**Iraqi Journal of Statistical Sciences**

http://stats.uomosul.edu.iq

# Using the State Space Model based on ARIMA Model for Air Temperature Forecasting

**Suha Saleem Mahjoob [1] Osamah Basheer Shukur[2]**

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

**Article information**

***Article history:***
Received July 7 ,2024
Revised: December 1,2024
Accepted: December 14 ,2024
Available online June 1,2025

***Keywords***:
ARIMA model ,State Space Model,Minimum Temperature ,Forecasting

***Correspondence:***
**Osamah Basheer Shukur**
drosamahannon@uomosul.edu.iq

**Abstract**

The high accuracy of forecasts with the temperature data is very important to control environmental damages such as desertification and water resources drought as well as it is important to control the uses of renewable energy and clean energy. Using the multiplicative seasonal integrated auto-regressive and moving average (SARIMA) model for forecasting with uncertainty problem in the modeling process especially with nonlinear data such as minimum temperatures will make the forecasting results become low in quality because ARIMA is a linear model. Improving the minimum temperature forecasting quality is the main aim for this study by using more suitable methods for modeling the data with the problem of uncertainty. In this study, the minimum temperature data for Mosul and Baghdad will be used as a case of study. The state space (SS) will be used based on the ARIMA model which can be called the hybrid ARIMA-SS model which will be used to solve the uncertainty problem caused by the non-linearity of temperature data. Therefore the forecasting results may be not accurate. Also, the climate data often suffers from heterogeneity, especially in non-tropical regions, due to the high difference between the hot and cold seasons of these data. Time stratified (TS) will be used to solve the problem of data heterogeneity. In the ARIMA-SS hybrid method ARIMA is used only for the purpose of specifying the input of the SS model. In this study, the SS model was used as a statistical method for estimating and forecasting the state space. The SS method is to combine observations with current forecasts values by using weights that reduce biases and errors. The ARIMA-SS hybrid model has been used to deal with uncertainty and improve the minimum temperature forecasting by handling it well. The performance of the ARIMA model and the ARIMA-SS hybrid model will be compared to determine which of them will perform with more accurate forecasts .The results showed that the ARIMA-SS hybrid model outperformed the ARIMA model and produced more accurate forecasts. Therefore, it is possible to conclude that ARIMA-SS hybrid model can be used to result better forecasting accuracy for the minimum temperature compared to the forecasting performance of the traditional ARIMA model.

DOI: 10.33899/iqjoss.2025.187756 

## 1. Introduction

The time series of climate variables depends on chaos and heterogeneity, which is the reason for the emergence of the problem of non-linearity and uncertainty. Time series is defined as a set of observations with a certain time sequence in which each observation depends on its predecessors, which generates a driving force used in forecasting the future of the studied phenomenon (1). The non-linearity of minimum temperature data makes forecasting a complex process and may also not lead to accurate forecasts. Some researchers have proposed to use the autoregressive model and integral moving averages (ARIMA), which is a traditional statistical method for forecasting single-variable time series data. It is possible to obtain the appropriate ARIMA model for the minimum temperature by following the methodology proposed by Box-Jenkins (2). The inaccuracy of

the ARIMA models ' forecast is caused by the fact that it is a linear time series model (3) whereas the temperature data is nonlinear data. To forecast time series, past and current values are used to forecast future values (4-6).
The state space (SS) was used based on the ARIMA model, which is called the ARIMA-SS hybrid model, to deal with random uncertainty caused by the non-linearity of minimum temperature data . The forecasting methods represented by ARIMA model and ARIMA-SS hybrid model (7) were used to forecast the minimum temperature of the cities of Mosul and Baghdad, since the data of the minimum temperature of Mosul and Baghdad are heterogeneous, therefore, the time stratified (TS) was used[8] to align the data in two seasons (the hot season and the cold season) to obtain more homogeneous data for forecasting the minimum temperatures for the daily data for the period (1/11/2011-31/10/2022). The autocorrelation function (ACF) and the partial autocorrelation function (PACF) were used to determine the ARIMA model in addition to plotting the time series to diagnose the stationarity of the time series and determine the rank of the model .From watching the results obtained after applying the double seasonal ARIMA model and the ARIMA-SS model. The results showed the superiority of the ARIMA-SS hybrid model over the doubled seasonal ARIMA model because it was more accurate in forecasting small temperatures based on the Mean Squared Error (MSE) scale (8). The autocorrelation function (ACF) and the partial autocorrelation function (PACF) were used to determine the ARIMA model in addition to plotting the time series to diagnose the stationarity of the time series and determine the rank of the model .

From watching the results obtained after applying the double seasonal ARIMA model and the ARIMA-SS model. The results showed the superiority of the ARIMA-SS hybrid model over the doubled seasonal ARIMA model because it was more accurate in forecasting minimum temperatures based on the Mean Squared Error (MSE) scale. Minimum-temperature data is classified within the interval variables and not within the relative variables Ratio variables, because zero in them is not absolute and does not represent the absence of minimum temperature, 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 theory.

The researchers (9) presented a non-Gaussian homogeneity methodology for analyzing time series. This method is based on the modeling of a non-Gaussian state space and is especially relevant for time series that cannot be satisfactorily analyzed by conventional time series models. The researchers (7) used the SS model to forecast the wind speed, which is non-linear, which leads to uncertainty. To obtain the best initial parameters of the SS model, AR will be used to create the SS model structure to handle random uncertainty and improve forecasting. The researchers (10) using state space models as a modeling framework for analyzing environmental time series data. It is commonly used to model population dynamics including population metabolic dynamics. it has a long history in fishery stock assessment and has recently been used as a means of analyzing sparse Biodiversity data moreover, it has been a preferred approach in the movement environment for more than a decade and has been increasingly used with biological registration.

## 2.Materials and Methods

### Data and Framework of This Study

In this study, the forecasting methods represented by the ARIMA model and the ARIMA-SS hybrid model were used to forecast the minimum temperature of Mosul and Baghdad, since the data of the minimum temperature of Mosul and Baghdad are heterogeneous, and the data were aligned in two seasons (hot season and cold season) to obtain more homogeneous data to forecast the minimum temperatures for daily data for the period (1/11/2011-31/10/2022). The framework includes the following paragraphs.

1. Identify the appropriate ARIMA model using the Box-Jenkins methodology.
2. Build the ARIMA-SS hybrid model to use in improving forecasts.
3. Compare the methods used to determine which model will provide the best forecasts

**Autoregressive Integrated Moving Average (ARIMA) model** (11, 12)

The ARIMA (p,d,q) model is one of the most prominent nonstationary time series, it can be expressed as as follows:

$$\phi(B)(1-B)^d Z_t = \theta(B)a_t \tag{1}$$

$$\phi(B)W_t = \theta(B)a_t \tag{2}$$

$$W_t = (1-B)^d Z_t$$

$$(1-\phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p)W_t = (1-\theta_1 B - \theta_2 B^2 - \cdots - \theta_q B^q)a_t \tag{3}$$

$$W_t = \phi_1 W_{t-1} + \phi_2 W_{t-2} + \cdots + \phi_p W_{t-p} - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \cdots - \theta_q a_{t-q} + a_t \tag{4}$$

$\phi_k$ , $\theta_k$ Are the parameters of the autoregression and moving averege at step k, respectively, and reflect the effect of changing the time series variable ( $Z_{t-k}$ ) and the random variable ( $a_{t-k}$ ) at step k, respectively, p indicates the rank of the usual self-regression model, q stands for the rank of the usual moving circles, d represents the number of usual differences necessary to achieve stationarity, B is the back-shift operator, $a_t$ is the White Noise with zero arithmetic mean and $\sigma_a^2$ variance , that is $a_t \sim i.i.d.N(0, \sigma_a^2)$ also:

$\phi(B)=(1-\phi_1 B-\phi_2 B^2-\cdots-\phi_p B^p)$

$\theta(B)=(1-\theta_1 B-\theta_2 B^2-\cdots-\theta_q B^q)$

The model AR (p) and the model MA (q) are special cases of ARIMA (p,d,q). AR (p) can be expressed in the form (ARIMA (p,0,0), and MA (q) can be expressed in the form ARIMA (0, 0,q). Seasonal time series data can be generalized and written as a multiplicative seasonal ARIMA model or ARIMA (p,d,q) (P,D,Q)s (4, 13). It includes seasonal and non-seasonal parameters and their differences in general as follows.

$$\phi(B)\Phi(B)(1-B^S)^D(1-B)^d Z_t=\theta(B)\Theta(B)a_t \quad (5)$$

$$\phi(B)\Phi(B)W_t=\theta(B)\Theta(B)a_t \quad (6)$$

where $W_t=(1-B^S)^D(1-B)^d Z_t$, $\Phi(B)=(1-\Phi_1 B^s-\Phi_2 B^{2s}-\cdots-\Phi_P B^{Ps})$, and $\Theta(B)=(1-\Theta_1 B^s-\Theta_2 B^{2s}-\cdots-\Theta_Q B^{Qs})$.

$\Theta_k$ and $\Phi_k$ are the seasonal parameters of the autoregressive and the moving averages at step k, respectively, P indicate the rank of the seasonal autoregressive model, Q stands for the rank of the seasonal moving media, D represents the number of seasonal differences necessary to achieve stationarity, s represents the seasonal period.

We can build the ARIMA model following the Box-Jenkins methodology (14), which is an iterative methodology that contains four consecutive steps: recognition, estimation, diagnostic examination and forecasting, which is the most famous and important application of time series. Based on time series observations, we can forecast future observations. Accurate forecasting plays an important role in various applicable areas based on time intervals. Recognition is the first step in the Box-Jenkins methodology and involves stationaries of mean and variance of time series data. Stationarity is a fundamental concept in time series, which indicates that the common probability is not affected by the temporal displacement of all observations backward or forward by an amount of K time interval. The first step in analyzing time series is to plot the time series data and verify the stationarity of the series. After performing the first step of the Box-Jenkins methodology and determining the time series model, the second step that follows is to estimate the parameters of the time series model using the Maximum Likelihood Estimation (MLE) method, and then the third step, which is to verify the validity of the estimated model, the most prominent diagnostic tests is to test the significance of the parameters and ACF of the residuals series. The parameters must be significant using the t - test. The remainder series are independent random variables with similar distributions $a_t \sim i.i.d.N(0,\sigma_a^2)$. The estimated time series model after passing the previous three steps is ready to forecast the time series .After plotting the time series data and achieving stationarity the mean and variance as mentioned earlier, determining the p and q ranks is the most important procedure for determining the model. The use of ACF and PACF is useful for determining the appropriate time series model and the P and Q ranks as shown in Table 1.

**Table 1. ACF and PACF patterns according to time series models**

| Model | ACF | PACF |
|---|---|---|
| AR (*p*) | Gradually decays | Suddenly breaks after lag p |
| MA (*q*) | Suddenly breaks after lag q | Gradually decays |
| ARIMA (*p*,*q*) | radually decays but goes to zero after la | radually decays but goes to zero after la |

## State Space (SS) model

State space models are commonly used in the modeling framework for analyzing time series data because they directly represent the automatic temporal correlation. The SS model is a kind of pyramid model (15, 16). The use of the ARIMA model as a linear statistical model for modeling nonlinear time series data will lead to the presence of random uncertainty (17), which reduces the accuracy of the forecast. To address random uncertainty, the SS method is used for more accurate forecasting due to its good performance in forecasting time series (18-20).

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

$$Z_t=AZ_{t-1}+Bu_{t-1}+e_{1,t} \quad (7)$$

$$Y_t=CZ_t+e_{2,t} \quad (8)$$

$Z_t$ is a state vector of dimension m, $u_t$ is a finite input vector, $Y_t$ is a vector of output observations, $e_{1,t}$ and $e_{2,t}$ are independent white noise vectors, A, B and C are constant matrices. Based on the ARIMA model in equation(6), the SE equation, which is equation(7), and the OE equation, which is equation(8), can be written in the state space, respectively, in the following form (21)

$$Z_t = AZ_{t-1} + Bu_{t-1} + C'a_t \qquad (9)$$

$$Y_t = CZ_t \qquad (10)$$

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

g is the number of boundaries of the delayed series of the variable $Z_t$ , j is the number of boundaries of the delayed remainder series , $Z_t$ is the state vector with Dimension r , $u_t$ is the vector of the delayed residuals series , $a_t$ is the rounded vectors of the current residuals.

$$Z_t = \begin{bmatrix} Z_{1,t} & Z_{2,t} & \ldots & Z_{r,t} \end{bmatrix}' \text{ , } u_{t-1} = \begin{bmatrix} u_{1,t-1} & u_{2,t-1} & \cdots & u_{r,t-1} \end{bmatrix}' \text{ , } A = \begin{bmatrix} P_1 & P_2 & P_3 & \cdots & P_r \\ 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}_{r \times r}$$

Matrix A is the state transition matrix with Dimension ( $r \times r$ )

$$B = \begin{bmatrix} R_1 & R_2 & R_3 & \cdots & R_r \\ 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}_{r \times r}$$

The Matrix B is the commutative matrix of dimension ( $r \times r$ )

$$C = \begin{bmatrix} 1 & 0 & 0 & \ldots & 0 \end{bmatrix}_{1 \times r}$$

The matrix C is the transition matrix of the observations by Dimension $(1 \times r)$, $Y_t$ are the observed cyclic vectors representing the SS output series, $a_t$ are the cyclic vectors of the current remainders, $(P_1, P_2, \ldots, P_r)$ and $(R_1, R_2, \ldots, R_r)$ are the parameter values of the delayed series of the variable $Z_t$ and the delayed series of residuals respectively in the ARIMA model equation. Equation(9) and equation(10) are complex in application and for simplification they can be reformulated as follows (4, 22).

$$Z_t = AZ_{t-1} + C'a_t \qquad (11)$$

$$\hat{y}_t = CZ_t \qquad (12)$$

$Z_t$ Is the state vector in dimension m, A the state transition matrix in dimension (m x m), C the observation transition matrix in dimension (1 x m), $\hat{y}_t$ the observed cyclic vectors representing the SS output Series, m is the number of the boundaries of the delayed series of the variable $Z_t$ and the sum of the series of delayed residuals of the variable $a_t$ on the right-hand side of the ARIMA data model equation after simplifying it and keeping only $Z_t$ on the left-hand side.

$$Z_t = \begin{bmatrix} Z_{1,t} & Z_{2,t} & \cdots & Z_{m,t} \end{bmatrix}' \qquad (13)$$

$$A = \begin{bmatrix} P_1 & P_2 & P_3 & L & P_m \\ 1 & 0 & 0 & L & 0 \\ 0 & 1 & 0 & L & 0 \\ M & M & M & O & M \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}_{m \times m} \qquad (14)$$

$$C = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \end{bmatrix}_{1 \times m} \qquad (15)$$

*$(P_1, P_2, \ldots, P_m)$* Is all the parameter values of the delayed series of the variable $Z_t$ *and all* the delayed remainder series of the variable $a_t$ on the right side of the ARIMA data model equation after simplifying it and keeping only $Z_t$ on the left side. The Matrix A, the row vector C and the variables on the right-hand side of the ARIMA data model equation will determine the input variables and the SS structure used in this research. The time stratified method (TS) was used, which is an analytical method that aligns the data chronologically according to seasonal influences, which are clearly manifested as influences on the behavior of the time series and the behavior of forecasting results. Time Stratified can be applied to different time series in the event that they include recurring seasonal time trends with the same context and influence and works to reach more homogeneous data

than the aggregate data and thus obtain more accurate results (23, 24). The mean squared error (MSE) is one of the measures used to express the forecasting accuracy of the data and the mean Squared Error (MSE) equation(8) can be written as follows:

$$MSE = \frac{\sum_{t=1}^{n} a_t^{\,2}}{n} \qquad (16)$$

$a_t = y_t - \hat{y}_t$ ; $t = 1, 2, \ldots, n$ , n: number of observations, $y_t$ real observation at time t, $\hat{y}_t$ the forecasting value for the observation y at the time t, $a_t$ is the residuals series at the time t.

## 3.Results and Conclusions

### Data and Framework Used in this Study

In this study, the forecasting methods represented by the ARIMA model and the ARIMA-SS hybrid model were used to forecast the minimum temperature of Mosul and Baghdad. The minimum temperature data of Mosul and Baghdad are heterogeneous, therefore, the TS method was used to obtain more homogeneous data for forecasting minimum temperatures for daily data for the period (1/11/2011-31/10/2022) (23, 24), which is by aligning the data for the hot season and the cold season separately, as the data for the hot season for consecutive months (may - October) for the years within the period (2012-2022). And the data of the cold season for consecutive months (November– April) for the years within the period (2011-2022). 1656 obseravtion were used for training for the warm season for both Mosul and Baghdad. 1629 observations were used for training for the cold season for both Mosul and Baghdad. 368 Test observations were used for the warm season for both Mosul and Baghdad and for the cold season 362 Test observations for both Mosul and Baghdad were used to forecast the minimum temperature. These data were modeled using the double seasonal ARIMA 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 six months, i.e. (s=6). The working steps will be as follows:

1. Build an ARIMA model based on the Box-Jenkins methodology.
2. Take the right-hand side of the ARIMA models as input to the SS model.
3. Initialize each of the following variables:

a. The target variable $y_t$ , which isthe left end of the ARIMA model. $Z_t$

b. The variable $X_t$ is the right-hand end of the ARIMA model except for the current residuals series. $a_t$

c. The variable $a_t$ is the current residuals series.

### ARIMA Model

The Box-Jenkins methodology in which the first step is to identify the model, one of the basic conditions of which is that the data is stationary. To detect the degree of (stationarity) of the data is done by plotting time series data, and it is also possible to use the ACF and PACF functions . To achieve stationarity, the usual and seasonal sequential differences are taken on the assumption that the variation is stationary, since the minimum temperature is a natural phenomenon that is not subject to ambient interference. One regular difference and one seasonal difference were taken, respectively, and stationarity was tested at each stage, and it turned out that the data was completely stationary after taking the two differences above. The determination of the model is also through the plotting of the functions ACF and PACF. Figure 1 shows ACF and PACF of the city of Mosul the hot season after stationarity, i.e. when (D=1,d=1, s=6).

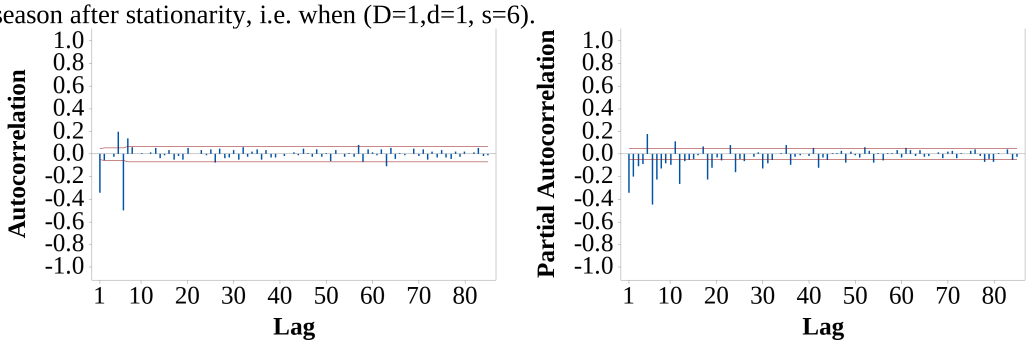


**Figure 1. ACF and PACF for the city of Mosul is the hot season when (D=1, d=1, s=6)**

It can be seen from Figure 1 that ACF shows a break after the usual Lag 2 and after the seasonal Lag 6, while PACF shows a decaying pattern in the normal and seasonal phases, so the best ARIMA model for the city of the hot season connector is ARIMA $(0,1,2)(0,1,1)_6$. Table 2. It shows the parameter values of the ARIMA model for the city of Mosul hot season and significant indicators..

**Table 2. Parameter values for the ARIMA $(0,1,2)(0,1,1)_6$ model**

| Parameter | Calc. t |
|---|---|
| 0.4861 | 19.83 |
| 0.1675 | 6.82 |
| 0.9817 | 14664.08 |

From Table 2. It turns out that all the values of the estimated parameters are significant (p-value is less than the level of significance 0.05), which are the parameters belonging to the ARIMA $(0,1,2)(0,1,1)_6$. model. The ACF function of the remainders was also tested and it turned out that they match the conditions of a good model. Thus, the ARIMA $(0,1,2)(0,1,1)_6$. will have successfully passed the diagnostic tests. The ARIMA $(0,1,2)(0,1,1)_6$. model of the minimum temperature of the city of Mosul for the hot season can be expressed by the following formula:

Let $W_t$ is the stationary series after taking difference procedures i.e. $W_t = (1-B)(1-B^6)Z_t$, and

$$(1-B)(1-B^6)Z_t = (1-\Theta_1 B^6)(1-\theta_1 B - \theta_2 B^2)a_t \quad (17)$$

After multiplying the brackets and making mathematical simplifications, we get the following:

$$Z_t = Z_{t-1} + Z_{t-6} - Z_{t-7} + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \Theta_1 a_{t-6} + \theta_1 \Theta_1 a_{t-7} + \theta_2 \Theta_1 a_{t-8} \quad (18)$$

After compensating for the values of the parameters found in table2. In equation(18), the equation of the ARIMA $(0,1,2)(0,1,1)_6$ model becomes as follows:

$$Z_t = Z_{t-1} + Z_{t-6} - Z_{t-7} + a_t - 0.4861 a_{t-1} - 0.1675 a_{t-2} - 0.9817 a_{t-6} + 0.4472 a_{t-7} + 0.1644 a_{t-8} \quad (19)$$

**Table 3. The MSE scale for Mosul city, the hot season for both training and testing periods for ARIMA model**

| Training Period | Testing Period |
|---|---|
| 4.4404 | 4.1050 |

The Box-Jenkins methodology in which the first step is to identify the model, one of the basic conditions of which is that the data be stationary. To detect the degree of stationarity of the data, time series data is plotted, and it is also possible to use the ACF and PACF functions . To achieve stationarity, the usual and seasonal sequential differences are taken on the assumption that the variation is stationary, since the minimum temperature is a natural phenomenon that is not subject to ambient interference. One regular difference and one seasonal difference were taken, respectively, and stationarity was tested at each stage, and it turned out that the data was completely stationary after taking the two differences above. The determination of the model is also through the plotting of the functions ACF and PACF. Figure 2. ACF and PACF for Mosul show the cold season after stationarity, i.e. when (D=1,d=1, s=6)

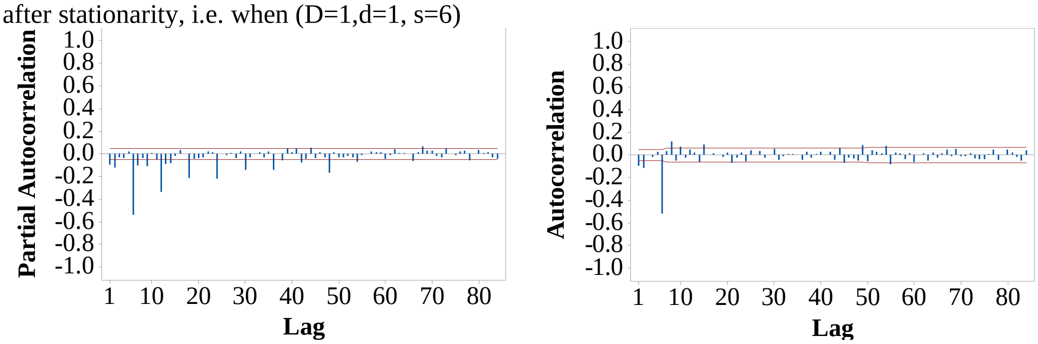


**Figure2. ACF and PACF for the city of Mosul, Cold Season (D=1,d=1,s=6)**

Figure 2 shows that ACF shows a break after the usual Lag 3 and after seasonal Lag 18, while PACF shows a break after the usual Lag 1 but shows decay in the seasonal phase, so the best ARIMA model for Mosul in the cold season is ARIMA $(1,1,2)(0,1,3)_6$. . Table 4. Shows the parameter values of the ARIMA $(1,1,2)(0,1,3)_6$. model for the city of Mosul cold season and significant indicators.

**Table 4. Parameter Values for the ARIMAmodel** $(1,1,2)(0,1,3)_6$

| Type | Parameter | Calc. t | p-value |
|---|---|---|---|
| $\phi_1$ | 0.6127 | 16.72 | 0.000 |
| $\theta_1$ | 0.8332 | 21.19 | 0.000 |
| $\theta_2$ | 0.0695 | 2.15 | 0.031 |
| $\Theta_1$ | 0.7818 | 951.10 | 0.000 |
| $\Theta_2$ | 0.1094 | 4.45 | 0.000 |
| $\Theta_3$ | 0.0922 | 3.76 | 0.000 |

From Table 4 it is clear that all the values of the estimated parameters are significant (p - value is less than the level of significance 0.05), which are the parameters belonging to the ARIMA $(1,1,2)(0,1,3)_6$. model . The ACF function of the remainders was also tested and it turned out that they match the conditions of a good model. Thus, the A ARIMA $(1,1,2)(0,1,3)_6$. model has successfully passed the diagnostic tests. The ARIMA $(1,1,2)(0,1,3)_6$. model of the minimum temperature of Mosul in the cold season can be expressed by the following formula:

$$(1-B)(1-B^6)(1-\phi_1 B)Z_t = (1-\theta_1 B-\theta_2 B^2)(1-\Theta_1 B^6-\Theta_2 B^{12}-\Theta_3 B^{18})a_t \qquad (20)$$

After multiplying the brackets and making mathematical simplifications, we get the following:

$$Z_t = (\phi_1+1)Z_{t-1}-\phi_1 Z_{t-2}+Z_{t-6}-(\phi_1+1)Z_{t-7}+\phi_1 Z_{t-8}+a_t-\theta_1 a_{t-1}-\theta_2 a_{t-2}-\Theta_1 a_{t-6}$$
$$+\theta_1\Theta_1 a_{t-7}+\theta_2\Theta_1 a_{t-8}-\Theta_2 a_{t-12}+\theta_1\Theta_2 a_{t-13}+\theta_2\Theta_2 a_{t-14}-\Theta_3 a_{t-18}+\theta_1\Theta_3 a_{t-19}+\theta_2\Theta_3 a_{t-20} \qquad (21)$$

After compensating for the values of the parameters found in table 3 in equation(21), the ARIMA $(1,1,2)(0,1,3)_6$ model becomes as follows:

$$Z_t = 1.6127Z_{t-1}-0.6127Z_{t-2}+Z_{t-6}-1.6127Z_{t-7}+0.6127Z_{t-8}+a_t-0.8332a_{t-1}$$
$$-0.0695a_{t-2}-0.7818a_{t-6}+0.6514a_{t-7}+0.0543a_{t-8}-0.1094a_{t-12}+0.0912a_{t-13}$$
$$+0.0076a_{t-14}-0.0922a_{t-18}+0.0768a_{t-19}+0.0064a_{t-20} \qquad (22)$$

Table 5. Below are the MSE values of the ARIMA $(1,1,2)(0,1,3)_6$ model's forecasts for the minimum temperature of Mosul in the cold season for the training and testing periods.

**Table 5. The MSE scale for Mosul city is the cold season for both training and testing periods**

| Training Period | Testing Period |
|---|---|
| 4.0958 | 4.1610 |

The determination of the model is also through the drawing of the functions ACF and PACF. Figure 3. ACF and PACF for the city of Baghdad show the hot season after stationarity.

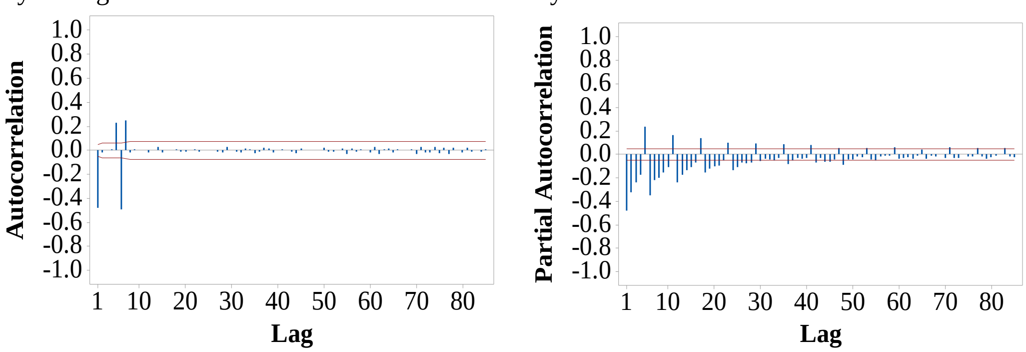


**Figure 3. ACF and PACF for the city of Baghdad, hot season (D=1,d=1, s=6).**

It is evidenced by Figure 3. The ACF shows a break after the usual Lag 1 and after the seasonal Lag 6, while the PACF shows a decay in the usual and seasonal phases, so the best ARIMA model for Baghdad in the hot season is ARIMA$(0,1,1)(0,1,1)_6$. Table 6. It presents the parameter values of the ARIMA$(0,1,1)(0,1,1)_6$ model for the city of Baghdad, the hot season and significance indicators.

**Table 6. Parameter Values for the ARIMA $(0,1,1)(0,1,1)_6$ model**

| Type | Parameter | Calc. t | p-value |
|---|---|---|---|
| $\theta_1$ | 0.8638 | 70.58 | 0.000 |
| $\Theta_1$ | 0.9854 | 23100.81 | 0.000 |

From Table 6. It turns out that all the values of the estimated parameters are significant (p - value is less than the level of 0.05), which are the parameters belonging to the ARIMA(0,1,1)(0,1,1)$_6$ model. The ACF function of the remainders was also tested and it turned out that they match the conditions of a good model. Thus, the ARIMA(0,1,1)(0,1,1)$_6$ model has successfully passed the diagnostic tests. The ARIMA(0,1,1)(0,1,1)$_6$ model of the minimum temperature of the city of Baghdad in the hot season can be expressed by the following formula:

$$(1-B)(1-B^6)Z_t = (1-\Theta_1 B^6)(1-\theta_1 B)a_t \qquad (23)$$

After multiplying the brackets and making mathematical simplifications, we get the following:

$$Z_t = Z_{t-1} + Z_{t-6} - Z_{t-7} + a_t - \theta_1 a_{t-1} - \Theta_1 a_{t-6} + \theta_1 \Theta_1 a_{t-7} \qquad (24)$$

After compensating for the values of the parameters found in table 6 in equation(24), the ARIMA(0,1,1)(0,1,1)$_6$ model becomes as follows:

$$Z_t = Z_{t-1} + Z_{t-6} - Z_{t-7} + a_t - 0.8638a_{t-1} - 0.9854a_{t-6} + 0.8512a_{t-7} \qquad (25)$$

Table 7. The following shows the MSE values for ARIMA(0,1,1)(0,1,1)$_6$ model forecasts for the minimum temperature of the city of Baghdad hot season for the training and testing periods.

**Table 7. The MSE scale for Baghdad city is the hot season for training and testing periods for ARIMA**

| Training Period | Testing Period |
|---|---|
| 32.5704 | 20.6798 |

The determination of the model is also through the drawing of the functions ACF and PACF. Figure 4. ACF and PACF for the city of Baghdad show the cold season after stationarity.

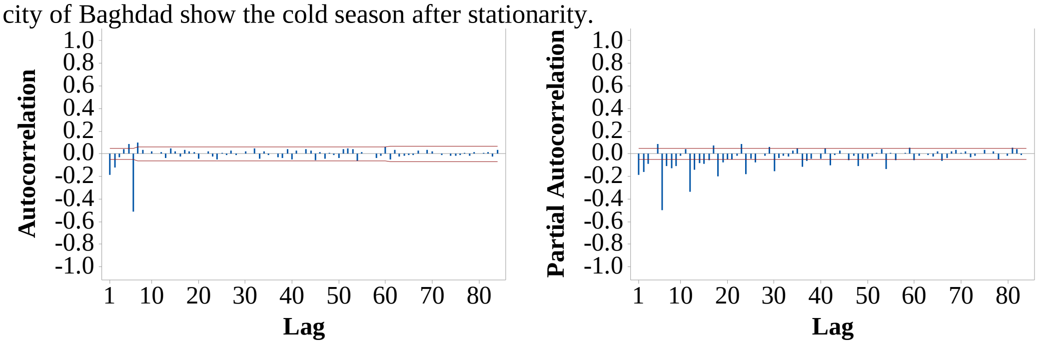


**Figure 4. ACF and PACF for the city of Baghdad cold season (D=1,d=1, s=6)**

It is evidenced by Figure 4. The ACF shows a break after the usual Lag 2 and after the seasonal Lag 6, while the PACF shows a break after the usual Lag 1 but shows a decay in the seasonal phase, so the best ARIMA model for Baghdad city in the cold season is ARIMA(0,1,2)(0,1,1)$_6$. Table 8 . Displays parameter values of ARIMA(0,1,2)(0,1,1)$_6$ model for Baghdad city of cold season and significant indicators.

**Table 8. Parameter Values for the ARIMA $(1,1,2)(0,1,1)_6$ model**

| Type | Parameter | Calc. t | p-value |
|---|---|---|---|
| $\phi_1$ | 0.4464 | 7.95 | 0.000 |
| $\theta_1$ | 0.7571 | 12.94 | 0.000 |
| $\theta_2$ | 0.0986 | 2.32 | 0.020 |
| $\Theta_1$ | 0.9806 | 7094.52 | 0.000 |

From Table 8. It turns out that all the values of the estimated parameters are significant (p - value is less than the level of significance 0.05), which are the parameters belonging to the ARIMA(1,1,2)(0,1,1)$_6$. model . The ACF function of the remainders was also tested and it turned out that they match the conditions of a good model. Thus, the ARIMA(1,1,2)(0,1,1)$_6$.

model has successfully passed the diagnostic tests. The ARIMA(1,1,2)(0,1,1)$_6$. model of the minimum temperature of the city of Baghdad cold season can be expressed by the following formula:

$$(1-B)(1-B^6)(1-\phi_1 B)Z_t = (1-\Theta_1 B^6)(1-\theta_1 B-\theta_2 B^2)a_t \quad (26)$$

After multiplying the brackets and making mathematical simplifications, we get the following:

$$Z_t = (\phi_1+1)Z_{t-1} - \phi_1 Z_{t-2} + Z_{t-6} - (\phi_1+1)Z_{t-7} + \phi_1 Z_{t-8} + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2}$$
$$-\Theta_1 a_{t-6} + \theta_1\Theta_1 a_{t-7} + \theta_2\Theta_1 a_{t-8} \quad (27)$$

After compensating for the values of the parameters found in table 7 in equation(27), the ARIMAmodel $(1,1,2)(0,1,1)_6$ becomes as follows:

$$Z_t = 1.4464Z_{t-1} - 0.4464Z_{t-2} + Z_{t-6} - 1.4464Z_{t-7} + 0.4464Z_{t-8} + a_t - 0.7574a_{t-1}$$
$$-0.0986a_{t-2} - 0.9806a_{t-6} + 0.7424a_{t-7} + 0.0967a_{t-8} \quad (28)$$

Table 9 below shows the MSE values of the ARIMA $(1,1,2)(0,1,1)_6$ model temperature forecasts for the city of Baghdad cold season for the training and test periods.

**Table 9. The MSE scale for Baghdad city is the cold season for training and testing periods for ARIMA**

| Training Period | Testing Period |
|---|---|
| 9.3753 | 11.8174 |

When forecasting using the ARIMA model for the training and testing periods, respectively, for the city of Mosul during the hot season, the MSE was equal to 4.4404 and 4.1050, respectively, as shown in Table 3. While for the city of Mosul during the cold season, the MSE was equal to 4.0958 and 4.1610, respectively, as in Table 5. For the city of Baghdad, the hot season MSE was 32.5704 and 20.6798 respectively in Table 7. For the city of Baghdad, the cold season MSE was 9.3753 and 11.8174, respectively, as shown in Table 9.

**ARIMA-SS Hybrid Approach**

The input of the SS model is the variables on the right-hand side of the equation of the doubled seasonal ARIMA model, excluding the variable $a_t$ after keeping $Z_t$ on the left-hand side. When taking the variables on the right side of the equation of the doubled seasonal ARIMA model regardless of the parameters and signals, the hybrid method is called ARIMA-SS without parameters, and when taking the variables on the right side of the equation of the doubled seasonal ARIMA model with parameters and signals (i.e. multiplying variables by parameters and signals), the hybrid method is called ARIMA-SS with parameters. The framework for completing the process of forecasting minimum temperatures using the ARIMA-SS hybrid model in the Matlab program is summarized as follows.

1. Modeling of minimum temperature time series data using the ARIMA model.
2. The input variables of the ARIMA-SS hybrid model are the same as the variables and their parameters at the right end of the final estimated ARIMA models, which include parameters, self-regression variables, moving averages, as well as the remainder variable.
3. Initialize the input variables and the target variable $Y_t$ , which is the variable $Z_t$ (the original series), the remainder and the two matrices referred to in equations 11 and 12 and insert them into (Matlab).
4. Build the ARIMA-SS hybrid model using the data configured in the previous step.
5. Calculation of the MSE error criterion.

When using the ARIMA-SS hybrid method to forecast the temperatures of Mosul in the hot season, equation(19) was the variables that determined the structure of the SS model. Table 10 below shows the MSE values of the ARIMA-SS hybrid model temperature forecasts for Mosul city during the hot season for the training and testing period .

**Table 10. The MSE scale for Mosul city is the hot season for training and testing periods for ARIMA-SS model.**

| | Training | Testing |
|---|---|---|
| **Without Parameters** | 1.5199 | 1.2575 |
| **With Parameters** | 1.4478 | 1.1744 |

The MSE value in Table 10 when using the ARIMA-SS hybrid method for forecasting the training period without parameters and with parameters, respectively, was 1.5199 and 1.4478, respectively, which is smaller than the MSE for the training period when using the ARIMA model for forecasting. For the test period without parameters and with parameters respectively, the MSE value was 1.2575 and 1.1744, respectively, which is smaller than the MSE for the test period when using the ARIMA model. Thus, it is clear that the use of the ARIMA-SS hybrid method for forecasting dealt with uncertainty and had more accuracy in forecasting than the ARIMA model, when using which the forecast results were not good, because the data in the hot season are more scattered and less homogeneous. Figure 5 and Figure 6 below show the extent of correspondence and harmony between the original time series variable and the forecast series for the training and testing periods respectively for the hot season of Mosul using the ARIMA -SS hybrid model with and without parameters.

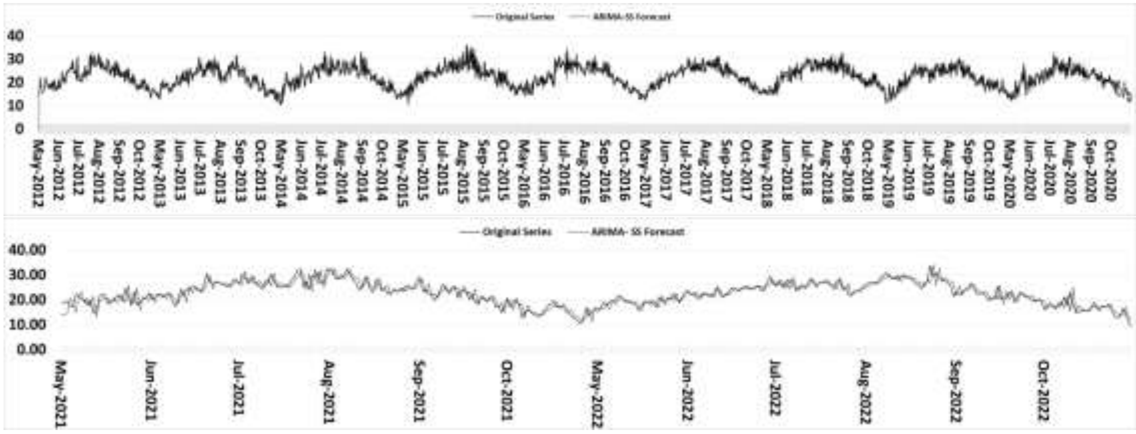


**Figure 5. The Fitting figure between the original time series and the forecast series for the training and testing periods respectively for the hot season of Mosul using the ARIMA-SS hybrid model without parameters**

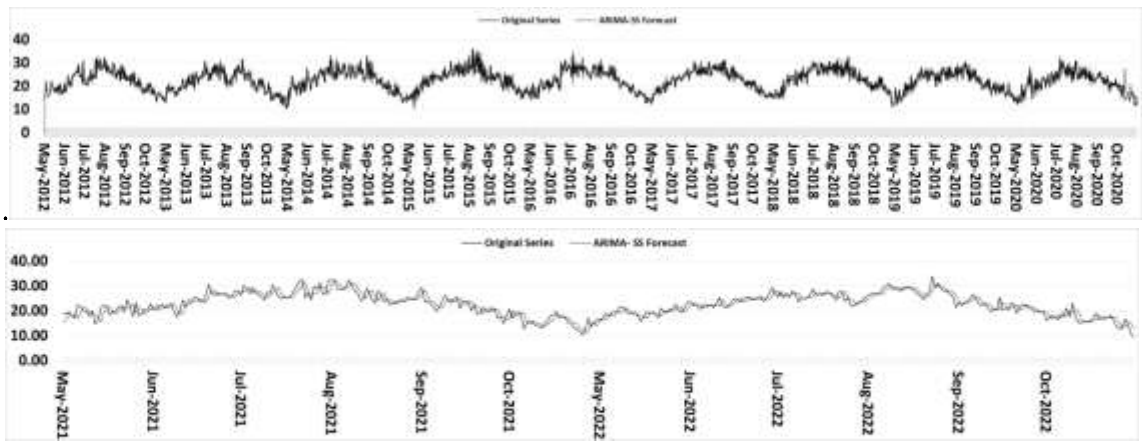


**Figure 6. The Fitting figure is formed between the original time series and the forecasting series for the training and testing periods respectively for the hot season of the conductor using the ARIMA-SS hybrid model with the parameters**

It can be seen from Figure 5 and Figure 6 that the ARIMA-SS hybrid model performed well in forecasting the minimum temperatures of Mosul in the hot season due to the convergence of the original time series with the forecasting series of the ARIMA -SS hybrid model without parameters and with parameters for both training and testing periods. When using the ARIMA-SS hybrid method to forecast the temperatures of Mosul in the cold season, equation(22) was the variables that determined the structure of the SS model .Table 11 below shows the MSE values of the ARIMA-SS hybrid model temperature forecasts for Mosul cold season for the training and test periods.

**Table 11. The MSE scale for Mosul city is the cold season for both training and testing periods for ARIMA-SS**

| Training Period | | Testing period | |
|---|---|---|---|
| **Without Parameters** | 6.1452 | **Without Parameters** | 3.0313 |
| **With Parameters** | 6.1426 | **With Parameters** | 2.8891 |

The MSE in Table 11 when using the ARIMA-SS hybrid method for forecasting the training period without parameters and with parameters, respectively, was 6.1452 and 6.1426, respectively, which is greater than the MSE for the training period when using the ARIMA model for forecasting. For the test period without parameters and with parameters respectively, the MSE values were 3.0313 and 2.8891, respectively, which is smaller than the MSE for the test period when using the ARIMA model. Thus, it becomes clear that the use of the ARIMA-SS hybrid method for forecasting in the training period had less accuracy in forecasting than the ARIMA model, which forecast results when using it were good, because the data for the city of Mosul for the cold season are less scattered and more homogeneous, while in the test period the forecast results when using the ARIMA-SS hybrid method for forecasting were more accurate than the ARIMA model. Figure 7 and figure 8 below show the extent of correspondence and harmony between the original time series variable and the forecast series for the training and testing periods respectively for the cold season of Mosul using the ARIMA-SS hybrid model with and without parameters.

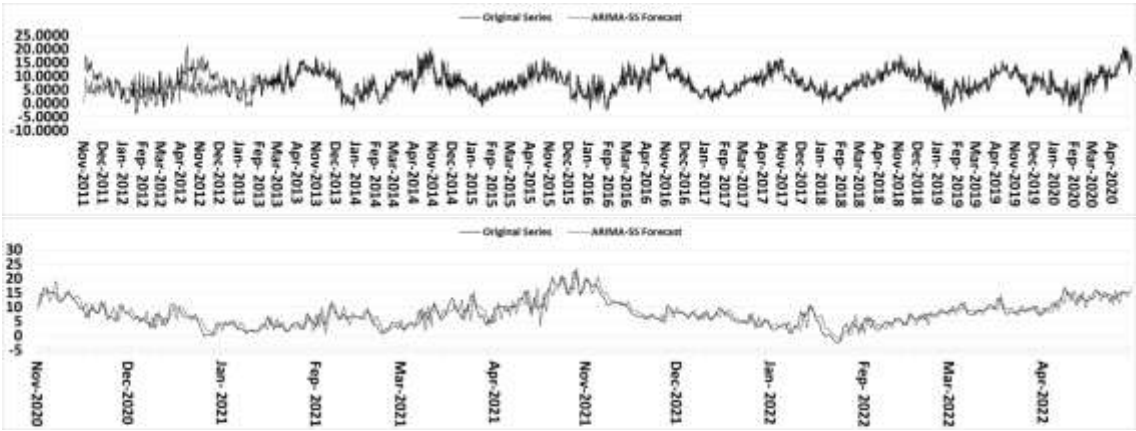


**Figure 7. The Fitting figure between the original time series and the forecasting series of the training and testing periods respectively for the cold season of Mosul using the ARIMA-SS hybrid model without parameters.**

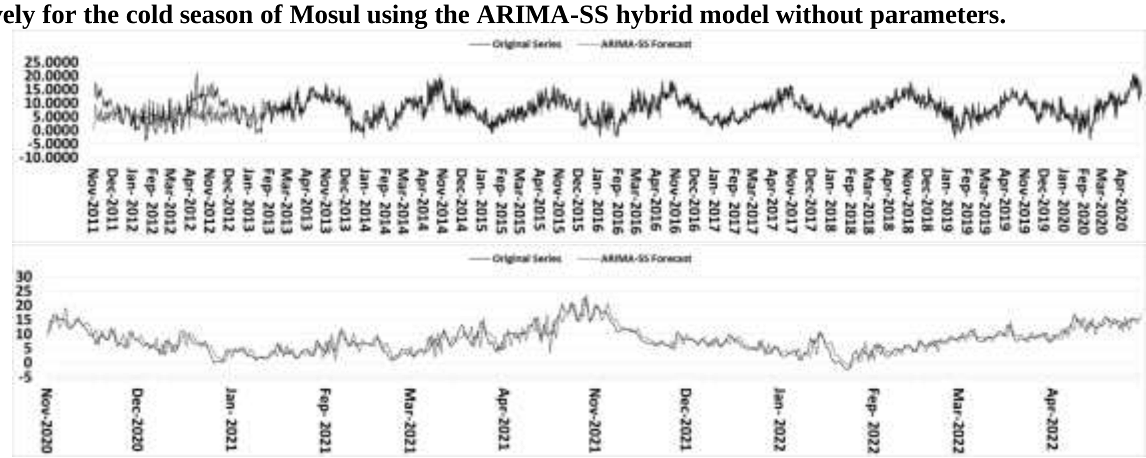


**Figure 8. The Fitting figure is formed between the original time series and the forecasting series for the training and testing periods respectively for the cold season of Mosul using the ARIMA-SS hybrid model with parameters.**

It is shown in Figure 7 and Figure 8 that the ARIMA-SS hybrid model performed well in forecasting the minimum temperatures of Mosul in the cold season due to the convergence of the original time series with the forecasting series of the ARIMA-SS hybrid model without parameters and with parameters for both training and testing periods .

When using the ARIMA-SS hybrid method to forecast the temperatures of the city of Baghdad in the hot season, equation(25) was the variables that determined the structure of the SS model. Table 12 below shows the MSE values for the ARIMA-SS hybrid model temperature forecasts for Baghdad during the hot season for the training and testing periods

**Table 12. The MSE scale for Baghdad city is the hot season for both training and testing periods for ARIMA-SS**

| Training Period | | Testing period | |
|---|---|---|---|
| **Without Parameters** | 1.5712 | **Without Parameters** | 6.7694 |
| **With Parameters** | 1.5875 | **With Parameters** | 6.7697 |

The MSE value in Table 12 when using the ARIMA-SS hybrid method for forecasting the training period without parameters and with parameters, respectively, was 1.5712 and 1.5875, respectively, which is smaller than the MSE for the training period when using the ARIMA model for forecasting. For the test period without parameters and with parameters respectively, the MSE value was 6.7694 and 6.7697, respectively, which is smaller than the MSE for the test period when using the ARIMA model for forecasting. Thus, it is clear that the use of the ARIMA-SS hybrid method for forecasting dealt with uncertainty and had more accuracy in forecasting than the ARIMA model, when using which the forecast results were not good, because the data in the hot season are more scattered and less homogeneous. Figure 9 and Figure 10 below show the extent of correspondence and harmony between the original time series variable and the forecast series for the training and testing periods respectively for the hot season of Baghdad using the ARIMA-SS hybrid model with and without parameters.

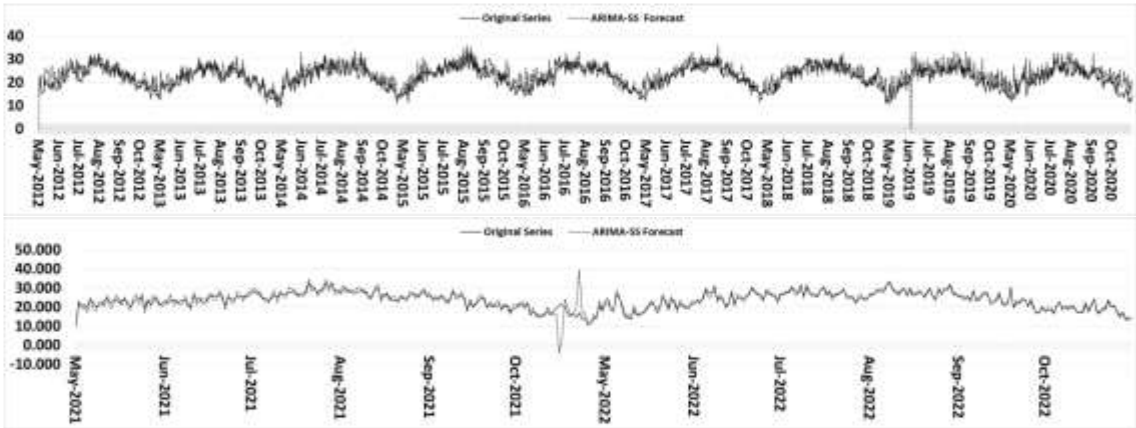


**Figure 9. The Fitting figure between the original time series and the forecasting series for the training and testing periods respectively for the Baghdad hot season using the ARIMA-SS hybrid model without parameters.**

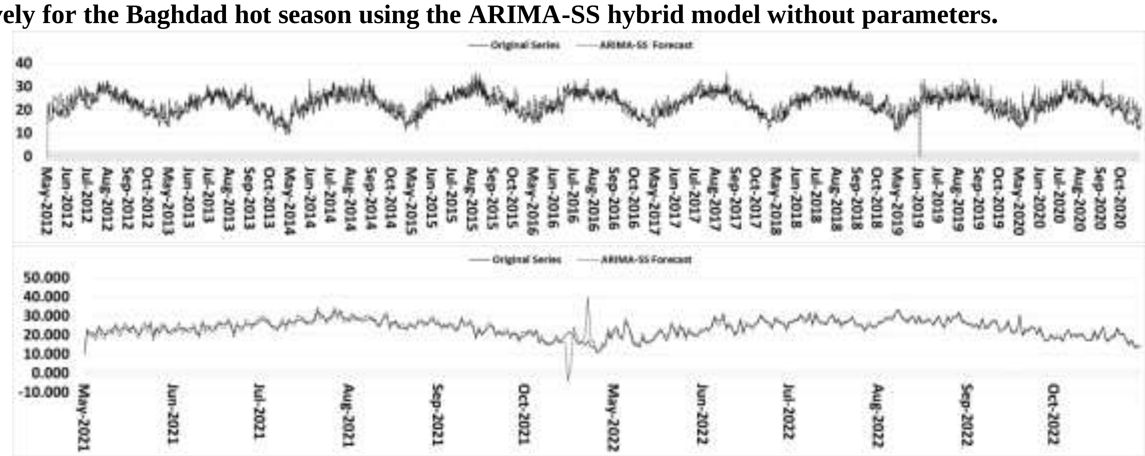


**Figure 10. The Fitting figure between the original time series and the forecasting series for the training and testing periods respectively for the Baghdad hot season using the ARIMA-SS hybrid model with parameters.**

It is shown in Fig. 9 and fig. 10 that the ARIMA-SS hybrid model performed well in forecasting the minimum temperatures of Baghdad during the hot season due to the convergence of the original time series with the forecasting series of the ARIMA-SS hybrid model without parameters and with parameters for both training and testing periods .

When using the ARIMA-SS hybrid method to forecast the temperatures of the city of Baghdad in the cold season, equation(28) was the variables that determined the structure of the SS model .Table 13 below shows the MSE values of the ARIMA-SS hybrid model temperature forecasts for the city of Baghdad cold season for the training and test periods.

**Table 13. MSE scale for Baghdad city is the cold season for both training and testing for ARIMA-SS**

| **Training Period** | | **Testing period** | |
|---|---|---|---|
| **Without Parameters** | 20.9981 | **Without Parameters** | 6.4420 |
| **With Parameters** | 21.0414 | **With Parameters** | 6.4414 |

The MSE value in Table 13 when using the ARIMA-SS hybrid method for forecasting the training period without parameters and with parameters, respectively, was 20.9981 and 21.0414, respectively, which is greater than the MSE for the training period when using the ARIMA model for forecasting. For the test period without parameters and with parameters respectively, the MSE values were 6.4420 and 6.4414, respectively, which is smaller than the MSE for the test period when using the ARIMA model. Thus, it becomes clear that the use of the ARIMA-SS hybrid method for forecasting in the training period had a lower accuracy in forecasting than the ARIMA model, which forecast results when using it were good, because the data for the city of Baghdad for the cold season are less scattered and more homogeneous, while in the test period the forecast results when using the ARIMA-SS hybrid method for forecasting were more accurate than the ARIMA model. Figure 11 and Figure 12 below show the extent of congruence and harmony between the original time series variable and the forecast series for the training and testing periods respectively for the cold season of the city of Baghdad using the ARIMA-SS hybrid model with and without parameters.

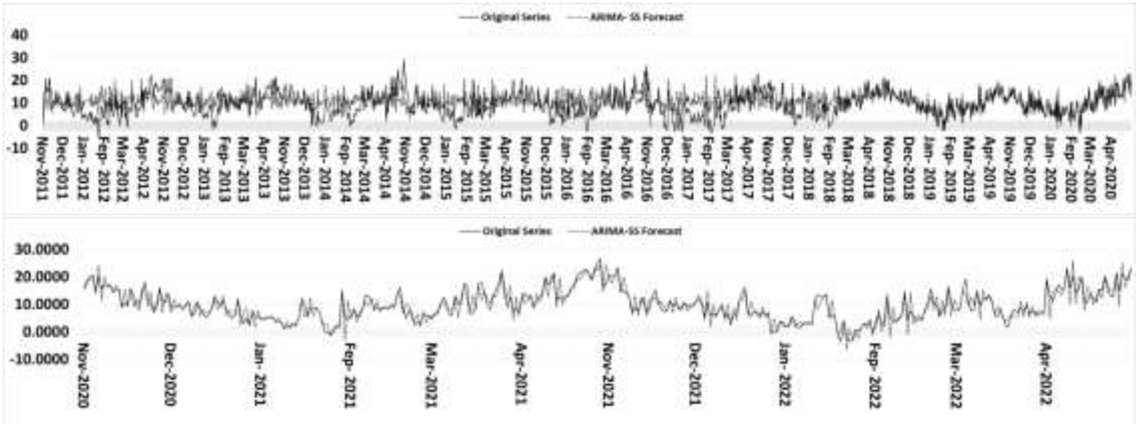


**Figure 11. The Fitting figure between the original time series and the forecasting series for the training and testing periods respectively for the cold season of Baghdad using the ARIMA-SS hybrid model without parameters.**

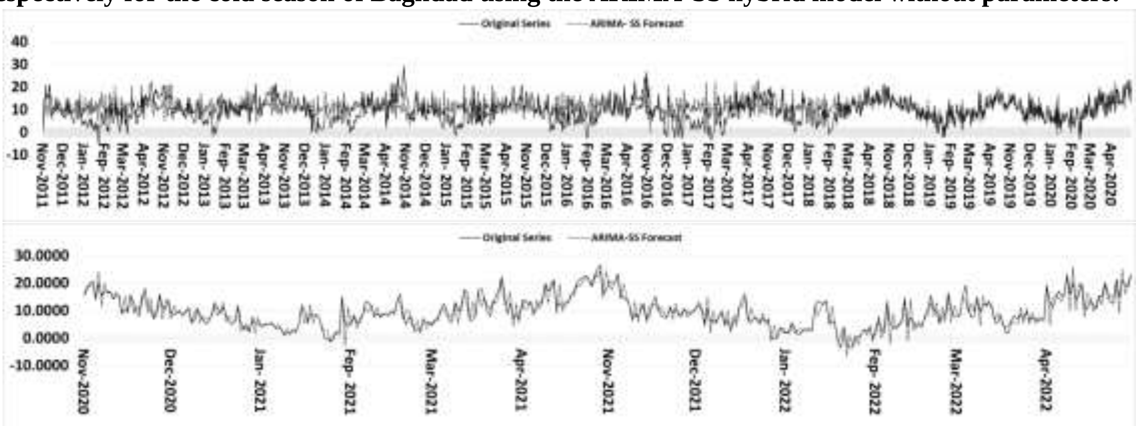


**Figure 12. The Fitting figure between the original time series and the forecasting series for the training and testing periods respectively for the cold season of Baghdad using the ARIMA-SS hybrid model with parameters.**

It is shown in Figure 11 and Figure 12 that the ARIMA-SS hybrid model performed well in forecasting the minimum temperatures of the city of Baghdad in the cold season due to the convergence of the original time series with the forecasting series of the ARIMA-SS hybrid model without parameters and with parameters for both training and testing periods.

## Conclusions

The ARIMA-SS hybrid method was used to improve the accuracy of minimum-temperature forecasting. Two sets of minimum-temperature data were used for the cities of Mosul and Baghdad. The results showed that ARIMA and ARIMA-SS hybrid models were effective. However, the MSE forecast results indicated that the ARIMA-SS hybrid model was the most effective tool for improving the accuracy of minimum-temperature forecasts, especially after aligning the data using Time Stratified (TS) to eliminate the problem of heterogeneity. But in the cold season for the cities of Mosul and Baghdad, for the training period, the ARIMA model had more accuracy in forecasting than the ARIMA-SS hybrid model, because the data in the cold season were less scattered and more homogeneous. Also, the use of parameters or not in the configuration of the input structure of the hybrid model did not have a significant difference in the efficiency and accuracy of forecasts equally. The advantages of the hybrid model were to deal with the problem of uncertainty that is caused by the non-linearity of the data. Therefore, it is possible to conclude that the use of the ARIMA-SS hybrid model will result in better forecasting accuracy for the minimum temperature compared to the forecasting performance of the traditional ARIMA model.

## References

1. Wei, W.W. and D.O. Stram, *Disaggregation of time series models.* Journal of the Royal Statistical Society: Series B (Methodological), 1990. **52**(3): p. 453-467.
2. Box, G.E. and G.M. Jenkins, *Time series analysis. Forecasting and control.* Holden-Day Series in Time Series Analysis, 1976.
3. Cadenas, E. and W. Rivera, *Wind speed forecasting in three different regions of Mexico, using a hybrid ARIMA–ANN model.* Renewable Energy, 2010. **35**(12): p. 2732-2738.

4. Gould, P.G., et al., *Forecasting time series with multiple seasonal patterns.* European Journal of Operational Research, 2008. **191**(1): p. 207-222.
5. Donner, R.V. and S.M. Barbosa, *Nonlinear Time Series Analysis in the Geosciences: Applications in Climatology, Geodynamics and Solar-Terrestrial Physics*. Vol. 112. 2008: Springer.
6. Wang, L., et al., *Design of sparse Bayesian echo state network for time series prediction.* Neural Computing and Applications, 2021. **33**: p. 7089-7102.
7. Shukur, O.B. and M.H. Lee, *Daily Wind Speed Forecasting Through Hybrid AR-ANN and AR-KF Models.* Jurnal Teknologi, 2015. **72**(5): p. 89–95.
8. Hyndman, R.J. and A.B. Koehler, *Another look at measures of forecast accuracy.* International journal of forecasting, 2006. **22**(4): p. 679-688.
9. Takanami, T., et al., *Multivariate time series model to estimate arrival times of S waves.* Methods and Applications of Signal Processing in Seismic Network Operations, 2003: p. 13-39.
10. Auger-Méthé, M., et al., *An introduction to state-space modeling of ecological time series.* arXiv preprint arXiv:2002.02001, 2020: p. 1-61.
11. Wei, W.W., *Time series analysis: univariate and multivariate.* Methods. Boston, MA: Pearson Addison Wesley, 2006.
12. Taneja, K., et al., *Time series analysis of aerosol optical depth over New Delhi using Box–Jenkins ARIMA modeling approach.* Atmospheric Pollution Research, 2016. **7**(4): p. 585-596.
13. Shadab, A., S. Said, and S. Ahmad, *Box–Jenkins multiplicative ARIMA modeling for prediction of solar radiation: a case study*. International Journal of Energy and Water Resources, 2019. **3**: p. 305-318.
14. Makridakis, S. and M. Hibon, *ARMA models and the Box–Jenkins methodology.* Journal of forecasting, 1997. **16**(3): p. 147-163.
15. Auger-Méthé, M., et al., *A guide to state–space modeling of ecological time series.* Ecological Monographs, 2021. **91**(4): p. e01470.
16. Cressie, N., et al., *Accounting for uncertainty in ecological analysis: the strengths and limitations of hierarchical statistical modeling.* Ecological Applications, 2009. **19**(3): p. 553-570.
17. Shukur, O.B., *Artifical Neural Network and Kalman Filter Approaches Based on ARIMA for Daily Wind Speed Forecasting*. 2015, Universiti Teknologi Malaysia.
18. Galanis, G., et al. *Applications of Kalman filters based on non-linear functions to numerical weather predictions*. in *Annales geophysicae*. 2006. Copernicus Publications Göttingen, Germany.
19. Louka, P., et al., *Improvements in wind speed forecasts for wind power prediction purposes using Kalman filtering.* Journal of Wind Engineering and Industrial Aerodynamics, 2008. **96**(12): p. 2348-2362.
20. Malmberg, A., U. Holst, and J. Holst, *Forecasting near-surface ocean winds with Kalman filter techniques.* Ocean Engineering, 2005. **32**(3-4): p. 273-291.
21. Costa, A., et al., *A review on the young history of the wind power short-term prediction.* Renewable and Sustainable Energy Reviews, 2008. **12**(6): p. 1725-1744.
22. Madsen, H., *Time Series Analysis*. 2007: CRC Press.
23. Malig, B.J., et al., *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, 2016. **124**(6): p. 745-753.
24. Tobias, A., B. Armstrong, and A. Gasparrini. *Analysis of time-stratified case-crossover studies in environmental epidemiology using Stata*. in *United Kingdom Stata Users' Group Meetings 2014*. 2014. Stata Users Group.

# استخدام نموذج فضاء الحالة بالاعتماد على نموذج ARIMA للتنبؤ بدرجة حرارة الهواء

سهى سليم محجوب و أسامة بشير شكر

قسم الإحصاء والمعلوماتية، كلية علوم الحاسوب والرياضيات، جامعة الموصل، موصل ،العراق.

**الخلاصة:** ان الدقة العالية للتنبؤ ببيانات درجة الحرارة مهم جداً للسيطرة على الاضرار البيئية المتمثلة بالتصحر وجفاف مصادر المياه كما هو مهم للسيطرة على استخدامات الطاقة المتجددة والطاقة النظيفة . ان استخدام نـموذج الانحدار الذاتي والمتوسطات المتحركة التكاملي Integrated Auto–regressive and Moving Average (ARIMA) الموسمي المضاعف للتنبؤ في حال وجود عدم التأكدية في عملية النمذجة خصوصاً مع البيانات ذات النمط غير الخطي مثل درجات الحرارة الصغرى مما يجعل من النتائج التنبؤية ان تكون قليلة الدقة عند استخدام النموذج الخطي ARIMA. ان تحسين دقة التنبؤ بدرجات الحرارة الصغرى هو الهدف الرئيسي لهذه الدراسة من خلال استخدام طرق أكثر ملائمة وذلك لنمذجة البيانات مع وجود مشكلة عدم التأكدية. في هذه الدراسة، سيتم استخدام بيانات درجة الحرارة الصغرى لمدينتي الموصل وبغداد. سيتم استخدام فضاء الحالة State Space (SS) بالاعتماد على نموذج ARIMA والذي يسمى نموذج ARIMA–SS الهجين وذلك لمعالجة مشكلة عدم التأكدية التي سببها عدم خطية بيانات درجة الحرارة وبالتالي تكون نتائج التنبؤ قليلة الدقة. كذلك فان البيانات المناخية تعاني غالباً من عدم التجانس خصوصاً في المناطق غير الاستوائية وذلك للاختلاف العالي بين المواسم الحارة والمواسم الباردة لتلك البيانات وسيتم استخدام التراصف الزمني Time Stratified (TS) لمعالجة مشكلة عدم تجانس البيانات. في الاسلوب ARIMA–SS الهجين يستخدم ARIMA فقط لغرض تحديد ادخال نموذج SS . في هذه الدراسة، تم استخدام نموذج SS وهو اسلوب احصائي لتقدير فضاء الحالة والتنبؤ بها. يتلخص اسلوب SS في دمج المشاهدات مع قيم التنبؤ الحالية عبر استخدام اوزان تقلل من التحيزات والاخطاء. تم استخدام النموذج الهجين ARIMA–SS للتعامل مع عدم التأكدية وتحسين التنبؤ بدرجة الحرارة الصغرى من خلال التعامل الجيد معها. سيتم مقارنة أداء نموذج ARIMA ونموذج ARIMA–SS الهجين لتحديد أي منها يكون اداؤه بالتنبؤات الأكـثر دقة. واظـهرت النتـائج ان نموذج ARIMA–SS الهجين أفضل اداءً من نموذج ARIMA وانتج تنبؤات أكثر دقة. لذلك فمن الممكن استنتاج ان استخدام نموذج ARIMA–SS الهجين سينتج عنه دقة تنبؤية أفضل لدرجة الحرارة الصغرى مقارنة بأداء تنبؤ نموذج ARIMA التقليدي.

**الكلمات المفتاحية:** نموذج اريما، نموذج فضاء الحالة،درجات الحرارة الصغرى، التنبؤ.