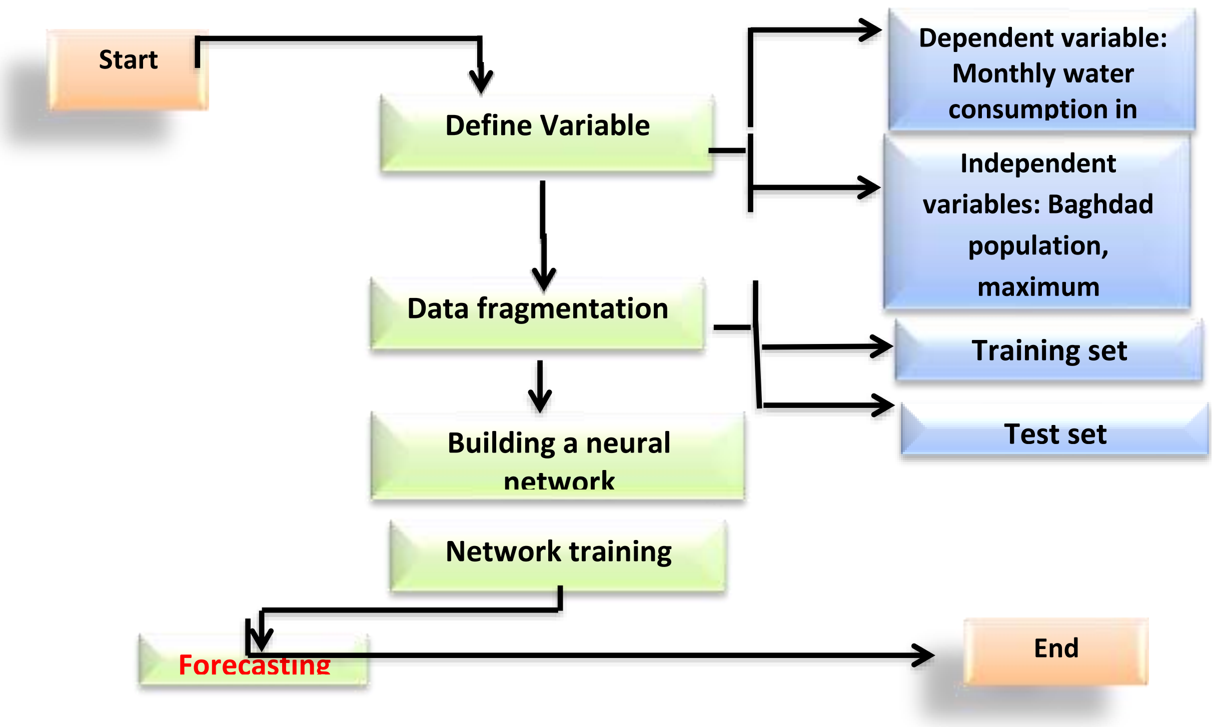
.**Source:** Prepared by the researcher based on the outputs of SPSS Version 26



Shape (7) Actual and predicted values of monthly water consumption in Baghdad city using the Multiple Regression model.

4.2.4 Prediction using MLP-Neural Networks

Monthly water consumption in Baghdad city is predicted using a multi-layer neural network model.MLP-NN, based on the prediction results of the factors used, which are (the population of Baghdad, maximum temperatures, and humidity), and diagram (2) shows the stages of applying the MLP-Neural Networks model in predicting water consumption.



The plan (2) The sequence of stages of applying the MLP-Neural Networks model in predicting water consumption.

Source: Prepared by the researcher.

- **Definition of variables:**

At this stage, the variables are defined (specified), where the dependent variable is represented by(Yi) represents the monthly water consumption in Baghdad city, and the independent variables (Xi) represent the population of Baghdad city, the maximum temperature, and the monthly humidity.

- **Data fragmentation:**

It consists of determining the training sample consisting of70% of the data and 30% are left for the test sample.

- **Determine the structure of the neural network (neural network architecture):**

Multiple experiments were conducted using multi-layer neural network technology.(MLP-NN), by using the factors affecting consumption (Baghdad population, maximum temperatures, monthly humidity) as inputs. Experiments were also conducted on models with one and two hidden layers to reach the best model for predicting monthly water consumption in Baghdad. The models will be compared to choose the best model according to the lowest error rate in the training set and the testing set. Table (6) shows the neural network experiments.

Table (6) MLP-Neural Network experiments models.

Experience	Training		Testing		Relative training error	Relative error of test	Hidden layers
1	P	69.6%	P	30.4%	22.5%	13.6%	1
	N	87	N	38			
2	P	75.2%	P	24.8%	23%	10.1%	1
	N	94	N	31			
3	P	68%	P	32%	26.5%	9.5%	1
	N	85	N	40			
4	**P**	**78.4%**	**P**	**21.6%**	**22.9%**	**8.1%**	**1**
	N	**98**	**N**	**27**			
5	P	67.2%	P	32.8	25.7%	13.5%	2
	N	84	N	41			
6	P	68.8%	P	31.2%	22.2%	14.2%	2
	N	86	N	39			
7	P	64.8%	P	35.2%	25.1%	14%	2
	N	81	N	44			
8	P	77.6%	P	22.4%	22.1%	10.7%	2
	N	97	N	28			

Source: Prepared by the researcher based on the outputs of SPSS Version 26

Based on the results of the table (19-3), we note that the model in experiment No. (4), which is MLP(3,3,1), has the lowest error rate in the training and testing set compared to the other experimental models, as the relative error of the model reached (0.229) in the training set and (0.081) in the testing set, with one hidden layer. Therefore, the model can be chosen as the best model for predicting monthly water consumption in the city of Baghdad using (MLP-NN). Table (7) shows the results of the MLP(3,3,1) model, which is a candidate for prediction.

Table (7) Results of the MLP(3,3,1) neural network model candidate for prediction.

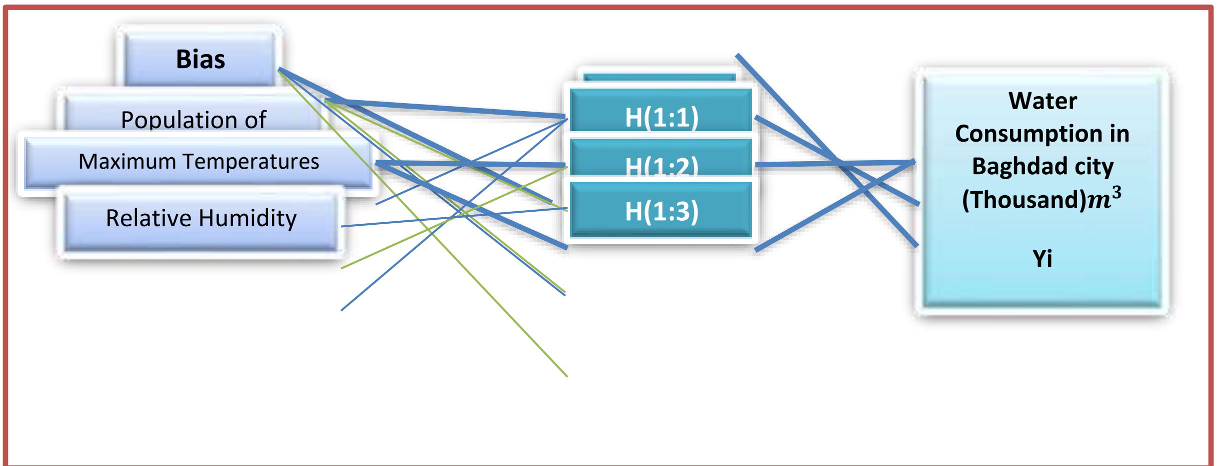
Input Layer	Training		Testing		Hidden Layer(s)		Relative training error	Relative error of test	Independent Variable Importance			Output Layer
X1	p	78.4 %	P	21.6 %	N	1	0.229	0.081	X_i	Importance	Normalized Importance	

X2	N		N		NU	3			X_1	51.9%	100%	Yi
									X_2	38%	73.2%	
X3		98		27					X_3	10.1%	19.5%	

Source: Prepared by the researcher based on the outputs of SPSS Version 26

The table shows (7) Three factors were used in the (Input Layer): () which represents the population of Baghdad, () which represents the maximum temperature, and () the monthly humidity. The monthly dataset in the model (a total of 125) was divided into two groups, the first is the model training set, which consists of (98) monthly records representing 78.4% of the dataset, and the second is the model testing set, which covered (27) records, representing 21.6% of the data. The model had one hidden layer and three units, while the relative error of the model was (0.229 in the training set, and 0.081 in the testing set). The model shows the relative importance of the population factor of Baghdad, which reached (100%) with an impact rate of (51.9%), therefore, the () factor is considered the most important and influential on monthly water consumption in Baghdad, which means that consumption increases directly with the increase in the population in Baghdad. While the relative importance of the maximum temperature factor () reached (73.2%) with an impact rate of (38%), and it ranks second in terms of importance and impact on water consumption. This means that with the rise in temperatures in the city, there is an increase in the quantities of water consumed in the city. The relative importance of the monthly humidity factor was (19.5%) with an impact rate of (10.1%).$X_1 X_2 X_3 X_1 X_2$

• The multi-layer neural network architecture can be defined for the model.MLP(3,3,1) is based on the results of the candidate prediction model, and Diagram (4) shows the structure of the network model (neural network architecture).



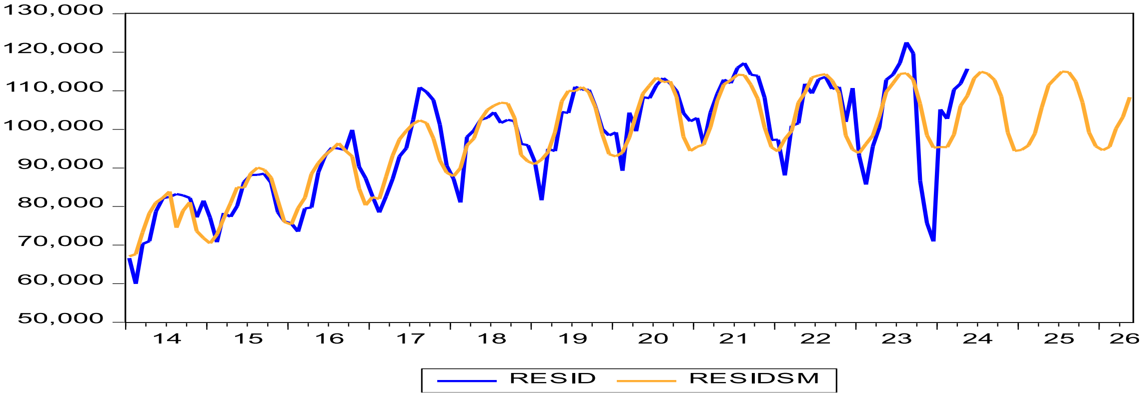
factors used (population, maximum temperatures, and humidity), and the table (8) It shows the predictive values of the monthly water consumption quantities in the city of Baghdad using the MLP(3,3,1) model. Figure (10) also shows the actual values and predictive values of the water consumption quantities.

Table (8) Predictive values of monthly water consumption quantities in the city of Baghdad (Thousand) for the period from June 2024 to May 2026 using the MLP(3,3,1) multi-layer neural network model.m^3

Predicted Values UsingMLP(3,3,1)					
June 2024	113278	February 2025	95807	October 2025	106997
July 2024	114975	March 2025	98928	November 2025	99246
August 2024	114454	April 2025	105646	December 2025	95623
September 2024	112780	May 2025	111445	January 2026	94580
October 2024	108549	June 2025	113315	February 2026	95445
November 2024	99144	July 2025	115021	March 2026	100213

December 2024	94410	August 2025	114855	April 2026	103209
January 2025	94679	September 2025	112461	May 2026	108338

Source: Prepared by the researcher based on the outputs of SPSS Version 26



Shape (10) Actual and predicted values of monthly water consumption quantities in Baghdad city using MLP(3,3,1) model.

3.4 Evaluating prediction models

After predicting the quantities of water consumption in the city of Baghdad using different prediction models, the accuracy measures were calculated for all prediction models in order to determine the best model among the models, and the model with the lowest error rate was chosen, and the table (9) Shows the results of the accuracy measures for the models.

Table (9) Results of error measures (MSE, MAD, MAPE) for the models used to predict monthly water consumption quantities in Baghdad city.

	Error measures		
Model	MSE	MAD	MAPE
Multiple Regression	54483.64	534.24	6%
MLP(3,3,1)	36004.99	427.06	5%

Source: Prepared by the researcher based on the laws of precision measurements.(MSE, MAD, MAPE).

It is clear from the table above that**MLP(3,3,1)**May outperform**Multiple Regression**Depending on the error measures used, which are:Mean Squared Error (MSE), Mean Absolute Error (MAD), Mean Absolute Percentage Error (MAPE), and thus this model is the most accurate for predicting monthly water consumption in Baghdad.

5. CONCLUSIONS AND RECOMMENDATIONS

5.1 Conclusions

1. Comparison between actual and predicted water consumption values in Baghdad showed a slight discrepancy, indicating the accuracy of the prediction results of the MLP-NN model compared to the Multiple Regression model. This is attributed to the ability of MLP-NN to represent complex and non-linear relationships in the data.

.2The accuracy indicators of the model were shown to beMLP-NN outperformed the model with MAPE = 5%), while the Multiple Regression model had MAPE = 6%), which reflects the effectiveness of MLP-NN in predicting water consumption in Baghdad city.

3. The results showed that population size and maximum temperatures have the greatest impact on water consumption in Baghdad. These factors explained, through the Multiple Regression model, 70.3% of the changes in water consumption, which confirms the importance of these factors in explaining consumption behavior in Baghdad.

5.2 Recommendations

.1It's essential to provide daily consumption data in Baghdad to increase forecast accuracy. Additionally, data on other factors, such as network leaks, will help improve model accuracy and provide more comprehensive insights into water consumption and the factors affecting it.

2. Expanding the use of artificial intelligence models through hybrid models, with the aim of achieving higher accuracy in predicting water consumption and enhancing their reliability in various situations in the city of Baghdad.
.3The results of this research will assist decision-makers in making various decisions regarding future project planning and adopting methods that seek to improve consumption management more efficiently, conserve water resources, and ensure their sustainability in the city of Baghdad.
.4These models are being used in cities similar to Baghdad in terms of drought, climate change, and population, to obtain predictive values that help understand future patterns of water consumption and identify the factors affecting consumption.

REFERENCES

Al-Kayyal, Mukhtar Ahmed, Amin, Magdy Shaaban, & Mahmoud Mohamed Ibrahim. (2024). The relative contribution of some psychological variables in predicting academic performance among university students: A comparative study using artificial neural network and multiple linear regression. Journal of Psychological Counseling. 77(4), 201-259.
Basyouni, Abdel Rahim Awad Abdel Khaleq.(2023). A comparative study of methods for treating the problem of linear duplication, applied to internal migration in Egypt. Trade and Finance. 43(1), 1180-1213.
Al-Zuhairi, H. H., Mozan, Z. H., & Sayah, H. R. (2023). Employing Time-Series Forecasting to Manage the Traffic Crisis in Baghdad Governorate/(Case Study). Al-Kut University College Journal, (Special issue).
Hamid, E.Y., & Abushaba, M.H.O. (2020). Use of Exponential Holt Model and the Box-Jenkins Methodology in Predicting the Time Series of Cement Production in Sudan. Journal of Al-Qadisiyah for computer science and mathematics, 12(1), Page-42.
Sant'Ana, D. R. (2021). Água eo ambiente construído.
Younis, H. I. (2016). Water Demand Forecasting In Al-Ain City, United Arab Emirates
Gil-Gamboa, A., Paneque, P., Trull, O., & Troncoso, A. (2024). Medium-term water consumption forecasting based on deep neural networks. Expert Systems with Applications, 247, 123234.
Lima, A.D., Trage, D.R., de Souza, F.F., & Pagani, R.N. (2023). Water demand forecast: a literature review. Research, Society and Development, 12(4), e4312440819-e4312440819.
Zheng, Y., Zhang, W., Xie, J., & Liu, Q. (2022). A water consumption forecasting model by using a nonlinear autoregressive network with exogenous inputs based on rough attributes. Water, 14(3), 329.
Saeed, AM & Aqla, IS (2018) "Effect of using quantitative methods to forecast demand in improving the performance of the supply chain" A case study in one of the industrial organizations Journal. of Economics and Administrative Sciences, 24(106), 295-295.
Montgomery, D. C., Jennings, C. L., & Kulahci, M. (2015). Introduction to time series analysis and forecasting. John Wiley & Sons.
Hamad, B. A. (2023). Combining Cluster Analysis with Multiple Linear Regression Analysis to Create the Most Accurate Prediction Model for Evaporation in the Kurdistan Region of Iraq. IRAOI JOURNAL OF STATISTICAL SCIENCES, 20(2).
Al-Sharoot, M.H., & Abdoon, E.Y. (2017). Prediction by using Artificial Neural Networks and Box-Jenkins methods: Comparison Study. Journal of Al-Qadisiyah for computer science and mathematics, 9(2), Page-1.
Khalil, D. M., & Hamad, S. R. (2023). A Comparison of Artificial Neural Network Models and Time Series Models for Forecasting Turkey's Monthly Aluminum Exports to Iraq. Journal of Survey in Fisheries Sciences, 10(1S), 4262-4279.
Awad, M., & Zaid-Alkelani, M. (2019). Prediction of water demand using artificial neural networks models and statistical model. International Journal of Intelligent Systems and Applications, 11(9), 40.
Ubah, JI, Orakwe, LC, Ogbu, KN, Awu, JI, Ahaneku, IE, & Chukwuma, EC (2021). Forecasting water quality parameters using artificial neural network for irrigation purposes. Scientific Reports, 11(1), 24438.
Kassem, A. A., Raheem, A. M., & Khidir, K. M. (2020). Daily streamflow prediction for Khazir river basin using ARIMA and ANN models. Zanco Journal of Pure and Applied Sciences, 32(3), 30-39.
Al-Sharoot, M.H., & Abdoon, E.Y. (2017). Prediction by using Artificial Neural Networks and Box-Jenkins methods: Comparison Study. Journal of Al-Qadisiyah for computer science and mathematics, 9(2), Page-1.
Agrawal, R. K., & Adhikari, R. (2013). An introductory study on time series modeling and forecasting. Nova York: CoRR, 200-212.
Al-Badiri, K. H., & Al-Kaabi, A. A. (2024). Forecasting common stock returns using time series models and neural networks.
Salaheddine, S. H., NAMANE, M. A., & MALIOUI, A. (2024). Using Artificial Neural Network Model to Forecast the Electricity Consumption in Algeria. Dirasat Journal Economic Issue, 15(2), 89-104.
Saleh, R. F., Kaml, B. S., & Hameed, L. M. (2024). Utilizing Fuzzy TOPSIS for Sustainable Development: ACaseStudyin Selection of Airport Location. Journal of Information Systems Engineering and Management, 9(4), 30116.