IRAQI STATISTICIANS JOURNAL

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

ISSN: 3007-1658 (Online)

Chaotic Time Series Forecasting by using Echo State Network and Autoregressive Model

Shahla Tahseen Hasan[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)
Autoregressive (AR) model
Chaotic Time Series
Maximum wind speed
Univariate time series Forecasting

ABSTRACT

Chaotic time series forecasting such as maximum wind speed rates is of great importance in the fields of meteorology and renewable energy to reduce and control the harmful negative effects. The problem of wind speed is that it is affected by several interrelated factors such as temperature and atmospheric pressure, which are characterized by non-linearity through the influence of time series on differences that may be a cause of the emergence of uncertainty problems, which makes it difficult to model using traditional univariate time series methods. Echo State Network (ESN) is a neural network specialized in time series forecasting 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 its ability to model nonlinearly. This study presents the use of the Autoregressive (AR) model and then its hybridization with the deep echo state network, which is called the AR-ESN hybrid method by using the optimal structure of the AR model to determine the optimal inputs to the ESN network as the main contributions to solving the prediction problems for real data forecasts. The use of ESN as a proposed forecasting method is to improve the forecasting efficiency to reduce the risks associated with extreme weather fluctuations compared with traditional forecasting results. The results indicate that the ESN model based on AR model can contribute to increasing the forecasting accuracy of maximum wind speed compared with traditional models by using mean absolute percentage error (MAPE) as one of the criteria the forecasting accuracy.

1. Introduction

Time series are defined as a set of observations in a specific time order, where each value depends on the previous one, which generates momentum used in future forecasting of the studied phenomenon. In particular, time series of climate variables are affected by chaos and variation, which may be a cause of uncertainty problems.

The uncertainty of maximum wind speed data makes the forecasting process complicated and may not lead to accurate forecastings as well. Some researchers have proposed the AR method, which is the traditional method for forecasting univariate time series data. The inaccuracy of forecasting in AR models is due to the fact that they are linear models, while maximum wind speed data is characterized by nonlinearity. To forecast time series, past and current values are used to predict future values[2][5].

AR model based on ESN model based on Echo State Network (ESN) was used to deal with the nonlinear nature of maximum wind speed data. Conventional forecasting methods such as AR and AR model methods based on ESN model based on Echo State Network (ESN) were used to predict the maximum wind speed of Mosul.

[*] Corresponding authorE-**mail** address: drosamahannon@uomosul.edu.iq
https://doi.org/10.62933/faqd2g68

Due to the difference in maximum wind speed data, Time Stratified (TS) method was applied to divide the data into two seasons (hot season and cold season) to obtain more homogeneous data to predict the maximum wind speed for daily data for the period from (2012/07/16 to 2023/01/17). Autocorrelation function (ACF) and Partial Autocorrelation Function (PACF) were used to determine the appropriate ranks of the AR model and plot the time series to diagnose their stability.

By analyzing the results derived from applying the AR model and AR model based on ESN model based on Echo State Network (ESN), the AR model based on ESN model outperforms the AR model in its accuracy in predicting maximum wind speed using the mean square error (MSE).

The State Space (SS) model was also used to determine the basic equation structure of the Echo Stat Network (ESN) method. The maximum wind speed data can be classified into interval variables, not ratio variables, as zero in these data is not absolute and does not mean the absence of maximum wind speed. Therefore, parametric tests can be used provided that the data distribution is normal, or the sample size is large enough (minimum 30) according to the central limit theorem. Researchersused state space models as a modeling framework for analyzing environmental time series data. They are commonly used to model population dynamics including population metabolism dynamics. In this study, researchersmodeled and estimated climate data using the state space model. The study specifically aimed to identify the trend movement pattern, i.e. increase or decrease in the occurrence of climate change. The state space of climate projections was used to model the climate dataset. The researcher[2] used a diagnostic method for discrete-time state space modeling, where temperature fluctuations were modeled as a state space system with temperature time series as inputs, with the aim of providing a tool for predicting future scenarios for minimum and maximum temperatures. This study uses a forecasting-focused methodology to diagnose the system, with the overall goal of developing a realistic and dynamic system model. Temperature data recorded in the Makkah region of Saudi Arabia were used to test the accuracy and robustness of the method used. Kim and King [23] used a new method for predicting time series data known as AD-Dee ESN, which relies on additive decomposition (AD) applied to the series as a pre-step to the Deep ESN network. This method aims to deal with large, multidimensional, and nonlinear time series, where artificial neural networks are used to model and predict time series with nonlinear data. In a similar vein, researchers Viehweg et al. [used a systematic procedure to study the hyperparameters of the ESN method, where they considered the effect of randomly generated weight parameters in order to improve the forecasting of chaotic time series, such as climate time series. Researchers Peng et al. [also analyzed the components of non-seasonal time series using an AR model with wavelet transform to analyze the components of seasonal time series. ESN was used to analyze the components of non-seasonal time series, where the researchers combined the modeling results using the AR model based on the ESN model to obtain the desired forecasting, which is an approach different from the methodology used in our research. In our research, we adopted a new methodology to integrate ESN with the AR model by modeling the climate data of maximum wind speed using the AR model first. Then, the structure of AR models was used to determine the structure of the variables involved in the ESN method. All variables listed on the right side of each AR model (i.e., the variables after multiplying them by their parameter values and signs) will be adopted except for the original residual variable. We also adopted a second approach that takes the same variables on the right side of the AR models but without taking the values and signs of their parameters. Thus, there will be two hybrid approaches: AR based on the ESN model with parameters and AR based on the ESN model without parameters.

2. Methodology

2.1 Autoregressive (AR) Model

In the autoregressive (AR) model, the current values of the time series are expressed as a linear function of the previous values. Assuming that the time series is denoted by the symbol Y_t at time t, the autoregressive model of order p ($AR(p)$), which takes p into account the previous values, can be represented by the following equation:

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \cdots + \phi_p Y_{t-p} + a_t \quad (1)$$

Where:

Y_t: The current value of the time series at time.

$\phi_1, \phi_2, \ldots, \phi_p$: Are the coefficients that represent the extent to which each of the previous values affects the current value.

$Y_{t-1}, Y_{t-2}, \ldots, Y_{t-p}$: Are the previous values of the time series.

a_t: Is the random error (noise) at time and is assumed to be normally distributed with zero mean and constant variance $a_t \sim i.i.d.N(0, \sigma_a^2)$.

Examples of model equations with different orders:

1. AR(1) model:

If $p = 1$, then the equation becomes:

$$Y_t = \phi_1 Y_{t-1} + a_t \quad (2)$$

Where the current value depends only on the previous value.

2. AR(2) model:

If $p = 2$, then the equation becomes:

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + a_t \quad (3)$$

Here the current value depends on the two previous values. Each ϕ coefficient reflects the magnitude of the influence of the corresponding past value on the current value. If the coefficient value is high, it means that the past values have a strong influence on the current values, making the model suitable for forecasting in time series that depend heavily on past history.

2.2 State Space (SS) model

Using the AR model as a linear statistical model to model nonlinear time series data will introduce random uncertainty, which reduces the forecasting accuracy and deal with the nonlinearity of the data. To deal with random uncertainty, the SS method will be used for more accuracy in forecasting due to its good performance in time series forecasting [10; 12; 3-4].

The SS method can be presented as a statistical method for estimating and predicting unmeasured state space equations. The SS method is to combine observations and current forecasting values using weights that reduce biases and errors. The state equation (SE) and the observation equation (OE) form an equation system that can be called the linear model of the state space and can be written as follows:

$$Z_t = AZ_{t-1} + Bu_{t-1} + e_{1,t} \quad (4)$$

$$Y_t = CZ_t + e_{2,t} \quad (5)$$

Z_t is the m-dimensional state vector, u_t is a specific input vector, Y_t is an output observation vector, $e_{1,t}$ and $e_{2,t}$ are independent white noise vectors, A, B, and C are constant matrices.

The SE equation, which is equation (4), and the OE equation, which is equation (5), can be written in the state space respectively as follows:

$$Z_t = AZ_{t-1} + Bu_{t-1} + C'a_t \quad (6)$$

$$Y_t = CZ_t \quad (7)$$

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

g is the number of the lagged series of the variable Z_t, j is the number of the lagged series of the residuals, Z_t is the state vector with dimension r, u_t is the lagged series vector of the residuals, a_t is the rotated vector of the current residuals.

Equations (6) and (7) are complicated in application, and for simplicity they can be reformulated as follows [1].

$$Z_t = AZ_{t-1} + C'a_t \quad (8)$$

$$\hat{y}_t = CZ_t \quad (9)$$

$$a_t = Y_t - \widehat{Y}_t \quad (10)$$

Z_t is the m-dimensional state vector, A is the (m x m) dimension state transition matrix, C is the dimension display transition matrix (m x 1), $\hat{y}_t$ the displayed rotation vectors represent the SS output series, m is the number of lagged series terms of Z_t, and all the lagged series of the residuals series a_t are one of the right side of the AR data model equation after simplifying and keeping Z_t only the left side.

$$Z_t = \left[Z_{1,t} \quad Z_{2,t} \quad \cdots \quad Z_{m,t} \right]' \tag{11}$$

$$A = \begin{bmatrix} P_1 & P_2 & P_3 & \cdots & P_m \\ 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0 \end{bmatrix}_{m \times m} \tag{12}$$

$$C = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \end{bmatrix}_{1 \times m} \tag{13}$$

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

2.3 Echo state network (ESN)

ESN is an advanced computing technique that aims to deal with time-dependent data, making it very effective in handling time series forecasting tasks. Its design is the most accurate among standard neural networks, its learning process is relatively simple, and it has powerful computing power to solve nonlinear problems. Since its inception as a deep learning neural network, ESN has become a favorite choice for researchers in the field of time series forecasting [1].

ESN methodology uses a large set of randomly initialized recurrent neurons known as reservoirs, which act as a hidden layer. The randomly generated recurrent reservoir simplifies the training of a recurrent neural network, of which ESN is one type. Assuming that the dimension of the input variables is (p) and the number of hidden nodes in ESN is (m), the idea of ESN is to train W_{out} the output weights, while W_{in} the input weights and W_z reservoir weights are determined 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].

ESN is a neural network trained using linear regression method, which enables it to overcome the shortcomings of traditional neural network models. It also shows high-quality performance and faster learning speed [11]. The modeling process of ESN is described using the following two equations:

$$Z_t = f(W_z Z_{t-1} + W_{in} X_t) \tag{14}$$

$$\hat{y}_t = W_{out} Z_t \tag{15}$$

As shown in equations (14) and (15), which are derived from the state space model specialized in time series from the SE equation, which is equation (8), and the OE equation, which is equation (9), with the integration or formation of a nonlinear function such as the tangent function.

t represents the number of time steps, X_t represents the rotated input vector with dimension (1 x t) and is the series of residues in the SS model in equation (8), which is denoted by a_t, Z_t represents the rotated matrix of internal states with dimension (m x t) located in the hidden layer, $\hat{y}_t$ represents the output vector with dimension (t1 x) of the neural network. The output weights can be calculated using the general inverse method with the following formula:

$$W_{out} = \hat{y}_t' Z_t^{+} = \hat{y}_t' (Z_t' Z_t)^{-1} Z_t' \tag{16}$$

$W_z = A$, $W_{in} = C'$, $W_{out} = C$

$z = W_z Z_{t-1} + W_{in} X_t$

$f(.)$ is a nonlinear function and in this research the tan function was used, and its formula is:

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

2.4 Time Stratification (TS)

Time Stratified Method (TS) is an analytical approach used to align data temporally according to seasonal effects that clearly affect the behavior of the time series and forecast results.

Time Stratified Method can be applied to different time series if they include repeated seasonal time trends with the same context and effect, and it works to reach more homogeneous data than totsl data and thus obtain more accurate results [6; 12].

2.5 Mean Absolute percentage Error (MAPE)

MAPE is one of the criteria used to express the accuracy of data forecasting, and it can be written as follows:

$$MSE = \frac{\sum_{t=1}^{n} a_t^2 / y_t}{n} \quad (18)$$

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

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

3. Results and discussion

3.1 Data Description and Study Framework

In this aspect, forecasting methods will be relied upon, which include the AR model and the ESN model based on the AR model based on the Echo State Network (ESN), to predict the maximum wind speed in Mosul city obtained from the Ministry of Agriculture/Agricultural Meteorology Center, Nineveh Governorate - Mosul Station within the coordinates of longitude E (43.16°) and latitude N (36.33°). The maximum wind speed data in Mosul city suffers from the problem of non-conformity. Therefore, the TS method was used to harmonize the data across two seasons (hot season and cold season) in order to obtain more homogeneous data to predict the maximum wind speed for daily data for the period (2012/07/16-2023/01/17). The hot season data includes consecutive months (May - September) for the years falling within the period (2013 - 2022). While the cold season data covers consecutive months (November - March) for the years falling within the period (2012-2023). 1200 observations were used to train the hot season model for Mosul city. 1200 observations were used to train the cold season model for Mosul city. Also, 300 test observations were used for the hot season for Mosul city, and for the cold season, 300 test observations for Mosul city were used to predict the maximum wind speed. These data were modeled using the AR model. Based on the above, the seasonality of the data and its models will be assumed, as each seasonal cycle consists of five months, i.e. (s = 5). The framework includes each of the following steps:

1. Using the Time Trasefat (TS) method to convert the data into two hot and cold seasons.
2. Dividing the data for each season into training data and test data.
3. Modeling the data for the training period using the autoregressive AR model for time lags (2,3,4,5).
4. Using the Echo State Network (ESN) based on the AR 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 AR-ESN.

The general framework of the study can be formulated as in Figure (1). The maximum wind speed variable for the hot and cold seasons for the training and testing periods can be described in Figure (2) and Figure (3).

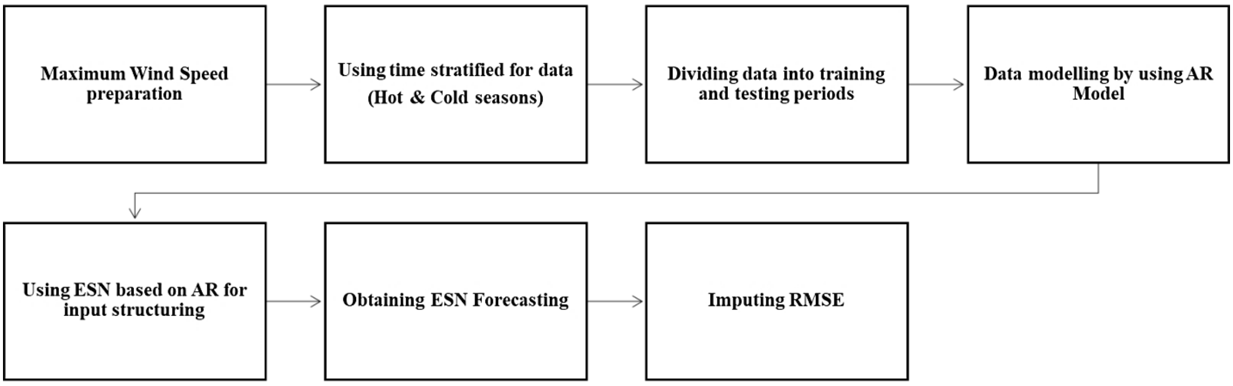


Figure 1. General framework for the research

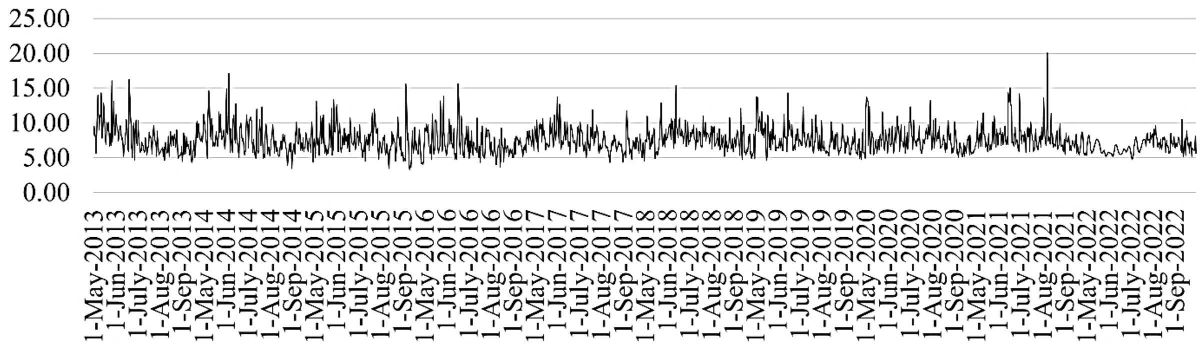


Figure 2. time series plot of the maximum wind speed of hot season

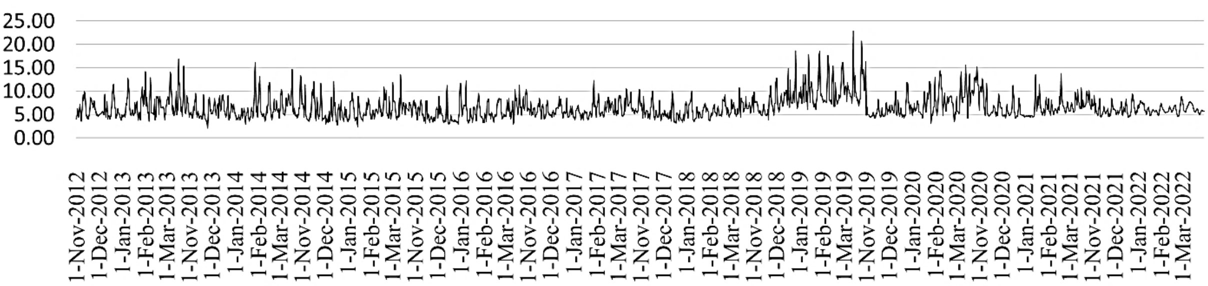


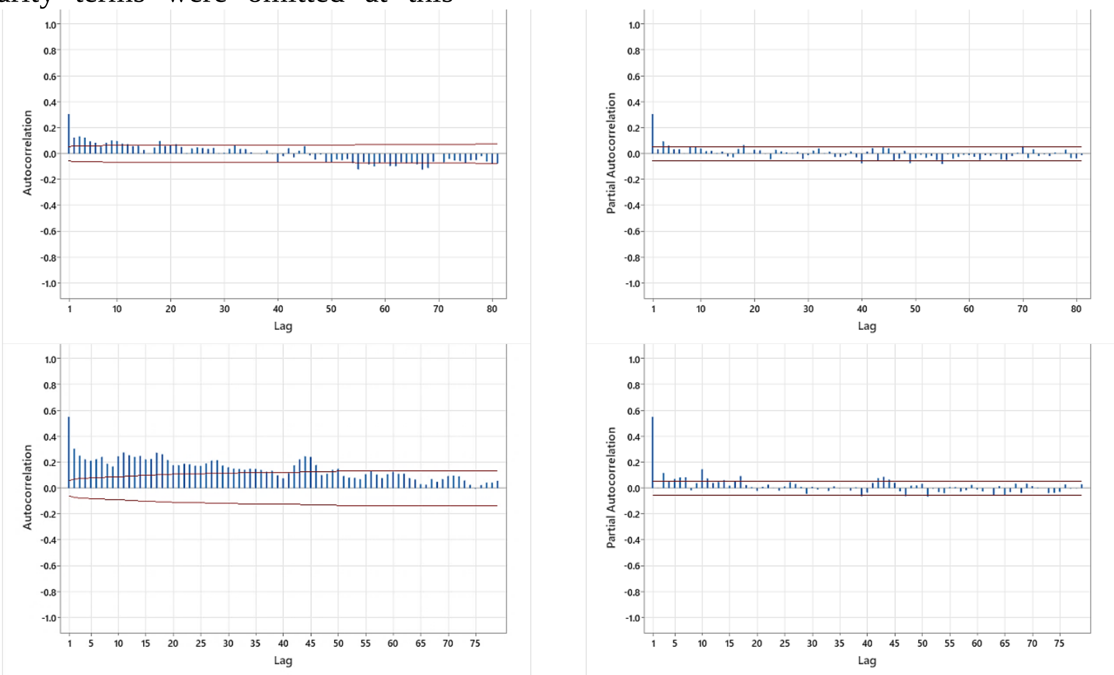
Figure 3. time series plot of the maximum wind speed of cold season

3.2 Autoregressive (AR) Model

The first step is to specify the AR model, where the partial autocorrelation function (PACF) of the original series reflects the number of significant orders for the AR model. The stationarity terms were omitted at this stage, because the AR model was only used to specify the input layer structure of the ESN. The ACF and PACF functions for the maximum wind speed data for the hot and cold season are inserted below in Figure (4).

Hot

Cold



a. ACF b. PACF

Figure 4. ACF and PACF of the maximum wind speed data for the hot and cold seasons.

From Figures (4) above, it is possible to conclude that the best AR(p) models are when (p=2– p=5) and the significant results of these models are as inserted in the tables (1,2) below.

Table 1: The significance of AR(2) –AR(5) for hot season

Type	Parameters	T-Value	P-Value	
AR 1	0.6173	23.80	0.000	**AR(2)**
AR 2	0.3552	13.70	0.000	
AR 1	0.5078	19.24	0.000	**AR(3)**
AR 2	0.1638	5.54	0.000	
AR 3	0.3101	11.75	0.000	
AR 1	0.4407	16.26	0.000	**AR(4)**
AR 2	0.1287	4.41	0.000	
AR 3	0.2002	6.85	0.000	
AR 4	0.2164	7.98	0.000	
AR 1	0.4068	14.83	0.000	**AR(5)**
AR 2	0.0972	3.31	0.001	
AR 3	0.1802	6.20	0.000	
AR 4	0.1472	5.01	0.000	
AR 5	0.1570	5.72	0.000	

Table 2: The significance of AR(2) –AR(5) for cold season

Type	Parameters	T-Value	P-Value	
AR 1	0.7499	26.58	0.000	**AR(2)**
AR 2	0.2027	7.18	0.000	
AR 1	0.6966	25.03	0.000	**AR(3)**
AR 2	0.0069	0.20	0.841	
AR 3	0.2622	9.41	0.000	
AR 1	0.6516	22.90	0.000	**AR(4)**
AR 2	0.0057	0.17	0.865	
AR 3	0.1442	4.26	0.000	
AR 4	0.1706	5.98	0.000	
AR 1	0.6248	21.91	0.000	**AR(5)**
AR 2	-0.0176	-0.52	0.601	
AR 3	0.1435	4.29	0.000	
AR 4	0.0673	2.00	0.046	
AR 5	0.1595	5.58	0.000	

From above tables (1,2), it is clear that the most of AR parameters are significant and the models can be fitted with the data regardless of the stationarity status of the series in figures (2,3) above. The AR models in tables (1,2) above can be written respectively based on equation (1) mathematically such as bellow.

$$Z_t = 0.6173Z_{t-1} + 0.3552Z_{t-2} + a_t \quad)19($$

$$Z_t = 0.5078Z_{t-1} + 0.1638Z_{t-2} + 0.3101Z_{t-3} + a_t \quad (20)$$

$$Z_t = 0.4407Z_{t-1} + 0.1287Z_{t-2} + 0.2002Z_{t-3} + 0.2164Z_{t-4} + a_t \quad (21)$$

$$Z_t = 0.4068Z_{t-1} + 0.0972Z_{t-2} + 0.1802Z_{t-3} + 0.1472Z_{t-4} + 0.1570Z_{t-5} + a_t \quad (22)$$

$$Z_t = 0.7499Z_{t-1} + 0.2027Z_{t-2} + a_t \quad)23($$

$$Z_t = 0.6966Z_{t-1} + 0.0069Z_{t-2} + 0.2622Z_{t-3} + a_t \quad (24)$$

$$Z_t = 0.6516Z_{t-1} + 0.0057Z_{t-2} + 0.1442Z_{t-3} + 0.1706Z_{t-4} + a_t \quad (25)$$

$$Z_t = 0.6248Z_{t-1} - 0.0176Z_{t-2} + 0.1435Z_{t-3} + 0.0673Z_{t-4} + 0.1595Z_{t-5} + a_t \quad (26)$$

3.3 State Space (SS) model

After obtaining AR(2–5) models of the maximum wind speed for hot and cold seasons, AR-SS models can be formulated respectively based on equations (8) and (9) such as follows.

$$\begin{bmatrix} Z_{1,t} \\ Z_{2,t} \end{bmatrix} = \begin{bmatrix} 0.6173 & 0.3552 \\ 1 & 0 \end{bmatrix}_{2\times2} \begin{bmatrix} Z_{1,t-1} \\ Z_{2,t-1} \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \end{bmatrix} a_t$$

$$Y_t = \begin{bmatrix} 1 \\ 0 \end{bmatrix}'_{1\times2} \times \begin{bmatrix} Z_{1,t} \\ Z_{2,t} \end{bmatrix} \quad (27)$$

$$\begin{bmatrix} Z_{1,t} \\ Z_{2,t} \\ Z_{3,t} \end{bmatrix} = \begin{bmatrix} 0.5078 & 0.1638 & 0.3101 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}_{3\times3} \begin{bmatrix} Z_{1,t-1} \\ Z_{2,t-1} \\ Z_{3,t-1} \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} a_t$$

$$Y_t = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}'_{1\times3} \times \begin{bmatrix} Z_{1,t} \\ Z_{2,t} \\ Z_{3,t} \end{bmatrix} \quad (28)$$

$$\begin{bmatrix} Z_{1,t} \\ Z_{2,t} \\ Z_{3,t} \\ Z_{4,t} \end{bmatrix} = \begin{bmatrix} 0.4407 & 0.1287 & 0.2002 & 0.2164 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}_{4\times4} \begin{bmatrix} Z_{1,t-1} \\ Z_{2,t-1} \\ Z_{3,t-1} \\ Z_{4,t-1} \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} a_t$$

$$Y_t = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}'_{1\times4} \times \begin{bmatrix} Z_{1,t} \\ Z_{2,t} \\ Z_{3,t} \\ Z_{4,t} \end{bmatrix} \quad)29($$

$$\begin{bmatrix} Z_{1,t} \\ Z_{2,t} \\ Z_{3,t} \\ Z_{4,t} \\ Z_{5,t} \end{bmatrix} = \begin{bmatrix} 0.4068 & 0.0972 & 0.1802 & 0.1472 & 0.1570 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}_{5\times5} \begin{bmatrix} Z_{1,t-1} \\ Z_{2,t-1} \\ Z_{3,t-1} \\ Z_{4,t-1} \\ Z_{5,t-1} \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} a_t$$

$$Y_t = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}'_{1\times5} \times \begin{bmatrix} Z_{1,t} \\ Z_{2,t} \\ Z_{3,t} \\ Z_{4,t} \\ Z_{5,t} \end{bmatrix} \quad (30)$$

$$\begin{bmatrix} Z_{1,t} \\ Z_{2,t} \end{bmatrix} = \begin{bmatrix} 0.7499 & 0.2027 \\ 1 & 0 \end{bmatrix}_{2\times2} \begin{bmatrix} Z_{1,t-1} \\ Z_{2,t-1} \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \end{bmatrix} a_t$$

$$Y_t = \begin{bmatrix} 1 \\ 0 \end{bmatrix}'_{1\times2} \times \begin{bmatrix} Z_{1,t} \\ Z_{2,t} \end{bmatrix} \tag{31}$$

$$\begin{bmatrix} Z_{1,t} \\ Z_{2,t} \\ Z_{3,t} \end{bmatrix} = \begin{bmatrix} 0.6966 & 0.0069 & 0.2622 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}_{3\times3} \begin{bmatrix} Z_{1,t-1} \\ Z_{2,t-1} \\ Z_{3,t-1} \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} a_t$$

$$Y_t = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}'_{1\times3} \times \begin{bmatrix} Z_{1,t} \\ Z_{2,t} \\ Z_{3,t} \end{bmatrix} \tag{32}$$

$$\begin{bmatrix} Z_{1,t} \\ Z_{2,t} \\ Z_{3,t} \\ Z_{4,t} \end{bmatrix} = \begin{bmatrix} 0.6516 & 0.0057 & 0.1442 & 0.1706 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}_{4\times4} \begin{bmatrix} Z_{1,t-1} \\ Z_{2,t-1} \\ Z_{3,t-1} \\ Z_{4,t-1} \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} a_t$$

$$Y_t = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}'_{1\times4} \times \begin{bmatrix} Z_{1,t} \\ Z_{2,t} \\ Z_{3,t} \\ Z_{4,t} \end{bmatrix} \quad)33($$

$$\begin{bmatrix} Z_{1,t} \\ Z_{2,t} \\ Z_{3,t} \\ Z_{4,t} \\ Z_{5,t} \end{bmatrix} = \begin{bmatrix} 0.6248 & -0.0176 & 0.1435 & 0.0673 & 0.1595 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}_{5\times5} \begin{bmatrix} Z_{1,t-1} \\ Z_{2,t-1} \\ Z_{3,t-1} \\ Z_{4,t-1} \\ Z_{5,t-1} \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} a_t$$

$$Y_t = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}'_{1\times5} \times \begin{bmatrix} Z_{1,t} \\ Z_{2,t} \\ Z_{3,t} \\ Z_{4,t} \\ Z_{5,t} \end{bmatrix} \tag{34}$$

The table (3) below shows the MAPE values for AR-SS for (2–5) lags of the maximum wind speed for hot and cold seasons.

Table 3: MAPE values of different SS-AR models

	Lags 2		**Lags 3**		**Lags 4**		**Lags 5**	
	Training	Testing	Training	Testing	Training	Testing	Training	Testing
Cold	11.9820	20.0841	11.1951	18.8023	10.5024	17.7789	10.2948	17.1773
Hot	9.4425	13.8324	8.1490	11.7331	7.4671	10.1601	6.3286	9.2672

3.4 Hybrid AR-ESN method

Based on the ESN model which is basically based on the state space model in equations (14) and (15) above and also based on equations of AR-SS above to determine the ESN inputs, the previous algorithm of ESN is performed and MAPE values can be seen such as in table (4) below.

Table 4: MAPE values of different AR-ESN models

	Lags 2		**Lags 3**		**Lags 4**		**Lags 5**	
	Training	Testing	Training	Testing	Training	Testing	Training	Testing
Cold	3.11954	9.98603	5.71713	9.97983	8.55145	10.01346	8.17095	9.96830
Hot	9.34131	9.55895	9.00557	9.54332	8.85217	9.54494	9.3015	9.47354

From the results presented in the tables (3,4), it is clear that the hybrid AR-ESN method is superior to the traditional method represented by the State Space model based on the AR model, as the MAPE values of the proposed hybrid method are lower. Therefore, its forecastings are more accurate, which reflects a clear superiority in improving forecasting and achieving the research objectives.

Hence, the maximum wind speed data as a time series with non-linear changes through the heterogeneity that was processed by TS after dividing the data into two relatively homogeneous seasons has achieved a significant improvement in the predictive results when using the hybrid ESN method based on AR, which reflects the efficiency and predictive effectiveness of the method.

4. Conclusions

From the results in the Results and Discussion Section, it is possible to conclude that the hybrid ESN method based on AR can be used to obtain the best forecasting, i.e. the possibility of using it to improve the predictive results compared to what is obtained from forecasting using traditional methods with climate variables in general and with the time series variable of maximum wind speed in particular using methods specific to predicting univariate time series.

References

[1] Adedotun, A., Latunde, T., & Odusanya, O. (2020). Modelling and forecasting climate time series with state-space model. *Journal of the Nigerian Society of Physical Sciences*, 149-159.

[2] Aljeddani, S. M. A. (2023). Forecasting temperature of the Saudi Arabian Province of Makkah using a discrete state–space modeling approach. *Journal of Umm Al-Qura University for Engineering and Architecture*, *14*(2), 79-94.

[3] Auger-Méthé, M., Newman, K., Cole, D., Empacher, F., Gryba, R., King, A. A., . . . Petris, G. (2020). An introduction to state-space modeling of ecological time series. *arXiv preprint arXiv:2002.02001*, 1-61.

[4] 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.

[5] Galanis, G., Louka, P., Katsafados, P., Pytharoulis, I., & Kallos, G. (2006). *Applications of Kalman filters based on non-linear functions to numerical weather predictions.* Paper presented at the Annales geophysicae.

[6] Gao, R., Du, L., Duru, O., & Yuen, K. F. (2021). Time series forecasting based on echo state network and empirical wavelet transformation. *Applied Soft Computing*, *102*, 107111.

[7] Gould, P. G., Koehler, A. B., Ord, J. K., Snyder, R. D., Hyndman, R. J., & Vahid-Araghi, F. (2008). Forecasting time series with multiple seasonal patterns. *European Journal of Operational Research*, *191*(1), 207-222.

[8] 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.

[9] Jaeger, H., & Haas, H. (2004). Harnessing nonlinearity: Predicting chaotic systems and saving energy in wireless communication. *science*, *304*(5667), 78-80.

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

[11] Li, J., Dai, Q., & Ye, R. (2019). A novel double incremental learning algorithm for time series

prediction. *Neural Computing and Applications, 31*, 6055-6077.

[12] Louka, P., Galanis, G., Siebert, N., Kariniotakis, G., Katsafados, P., Pytharoulis, I., & Kallos, G. (2008). Improvements in wind speed forecasts for wind power prediction purposes using Kalman filtering. *Journal of Wind Engineering and Industrial Aerodynamics, 96*(12), 2348-2362.

[13] 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.

[14] Madsen, H. (2007). *Time Series Analysis*: CRC Press.

[15] 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.

[16] Malmberg, A., Holst, U., & Holst, J. (2005). Forecasting near-surface ocean winds with Kalman filter techniques. *Ocean Engineering, 32*(3-4), 273-291.

[17] 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.

[18] 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.

[19] 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.

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

[21] 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.