Iraqi Statisticians Journal / Vol. 2 (2025): Special issue for ICSA2025 : 12-22

IRAQI STATISTICIANS JOURNAL

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

Using Echo State Network based on Ridge Regression for Time Series Forecasting

Raed Arif Snaan[1], and Osamah Basheer Shukur[2]

[1,2] Department of Statistics and Informatics, College of Computer Science and Mathematics, University of Mosul, Mosul, Iraq

ARTICLE INFO

Article history:
Received 23 November 2024
Revised 29 November 2024
Accepted 11 January 2025
Available online 17 April 2025

Keywords:
Echo State Network (ESN)
Ridge Regression (RR) model
Climatic Time Series
Evaporation Forecasting
Multivariate time series regression

ABSTRACT

Studying climate in terms of predicting evaporation quantities through changing many factors and climate variables is of great importance to reduce the risks of climate change and its impact on many environmental phenomena affecting human and plant health. Weather patterns and conditions will be analyzed for a limited period of time. In this study, Ridge Regression (RR) was used to identify components, address the problem of multicollinearity, and overcome the problem of multicollinearity of the reconstructed inputs. The problem of multicollinearity occurs when building a multiple linear regression model, i.e. the presence of a strong correlation between the predictive variables, and this correlation leads to an increase in the variance of the regression parameter estimates and obtaining unstable estimates. Forecasting chaotic time series represented by climate data affecting the evaporation time series often suffers from the problem of nonlinearity in addition to multicollinearity. The Echo State Network (ESN) is a neural network specialized in predicting time series after addressing the problem of synchronization with the time variable as a recurrent network to address time-dependent effects and accurate prediction of time series in addition to nonlinear modeling. RR provides accurate predictions due to its structural effect to avoid overfitting while ESN will address the nonlinearity problem in addition to the problems addressed by RR model which ESN structure is based on. The expected results show that ESN model based on RR significantly outperforms traditional models in multivariate time series forecasting.

1. Introduction

Future study helps to predict in large quantities how much the change of many factors and changes in climate affects. Weather patterns and conditions will be analyzed for a limited period of time (2012-2022) to understand the nature and changes of climate. These studies include the collection and analysis of climate data through the Agricultural Meteorology Center, Nineveh Governorate - Mosul Station at longitude E 43.16° and latitude N: 36.33°. RR and ESN models were used to simulate future climate changes and how climate changes affect the evaporation time series. The study data include evaporation amount, maximum and minimum temperatures and their rates, maximum and minimum relative humidity and their rates, solar radiation, wind direction and average wind speed and its highest speed. The researchers[3] addressed the study of wind erosion as a serious environmental problem resulting from the interaction between different climatic factors. The ridge regression (RR) method was used to investigate the relationship between storm parameters and index in one of the regions of Iran during the period 2000-2014. Researchers studied RR modeling to investigate rainfall trends and temporal patterns of variation. RR is used as an improved technique to improve forecast accuracy with the aim of reducing the problems of multicollinearity and overfitting through temporal variations of rainfall and trying to find improved patterns and trends. Researchers Kim and King [23] used a new method to forecast time series data by using Deep ESN to deal with time series with large

* Corresponding author E-**mail** address: drosamahannon@uomosul.edu.iq
https://doi.org/10.62933/aej59z81

number of observations and multidimensional that suffer from nonlinearity and that artificial neural networks are used to model and forecast time series with nonlinear data. Researchers Viehweg et al. investigated the hyperparameters of ESN method as a systematic procedure, which is very important to determine the effect of non-uniform generated weight parameters in order to improve the prediction of random time series such as climate time series. Researchers Peng et al. [divided the components of non-seasonal time series by RR model with wave change to analyze the components of seasonal time series. While ESN was relied upon to analyze the components of non-seasonal time series. Researchers added the modeling results using RR with ESN to achieve the required prediction, which is methodologically different from the procedures and steps included in this research.

The research aims to measure the effect of climate factors on the amount of evaporation and detect the presence of multicollinearity problem and its effect on the regression model to find out the appropriate model for the study by using the regression echo state network (RR-ESN).

By estimating a multiple linear regression model, there will be many problems that create an incorrect estimation process due to the multicollinearity problem, i.e. the predictive variables are not independent, meaning there is a strong correlation between them. The Ridge Regression model can be used and applied as a proposed method to address the problem of multicollinearity present in the study data.

2. Methodology

2.1 Ridge Regression (RR) Model

Linear models are widely used in various fields of science, and one of the linear models that is widely used in analyzing data in many medical, economic and other applied sciences is multiple linear regression. The problem of multicollinearity occurs when building a multiple linear regression model, i.e. there is a strong correlation between the predictive variables, and this correlation leads to an increase in the variance of the regression parameter estimates and obtaining unstable estimates and large standard errors. The estimates of the linear model parameters and their variance will not represent the phenomenon under study, which makes the results of the ordinary least squares method inaccurate. In order to address the problem of multicollinearity, it is necessary to identify the variable or variables causing this problem. Several methods have been proposed by many researchers, including the (Ridge Regression) method, which was proposed by (Horel & Kennard, 1970), which aims to get rid of this problem and reach a method that has the best accurate estimate values. It is useful when used for simulation, and because the estimates of the (Ridge Regression) method are the best because they have the smallest squares of error.

Time series are defined as a series of observations of a system, which are usually chaotic or nonlinear. This can lead to problems of nonlinearity and uncertainty. The nonlinearity of climate data makes predictions complex. Time series forecasting uses historical information to estimate future values, so research into time series forecasting is very important. Some neural network models have already been used in time series forecasting, including the Echo State Network (ESN). The presence of a perfect or imperfect linear relationship between two or more predictive variables in a regression model leads to the problem of multicollinearity, which leads to unstable estimates. In order to address this problem, it is necessary to identify the variables causing this problem. In order to diagnose the problem of multicollinearity in a model, it is necessary to calculate the following:

Variance Inflation Factor (VIF)

The variance inflation factor measure is one of the important methods to detect the presence of multicollinearity problem. It measures the inflation in the variance of the estimated parameters when there is a linear relationship between the predictive variables in the model. When the predictive variables are orthogonal, i.e. there is no relationship between them, this means that (VIF=1) and the more its value exceeds one, it indicates the presence of a linear relationship between those variables. If (VIFj > 10 where j=1,2,….,n) it indicates that there is a multicollinearity problem and it can be found by relying on the coefficient of determination. The following formula is used to find the values of (VIF):

$$VIF = \frac{1}{\left(1-R_j^2\right)} \qquad (1)$$

$$j = 1,2,.....,\text{n}$$

Where:

n: Number of predictive variables.

R_j^2 :coefficient of determination of the predictive variable x_i extracted from the regression of x_j on the rest of the predictive variables in the model. The problem of multicollinearity occurs when the relationship between the predictive variables is a strong correlation and the ridge regression estimators treat the problem, when the variance of the estimated parameters is large where the treatment is done by the ridge regression method and multicollinearity is almost complete.

The use of the traditional least squares (OLS) method in the process of estimating the multiple linear regression model is as follows:

$$\hat{Y} = Xb + e \qquad (2)$$

Where:

$\hat{Y}$: The vector of observations of the estimated dependent variable with dimensions (P x 1).

X : The matrix of observations of the predictive variables with dimensions (P x M)

e : The vector of random errors with dimensions (P x 1). Assuming that E(e) = 0 and Var(e) = Iσ

b: The vector of unknown parameters with dimensions (M x 1), and the (OLS) method is used according to the following formula:

$$\hat{b} = \left(X'X\right)^{-1} X'Y \qquad (3)$$

In case of multicollinearity problem, the estimators of the linear model parameters($\hat{b}$) and its variance will not represent the problem under study, and then it needs to be treated by the multicollinearity problem treatment method, which is Ridge Regression Method. The RR method shows that adding a small positive quantity to the main diagonal of the matrix ($X'X$), this value will reduce the variance of the estimated parameters by calculating part of the bias, but the mean square error for the RR estimators will be less than the mean square error for the traditional least squares method. The formula for the RR estimators is as follows:

$$\hat{b}_k = \left(X'X + kI\right)^{-1} X'Y \qquad (4)$$

Where:

I: Identity Matrix and when the value of k = 0, the estimators $\hat{b}$ will be converted to (OLS) method estimators

And the above equation (4) shows us that the value of k is added to the main diagonal in the information matrix($X'X$), and to choose the value of k there are several methods, which are:

There are two main methods for choosing the Ridge parameter: the Ridge Trace method and the K value estimation method. The following explains the two methods:

1- Ridge Trace Method

The graph is the one adopted in the ridge trace method by drawing the coordinates to choose the best K value, as the vertical coordinate represents the estimated parameters of the RR method at any value of K with the bias factor increasing gradually starting from zero within the chosen range and the values of k specified between (0,1) in the horizontal coordinate, as the researcher chooses the appropriate k value so that all estimates are fixed with an increase in k and the function is stable

2- Methods of Estimate k:

The Ridge Parameter estimation method is an unbiased method equivalent to the Ridge Trace method, as determining the best value of the

Ridge Parameter depends on the values of the dependent variable and finding the value of Ridge Regression (K) means significantly reducing the variance and increasing the value of the bias square. When the value of (k) is greater than zero, we get biased estimates and are more stable. Therefore, the values of k are relied upon at the stability point.

2.2 Echo State Network (ESN)

ESN is a technology for processing power, memory, network and storage to deal with time-dependent data, which makes it very important in addressing the importance of time series forecasting. Its design is accurate for standard neural networks, its learning process is relatively simple, and it has a strong computing power to get rid of non-linear problems. Researchers are interested in the echo state network as a means of predicting time series since its formation as a neural network for remote learning, ESN uses a large constellation of highly repetitive neurons with irregular configuration called reservoirs, which are weak rank. The recurrent reservoir is created by a number of hidden nodes and an ESN is one of its branches. Assume that the dimensions of the input variables are (p) and the artificial neural network has a number of hidden nodes (m). In ESN, the output weights W_{out} are trained while the input weights W_{in} and reservoir weights W_z are generated randomly without training, which makes ESN superior to other traditional neural networks.

Therefore, ESN is widely applied in many aspects including time series forecasting [1; 3; 2]. The State Equation (SE) and Observation Equation (OE) can be written as follows, respectively:

$$Z_t = AZ_{t-1} + Bu_{t-1} + C'a_t \tag{5}$$

$$Y_t = CZ_t \tag{6}$$

$$r = \max(g, j)$$

g: is the number of the delayed series of the variable Z_t, j: is the number of the delayed series of errors, Z_t : is the state vector of dimension r, u_t: is the vector of the delayed series of errors, a_t: is the circular vector of the current errors, Equations (5) and (6) are ambiguous in implementation, and for the sake of convenience, their explanation can be rotated as follows:

$$Z_t = AZ_{t-1} + C'a_t \tag{7}$$

$$\hat{y}_t = CZ_t \tag{8}$$

Z_t: is the m-dimensional state vector, A is the (m x m)-dimensional state transition matrix, C is the (m x 1)-dimensional width transition matrix, $\hat{y}_t$ width is the circular vector representing the output series, m is the number of lagged series terms of Z_t , and each lagged series of the residual series a_t is one of the right-hand side of the RR-ESN data model equation after simplifying it and keeping only Z_t the left-hand side.

$$Z_t = \begin{bmatrix} Z_{1,t} & Z_{2,t} & \cdots & Z_{m,t} \end{bmatrix}' \tag{9}$$

$$A = \begin{bmatrix} P_1 & P_2 & P_3 & \cdots & P_m \\ 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0 \end{bmatrix}_{m\times m} \tag{10}$$

$$C = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \end{bmatrix}_{1\times m} \tag{11}$$

$(P_1, P_2, \ldots, P_m)$ All parameter values for the lagged series Z_t and all time-lagged residual series a_t are included on the right side of the equation for the RR-ESN data model after simplification and only Z_t the left side is retained. The matrix A, the row vector C, and the variables in the RR-ESN data model equation will define the input variables and the ESN structure used in this research.

ESN is a neural network implemented by linear regression method, which is better than traditional neural network data generation based on gradient method and also has high quality and better learning speed[11]. At the same time, the computational cost of ESN is lower than other types of mathematical and sequential steps required to solve the traditional learning problem, especially for gradient descent methods or global search methods, according to the efficient method and has the ability to change to traditional recurrent networks (RNNs). The ESN modeling process is described using the following two equations:

$$Z_t = f(W_z Z_{t-1} + W_{in} X_t) \quad (12)$$

$$\hat{y}_t = W_{out} Z_t \quad (13)$$

t represents the number of time stages, and represents the rotating input vector of dimension (t × 1), and represents the rotating matrix of internal states of dimension (m × t) that is present in the hidden layer, and m represents the output vector of dimension (t × 1) of the neural network. The output weights can be calculated using the general inverse method with the following formula:

$$W_{out} = \hat{y}_t' Z_t^+ = \hat{y}_t' (Z_t' Z_t)^{-1} Z_t' \quad (14)$$

$$W_{out} = C, W_{in} = C', W_z = A \quad (15)$$

$$z = W_z Z_{t-1} + W_{in} X_t \quad (16)$$

It is a non-linear function, and in this research the tangent function and its formula were used:

$$f(z) = \tanh(z) = \frac{1 - e^{-2z}}{1 + e^{-2z}} \quad (17)$$

2.3 Time Stratification (TS)

The behavior of time stratification is a behavior that allows the data to be arranged in time according to seasonal changes that clearly appear as influences on the method of the time series and the behavior of the forecast results. The case of description of time stratification can be applied to heterogeneous time series if they contain seasonal time trends repeated under the same conditions of correlation and influence, and it works to reach more consistent data than the total data and thus obtain more accurate results [6; 12].

2.4 Root Mean Square Error (RMSE)

(RMSE) is one of the measures that is used to display ideas about the accuracy of data prediction, and the equation (RMSE) [4]is written as follows:

$$RMSE = \sqrt{\frac{\sum_{t=1}^{n} a_t^2}{n}} \quad (18)$$

$$t = 1, 2, \ldots\ldots\ldots\ldots, n, \quad a_t = y_t - \hat{y}_t$$

n: is the number of observations, y_t : the true value at time t, $\hat{y}_t$: the predicted value of observation y at time t, a_t : the error series at time t.

3. Results and discussion

3.1 Data Description and Study Framework

The parameters for Evaporation Prediction (RR-ESN) In this research, the forecasting methods represented by the evaporation amount model for Mosul city will be used, and since the climate factors data for Mosul city are not consistent. The weather patterns and conditions will be analyzed for a limited period of time (01/11/2012-30/03/2022). To understand the nature and changes of climate. These studies include collecting and analyzing climate data through the Agricultural Meteorology Center, Nineveh Governorate - Mosul Station at longitude E 43.16° and latitude N: 36.33°. RR and ESN models were used to simulate future climate changes and how climate changes affect the evaporation time series. The RR model was used as a commonly used traditional method, as the best RR model was obtained after several attempts to reach an RR-ESN model that fits the study data. In this study, meteorological data represented by climatic factors and evaporation amounts were used for one of the agricultural meteorological stations in Nineveh Governorate. Many researchers in previous studies used different weather data for prediction and concluded that the weather data is not linear at all, and therefore the RR model may be inaccurate in its prediction results due to this problem in the data. For these reasons, it is often suggested to use other non-linear methods that deal with this type of data better and thus give better prediction results compared to RR models. The general framework includes the following steps:

1- Using the Time Stratified (TS) method to convert the data into two seasons, hot and cold.
2- Dividing the data for each season into training data and test data.
3- Modeling the data for the training period using the RR model.

4- Using ESN based on the RR model to structure the network inputs, which are matrix A, vector C, vector of residuals, and input matrix. This method can be referred to as the hybrid RR-ESN.

The general framework of the study can be formulated as in Figure (1) below. The evaporation quantities variable for the hot and cold seasons for the training and test periods can be described as follows in the figures(2 & 3).

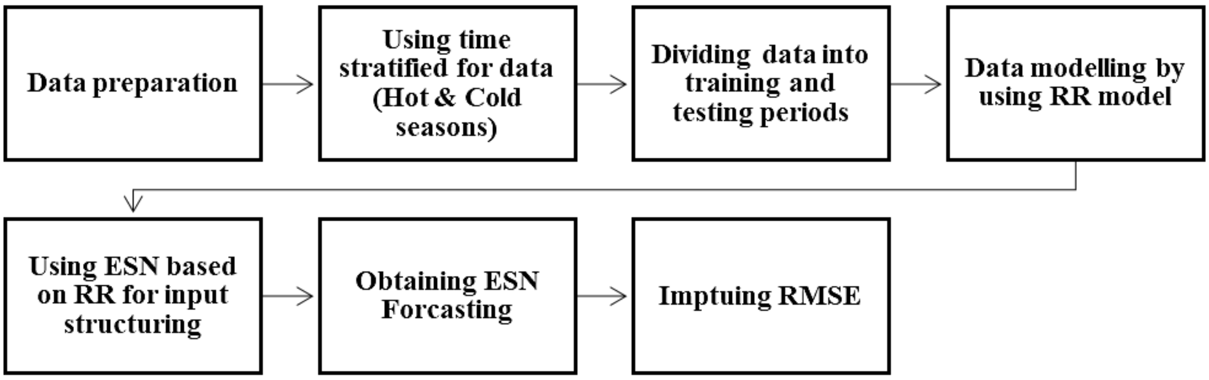

Figure 1. General framework for the research

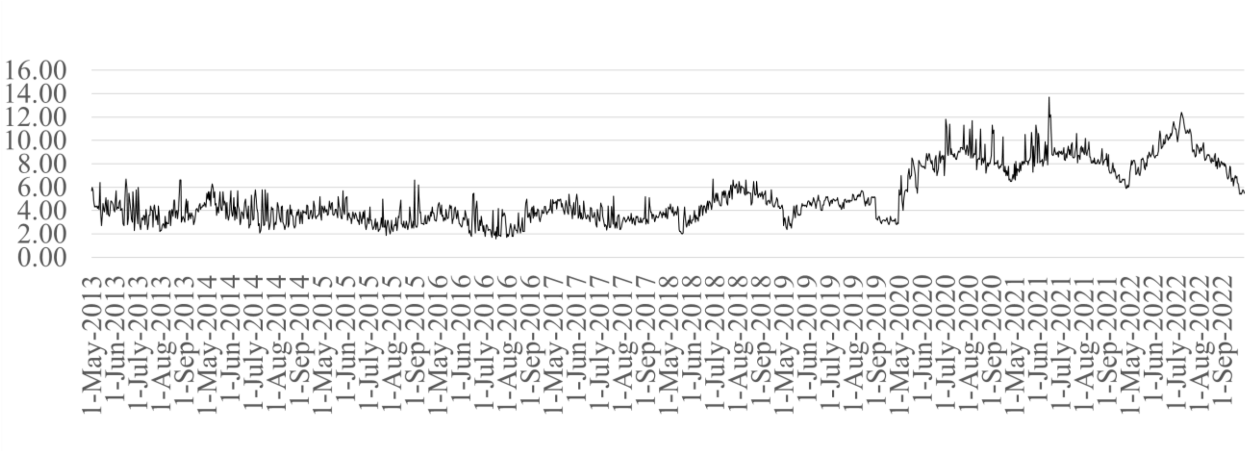

Figure 2. time series plot of the evaporation of hot season

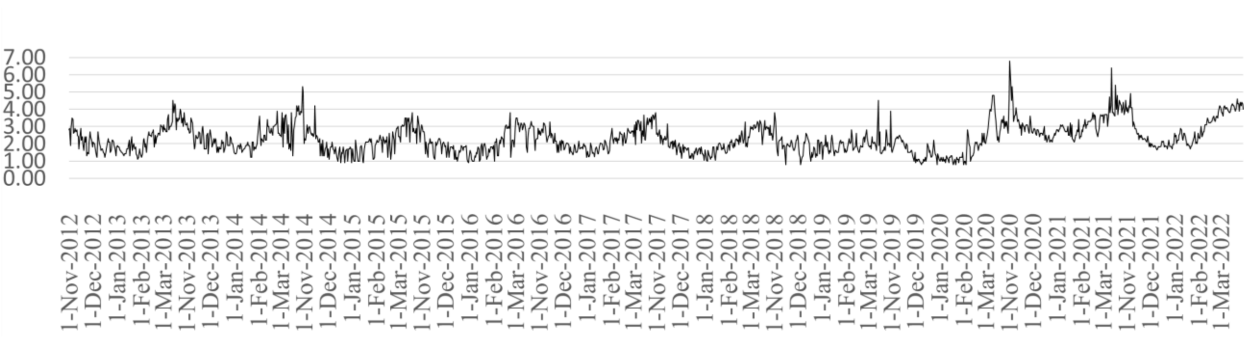

Figure 3. time series plot of the evaporation of cold season

In this aspect, the forecast cases that represent the model for the data in the two seasons (cold and hot) for Mosul city will be used to predict the evaporation quantities using the (RR-ESN) method under the influence of climate factors. The climatic factors data in Mosul city are heterogeneous. Therefore, the two-season method (hot season and cold season) was used to reach a more consistent forecast for the climatic factors for the daily data for the time period (01/11/2012 - 30/09/2022). The hot season data for the consecutive months (May - September) of the years are within the period (2013 - 2022). The cold season data for the consecutive months (November - March) of the years are within the period (2012-2022). 1200 observations were used to train for the hot season for Mosul city. 1200 observations were used to train for the cold season for Mosul city. 300 observations were used to test for the hot season for Mosul city, and 300 observations were used to test for Mosul city to predict the forecast quantities. These data were modeled using the seasonal multiplier (RR-ESN) model. Based on the above, the seasonality of the data and its models will be assumed on the basis that each seasonal cycle will consist of five months, i.e. (s = 5). The conclusions are that the significance of the model is affected by the problem of multicollinearity. The RR estimates are biased reflecting the biased least squares estimates, but the RR estimates are distinguished in terms of statistical significance.

3.2 Ridge Regression (RR) Model

The following research presents the results of applying Ridge Regression to time series data on the effect of climate factors on evaporation amounts, which represents the main dependent variable in the study and the climate factors affecting it during the period (2012-2022).

1- Detecting the problem of multicollinearity

We notice the insignificance of the estimated parameters of the predictive variables with each other, and in order to determine which of the predictive variables causes the problem of multicollinearity, we can benefit from the variance inflation factor, which appears high for two variables (temperature rate and relative humidity rate), which means that the relationship between the two variables with other predictive variables is a completely linear relationship, which indicates the presence of the problem of multicollinearity.

2- Using the ridge regression (RR) method in estimating the model equation for climate factors.

The ridge regression method was used in estimating the model parameters in order to address the problem of multicollinearity, as the ridge regression method depends on choosing the value of the ridge parameter (k), which works to stabilize the estimated parameters of the model as well as reduce the variance inflation factors (VIF) for these parameters to values close to one. The instability of the estimated parameters is observed when (k=0) represents the parameters estimated by the traditional least squares method, and then the estimated parameters move towards stability when (k=0.01), which means that the relationship between the predictive variables approaches orthogonality, i.e. after (0.01), and this conclusion is reinforced by the fact that the values of the variance inflation factor (VIF) decrease significantly and approach one when (k=0.05) for the two variables mentioned above. By using simple linear correlation or Pearson correlation as well as through the values of the Variance Inflation Factor (VIF), it was concluded that there is a multicollinearity problem between the variables (x_1, x_2, x_3) and (x_4, x_6) for the hot season data as in the tables below. The table (1) below show the highest values of the variance inflation factor (VIF) among the ten variables ($\boldsymbol{x_1}$, $\boldsymbol{x_2}$, $\boldsymbol{x_3}$, $\boldsymbol{x_4}$, $\boldsymbol{x_5}$, $\boldsymbol{x_6}$, $\boldsymbol{x_7}$, $\boldsymbol{x_8}$, $\boldsymbol{x_9}$, $\boldsymbol{x_{10}}$) as explanatory variables when y the evaporation variable is the dependent variable.

	y	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}
y	1	0.097	0.035	0.072	-0.039	-0.040	-0.044	0.042	0.837	0.252	-0.100
x_1		1	0.760	0.945	-0.592	-0.643	-0.687	0.261	-0.062	-0.187	-0.009
x_2			1	0.931	-0.565	-0.541	-0.629	0.248	0.027	-0.005	0.015
x_3				1	-0.617	-0.633	-0.703	0.271	-0.021	-0.107	0.002

$M =$	x_4					1	0.498	0.951	-0.120	0.104	0.141	-0.042
	x_5						1	0.741	-0.435	0.106	0.135	-0.030
	x_6							1	-0.247	0.118	0.158	-0.043
	x_7								1	0.088	0.069	0.061
	x_8									1	0.470	-0.199
	x_9										1	-0.295
	x_{10}											1

Table 1: The significance and VIF values of MLR for hot season

	Parameters	T-Value	P-Value	VIF
Constant	2.6010	5.63	0.000	
x_1	0.1068	3.18	0.002	9.28
x_3	-0.0808	-2.15	0.032	9.28
	Parameters	**T-Value**	**P-Value**	**VIF**
Constant	2.6010	5.63	0.000	
x_2	-0.1068	-3.18	0.002	7.55
x_3	0.1328	3.91	0.000	7.55
	Parameters	**T-Value**	**P-Value**	**VIF**
Constant	4.7020	19.03	0.000	
x_4	0.0057	0.38	0.701	10.54
x_6	-0.0196	-0.86	0.392	10.54

Therefore, we will resort to using ridge regression (RR) to solve the problem. The parameter values are as follows. The ridge regression equation can also be written as follows:

$$\begin{aligned}\hat{y} &= -0.6910 + 0.0030x_1 - 0.1995x_2 + 0.1795x_3 \\ &+ 0.0026x_4 - 0.0255x_5 - 0.0333x_6 - 0.0476x_7 \\ &+ 2.8042x_8 - 0.0838x_9 + 0.0014x_{10}\end{aligned} \quad (19)$$

For the cold season data, by using simple linear correlation or person correlation as well as through the values of the variance inflation factor (VIF), it was concluded that there is a multicollinearity problem between (x_1, x_2, x_3) and (x_4, x_6), as in the tables below. The table (2) below show the highest values of the variance inflation factor (VIF) among the ten variables (**x_1 , x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_9, x_{10}**) as explanatory variables when y the evaporation variable is the dependent variable.

		y	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}
	y	1	0.695	0.363	0.601	-0.180	-0.062	-0.167	0.718	0.107	0.152	0.120
	x_1		1	0.644	0.926	-0.368	-0.031	-0.293	0.459	-0.063	-0.033	0.152
	x_2			1	0.886	-0.302	0.001	-0.227	0.109	0.191	0.223	-0.181
	x_3				1	-0.373	-0.018	-0.290	0.333	0.056	0.090	0.003
$M =$	x_4					1	0.244	0.875	0.014	0.071	0.091	-0.027

x_5					1	0.682	0.058	0.315	0.239	-0.058
x_6						1	0.039	0.211	0.187	-0.049
x_7							1	-0.001	0.007	0.217
x_8								1	0.827	-0.311
x_9									1	-0.372
x_{10}										1

Table 2: The significance and VIF values of MLR for cold season

	Parameters	T-Value	P-Value	VIF
Constant	-0.0598	-0.92	0.360	
x_1	0.1627	17.85	0.000	6.98
x_3	-0.0599	-5.41	0.000	6.98
	Parameters	**T-Value**	**P-Value**	**VIF**
Constant	-0.0598	-0.92	0.360	
x_2	-0.1627	-17.85	0.000	4.63
x_3	0.26552	29.46	0.000	4.63
	Parameters	**T-Value**	**P-Value**	**VIF**
Constant	3.104	20.03	0.000	
x_4	-0.00897	-2.50	0.013	4.27
x_6	-0.00364	-0.67	0.502	4.27

Therefore, we will resort to using ridge regression (RR) to solve the problem. The parameter values are as follows. The ridge regression equation can also be written as follows:

$$\begin{aligned}\hat{y} &= 0.6549 + 0.1521x_1 + 0.0478x_2 - 0.1283x_3 \\ &+ 0.0049x_4 - 0.0070x_5 - 0.0105x_6 + 0.1097x_7 \\ &+ 0.0511x_8 + 0.0451x_9 - 0.0001x_{10}\end{aligned} \tag{20}$$

3.3 Hybrid RR-ESN method

Based on the ESN model which is basically based on the state space model in equations (7) and (8)above and also based on equation (19) and the ESN inputs will be structured so that matrix A and matrix C for the hot season will be as follows:

$$A = \begin{bmatrix} 0.0030 & -0.1995 & 0.1795 & 0.0026 & -0.0255 & -0.0333 & -0.0476 & 2.8042 & -0.0838 & 0.0014 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 \end{bmatrix}$$

$$C = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 \end{bmatrix}$$

And an Xtr matrix with dimension (number of variables * number of observations). As for the cold season, based on the ESN model, which is basically based on the state space model in equations (7)and (8) above, and also based on equation(20), the ESN inputs will be structured so that matrix A and matrix C are as follows:

$$A=\begin{bmatrix} 0.1521 & 0.0478 & -0.1283 & 0.0049 & -0.0070 & -0.0105 & 0.1097 & 0.0511 & 0.0451 & -0.0001 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 \end{bmatrix}$$

$$C=\begin{bmatrix}1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0\end{bmatrix}$$

And an Xtr matrix with dimension (number of variables * number of observations). As well as the results of applying these tests to the time series data shown in the following tables.

Below is a table representing the values of ridge regression (RR) and the root mean square error (RMSE) for the two seasons (hot and cold) based on 10 variables of climate factor.

Table 3: The RMSE values of RR and RR-ESN for training period

	RR	**RR-ESN**
Cold	4.3690	0.3804
Hot	4.7338	0.8285

Table 4: The RMSE values of RR and RR-ESN for testing period

	RR	**RR-ESN**
Cold	6.1861	0.9056
Hot	8.1286	2.5772

From the results presented in the tables (3 & 4), it is clear that the hybrid RR-ESN method is superior to the traditional method represented by the Ridge Regression model, as the RMSE values of the proposed hybrid method are lower. Therefore, its predictions are more accurate, which reflects a clear superiority in improving prediction and achieving the research objectives.

Hence, evaporation data as a dependent variable are significantly affected by changes in many climatic variables such as maximum and minimum temperatures and their rates, maximum and minimum relative humidity and their rates, solar radiation, wind direction and average wind speed and its highest speed. Represented by variables

$$x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_9, x_{10}$$

and this is through the effectiveness and efficiency of the hybrid RR-ESN method in improving predictions for multivariate time series, especially after using the TS time alignment method to achieve data homogeneity.

4. Conclusions

From the results in the Results and Discussion Section, it is possible to conclude the possibility of using the hybrid RR-ESN method to obtain the best predictions, i.e. the possibility of using it to improve the predictive results compared to what is obtained from predictions using traditional methods with climate variables in general and with the dependent evaporation variable and its impact on many climate variables and influences, especially when using methods specific to multivariate time series.

References

[1] Duan, H., & Wang, X. (2015). Echo state networks with orthogonal pigeon-inspired optimization for image restoration. *IEEE Transactions on Neural Networks and Learning Systems*, *27*(11), 2413-2425.

[2] Ebrahimi Khusfi, Z., Roustaei, F., Ebrahimi Khusfi, M., & Naghavi, S. (2020). Investigation of the relationship between dust storm index, climatic parameters, and normalized difference vegetation index using the ridge regression method in arid regions of Central Iran. *Arid land research and management*, *34*(3), 239-263.

[3] Grover, R., & Sharma, S. (2024). *Impact of Climate Change on Rainfall Pattern by using Ridge*

Regression Analysis. Paper presented at the 2024 International Conference on Computational Intelligence and Computing Applications (ICCICA).

[4] Han, M., & Xu, M. (2017). Laplacian echo state network for multivariate time series prediction. *IEEE Transactions on Neural Networks and Learning Systems, 29*(1), 238-244.

[5] Hyndman, R. J., & Koehler, A. B. (2006). Another look at measures of forecast accuracy. *International Journal of Forecasting, 22*(4), 679-688.

[6] Kim, T., & King, B. R. (2020). Time series prediction using deep echo state networks. *Neural Computing and Applications, 32*(23), 17769-17787.

[7] Ma, Q., Chen, E., Lin, Z., Yan, J., Yu, Z., & Ng, W. W. (2019). Convolutional multitimescale echo state network. *IEEE Transactions on Cybernetics, 51*(3), 1613-1625.

[8] Malig, B. J., Pearson, D. L., Chang, Y. B., Broadwin, R., Basu, R., Green, R. S., & Ostro, B. (2016). A time-stratified case-crossover study of ambient ozone exposure and emergency department visits for specific respiratory diagnoses in California (2005–2008). *Environmental health perspectives, 124*(6), 745-753.

[9] Peng, Y., Lei, M., Li, J.-B., & Peng, X.-Y. (2014). A novel hybridization of echo state networks and multiplicative seasonal ARIMA model for mobile communication traffic series forecasting. *Neural Computing and Applications, 24*, 883-890.

[10] Qiao, J., Li, F., Han, H., & Li, W. (2016). Growing echo-state network with multiple subreservoirs. *IEEE Transactions on Neural Networks and Learning Systems, 28*(2), 391-404.

[11] Tobias, A., Armstrong, B., & Gasparrini, A. (2014). *Analysis of time-stratified case-crossover studies in environmental epidemiology using Stata.* Paper presented at the United Kingdom Stata Users' Group Meetings 2014.

[12] Viehweg, J., Worthmann, K., & Mäder, P. (2023). Parameterizing echo state networks for multi-step time series prediction. *Neurocomputing, 522*, 214-228.

[13] Yang, C., Qiao, J., Wang, L., & Zhu, X. (2019). Dynamical regularized echo state network for time series prediction. *Neural Computing and Applications, 31*, 6781-6794.