Nanotechnology Perceptions
ISSN 1660-6795
www.nano-ntp.com

Forecasting Of Maximum And Minimum Temperatures Using The VAR Model

Researcher: Shajwan Mansur Hussien[1], Asst. Prof. Dr. Amer Fadhil Tawfeeq[2]

[1,2]*Faculty of Administration and Economics Kirkuk University.*

This paper aims to predict of the weather phenomenon in Kirkuk Governorate for the period between (January 1990–December 2022) by using the Vector Autoregressive (VAR) modeling. As this research deals with the presentation of a number of information criteria which are(AIC, SIC, HQC), we try to arrive the best Vector Autoregressive model and use it to predict the weather phenomena which are(maximum temperatures, minimum temperatures) in Kirkuk governorate for the next five years. The best model for maximum temperatures was found based on the value of the statistical criteria used.

Key words:
AR models, VAR models, Forecasting of maximum temperature, Forecasting of minimum temperature, Weather of Kirkuk, AIC, SIC, HQC.

1. Introduction:

Choosing the best model from among the maximum and minimum temperatures available to represent any phenomenon is one of the main issues in time series analysis because doing so ensures that the model accurately describes the phenomenon and, on the other hand, that future predictions coming out of it are accurate and realistic.

In many situations, including predicting the weather, it is critical to use VAR to predict future behavior. The best model may be chosen and evaluated using a range of model types that researchers who are interested in time series analysis have developed.

The vector autoregressive (VAR) models for the study and modeling of the time series for the weather in Kirkuk Governorate for the period from (January 1990-December 2022) and predicting future values of the phenomenon studied were the most significant time series models used in this study. Using various statistical criteria that were appropriate in this discipline, the best model was also chosen.

The vector autoregressive (VAR) model is one of the most efficient, flexible, and straightforward approaches for the analysis of multivariate time series. The univariate autoregressive model is logically extended by dynamic multivariate time series. It has been demonstrated that the VAR model is very useful for predicting and describing the dynamic behavior of weather time series. It typically provides predictions that are more accurate than those produced by univariate time series models and complicated simultaneous equations models. Forecasts utilizing VAR models may be quite flexible since they can be made conditional on the probable future courses of certain model variables.

Multivariate simultaneous equations models were frequently used for weather analysis until Sims (1980) offered vector autoregressive (VAR) models as alternatives. At the time, lengthier and more frequent observed weather time series were required, necessitating the need for models that reflected the dynamic structure of the variables. VAR models are ideal for this application. They frequently hold that all variables are endogenous from the start. They do this in order to respond to Sims' critique that several exogenous assumptions in simultaneous equations models are ad hoc and frequently not supported by well-developed theories. VAR models may be subject to limits based on statistical methods, such as the exogenous use of some variables.

2. <u>Theoretical aspect:</u>

2-1 Vector Autoregressive Model (VAR):

VAR models were first introduced by Sims (1980). One of the most effective, adaptable, and simple methods for the study of multivariate time series is the vector autoregressive (VAR) model. Dynamic multivariate time series are a logical extension of the univariate autoregressive model.

Let $\{Y_t = (Y_{1t}, \dots Y_{Kt})'; t \in Z\}$ be a K-variable random process. We say that the process $\{Y_t; t \in Z\}$ follows a p-order vector autoregressive model, or VAR(p), if:

$$Y_t = \nu + \phi_1 Y_{t-1} + \cdots + \phi_p Y_{t-p} + u_t \qquad (2\text{-}1)$$

Where p is a positive integer, ϕ_i are fixed $(K \times K)$ coefficients matrices, $\nu = (\nu_1, \dots, \nu_K)'$ is a fixed $(K \times 1)$ vector of intercept terms, $u_t = (u_{1t}, \dots, u_{Kt})'$ is a K-dimensional white noise with covariance matrix $\sum u$. The covariance matrix $\sum u$ is assumed to be nonsingular.

The VAR model has shown to be particularly effective for forecasting and characterizing the dynamic behavior of economic and financial time series. It frequently offers forecasts that are better than those from complex simultaneous equations models and univariate time series models, because they can be made conditional on the likely future courses of certain model variables, also forecasts from VAR models can be highly flexible. VAR models follow three fundamental principles:

- The system is devoid of both internal and external presuppositions.
- The model is not founded on a rigorous core economic theory.
- There is no zero-type restriction.

The fundamental goal of VAR modeling is to disclose connection between variables in terms of lags in addition to determining the one-way relationship between variables (Kearney and Monadjemi 1990):

- The procedure is straightforward; it is not essential to distinguish between internal and external variables. In a VAR, every variable is endogenous.
- Prediction is straightforward; each equation may be solved using the standard LS approach.

- In many instances, the estimates produced by this method is superior to those produced by simultaneous equation models with higher levels of complexity.

2-2 The First-Order Vector Autoregressive VAR(1):

The first order VAR models for this bivariate system is appeared the first time by Sims (1980), as in equation (2-2):

$$\begin{pmatrix} y_{1t} \\ y_{2t} \end{pmatrix} = \begin{pmatrix} \nu_1 \\ \nu_2 \end{pmatrix} + \begin{pmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{pmatrix} \begin{pmatrix} y_{1t-1} \\ y_{2t-1} \end{pmatrix} + \begin{pmatrix} u_{1t} \\ u_{2t} \end{pmatrix} \quad (2\text{-}2)$$

Or

$$\begin{aligned} y_{1t} &= \nu_1 + \phi_{11} y_{1t-1} + \phi_{12} y_{2t-1} + u_{1t} \\ y_{2t} &= \nu_2 + \phi_{21} y_{1t-1} + \phi_{22} y_{2t-1} + u_{2t} \end{aligned} \quad (2\text{-}3)$$

Where $\text{cov}(u_{1t}, u_{2s}) = \sigma_{12}$ for t = s, 0 otherwise. That each equation has the same regressors — lagged values of $y_{1t}$and y_{2t}. Hence, the VAR(p) model is just a seemingly unrelated regression (SUR) model with lagged variables and deterministic terms as common regressors. The Condition of Stationary is:

1. $(y_{1t}\,; y_{2t})$ are both stationary when the eigenvalues of ϕ are less than one in absolute value.
2. $(y_{1t}\,; y_{2t})$ are cointegrated when one eigenvalue is unity and the other eigenvalue is less than one in absolute value. They are both integrated of order one.
3. $(y_{1t}\,; y_{2t})$ are both integrated of order two if both eigenvalues of ϕ are unity.

2-3 The General Vector Autoregressive VAR(p):

The Vector Autoregressive model of order p is given by (Lütkepohl (2005):

$$y_t = \nu + \sum_{i=1}^{p} \phi_i \, y_{t-i} + u_t \quad (2\text{-}4)$$

Where $y_t = (y_{1t}, \dots, y_{Kt})'$ is a vector of dimensions $(K \times 1)$, being K the number of variables, each ϕ_i is a matrix of autoregressive coefficients, of dimensions $(K \times K)$; $\nu = (\nu_1, \dots, \nu_K)'$ is a vector of dimension $(K \times 1)$ that permits a non-zero mean $E(y_t)$; and $u_t = (u_{1t}, \dots, u_{Kt})'$ is a K—dimensional white noise, also referred to as the creativity process or error, that must imply $E(u_t) = 0$, $E(u_t\, u'_t) = \sum u$ and $E(u_t\, u'_s) = 0$ for $s \neq t$.

The VAR(p) model parameters are estimated in this work using the Least Square method. The model is written as:

$$Y = \beta Z + U \quad (2\text{-}5)$$

In matrix notation for this:

$$Y = (y_1, \dots, y_T),$$

$$\beta = (\nu, A_1, \dots, A_p),$$

$$Z_t = [1, y_t, \dots, y_{t-p+1}]',$$

$$Z = (Z_0, Z_1, \dots, Z_{T-1})$$

And

$$U = (u_1, \dots, u_T),$$

With T the total number of accessible observations for estimation. Then, autoregressive parameters are estimated as:

$$\hat{\beta} = (ZZ')^{-1}Z^{-1}Y \qquad (2\text{-}6)$$

The covariance matrix, meantime $\sum u$ of the withe noise u_t is estimated through the errors:

$$\hat{U} = Y - \hat{\beta}Z \qquad (2\text{-}7)$$

As

$$\widehat{\sum u} = \frac{\hat{U}\hat{U}'}{T-KP+1} \qquad (2\text{-}8)$$

2-4 Ordinary Least Square(OLS) for VAR(1) Model:

To derive the estimation formulas we explain equation (2-3) as follow:

$$y_{1t} = v_1 + \phi_{11}y_{1t-1} + \phi_{12}y_{2t-1} + u_{1t}$$
$$y_{2t} = v_2 + \phi_{21}y_{1t-1} + \phi_{22}y_{2t-1} + u_{2t}$$

Where $cov(u_{1t}, u_{2s}) = \sigma_{12}$ for $t = s$, 0 otherwise.

Two regressions with separate dependent variables and the same explanatory factors make up the model. By computing the ordinary least squares (OLS) estimator individually for each equation, we might estimate this model. It is assumed that a time series.
The premise is that a time series:

$$y_1 = [y_{11}, y_{21}]', \dots, y_T = [y_{1T}, y_{2T}]'$$

availability of the y variables. Moreover, a pre sample value:

$$y_0 = [y_{10}, y_{20}]'$$

is assumed to be a variable. Think about the first equation:

$$y_{1t} = v_1 + \phi_{11}y_{10} + \phi_{12}y_{2t-1} + u_{1t}; \quad t = 1, \dots, T$$
$$y_{11} = v_1 + \phi_{11}y_{10} + \phi_{12}y_{20} + u_{11}$$
$$y_{12} = v_1 + \phi_{11}y_{11} + \phi_{12}y_{21} + u_{12}$$
$$\vdots$$
$$y_{1T} = v_1 + \phi_{11}y_{1T-1} + \phi_{12}y_{2T-1} + u_{1T}$$

We define:

$$y_1 = [y_{11}, \dots, y_{1T}]',$$

$$X_1 = \begin{bmatrix} 1 & y_{1,0} & y_{2,0} \\ \vdots & \vdots & \vdots \\ 1 & y_{1,T-1} & y_{2,T-1} \end{bmatrix},$$

$$\pi_1 = [v_1, \phi_{11}, \phi_{12}]',$$

$$u_1 = [u_{11}, \dots, u_{1T}]',$$

Thus:

$$y_1 = X\pi_1 + u_1, \qquad (2\text{-}9)$$

The OLS estimator is $\hat{\pi}_1$:

$$\hat{\pi}_1 = (X'X)^{-1}X'y_1 \qquad (2\text{-}10)$$

In the second equation:

$$y_{2t} = \nu_2 + \phi_{21}y_{10} + \phi_{22}y_{2t-1} + u_{2t}; \quad t = 1, \dots, T$$

$$y_{21} = \nu_2 + \phi_{21}y_{10} + \phi_{22}y_{20} + u_{21}$$

$$y_{22} = \nu_2 + \phi_{21}y_{11} + \phi_{22}y_{21} + u_{22}$$

$$\vdots$$

$$y_{2T} = \nu_2 + \phi_{21}y_{1T-1} + \phi_{22}y_{2T-1} + u_{2T}$$

We define:

$$y_2 = [y_{21}, \dots, y_{2T}]',$$

$$\pi_2 = [\nu_2, \phi_{21}, \phi_{22}]',$$

$$u_2 = [u_{21}, \dots, u_{2T}]',$$

Thus:

$$y_2 = X\pi_2 + u_2, \qquad (2\text{-}11)$$

The OLS estimator is $\hat{\pi}_2$:

$$\hat{\pi}_2 = (X'X)^{-1}X'y_2 \qquad (2\text{-}12)$$

Therefore the OLS estimators:

$$\hat{\pi}_1 = (X'X)^{-1}X'y_1 \Rightarrow \pi_1 = [\nu_1, \phi_{11}, \phi_{12}]'$$

$$\hat{\pi}_2 = (X'X)^{-1}X'y_2 \Rightarrow \pi_2 = [\nu_2, \phi_{21}, \phi_{22}]'$$

2-5 Testing Stationary Time Series:

There are several ways to test if a time series is stationary. One popular technique nowadays is the unit roots test, which determines whether the series is stationary when its unit roots are located inside the unit circuit. The Augmented Dickey-Fuller test (ADF) is a crucial tool for identifying time series stationary.

Autoregressive models of the rank (p), which are based on the following formula, are used in the Augmented Dickey-Fuller test.

The constant term is not present in the initial equation

$$\Delta Y_t = \phi_1 Y_{t-1} + \sum_{j=1}^{p} \beta_j \, \Delta Y_{t-j} + \varepsilon_t \qquad (2\text{-}13)$$

The constant term appears in the second equation

$$\Delta Y_t = \phi_0 + \phi_1 Y_{t-1} + \sum_{j=1}^{p} \beta_j \, \Delta Y_{t-j} + \varepsilon_t \qquad (2\text{-}14)$$

The constant term and the temporal trend are included in the third equation

$$\Delta Y_t = \phi_0 + \phi_1 Y_{t-1} + \sum_{j=1}^{p} \beta_j \, \Delta Y_{t-j} + \delta_t + \varepsilon_t \qquad (2\text{-}15)$$

Where
Δ: The first difference coefficient , Which $\Delta Y_t = Y_t - Y_{t-1}$.
ϕ_0: constant expression.
δ_t: Trend in time.
β, ϕ: Parameters testing.
ε_t: White noise.
And the following hypothesis is tested:
$H_0: \phi_1 = 0$ (There is a unit root) Non-stationary.
$H_1: \phi_1 = 0$ (There is not a unit root) stationary.

The test statistics are as follows, please take not

$$t = \frac{\phi_1}{SE(\phi_1)}$$

The test statistic and the tabular values (Dickey-Fuller tables) are contrasted. The alternative hypothesis is accepted and the null hypothesis is rejected if the computed (t) value is higher than the tabular value, indicating that the series is stationary. The alternative hypothesis is accepted, indicating that the series is stationary, if the p-value is smaller than the designated level of significance.

2-6 Impulse Response Function (IRF):
We can demonstrate that the MA(∞) representation for the VAR(1) is:

$$\begin{pmatrix} y_{1t} \\ y_{2t} \end{pmatrix} = \begin{pmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{pmatrix} \begin{pmatrix} y_{1t-1} \\ y_{2t-1} \end{pmatrix} + \begin{pmatrix} u_{1t} \\ u_{2t} \end{pmatrix} \quad (2\text{-}16)$$

Or

$$z_t = \phi z_{t-1} + w_t \text{ (Reduced Forms)} \quad (2\text{-}17)$$

Where $z_t = \begin{pmatrix} y_{1t} \\ y_{2t} \end{pmatrix}$, $\phi = \begin{pmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{pmatrix}$, and $w_t = \begin{pmatrix} u_{1t} \\ u_{2t} \end{pmatrix}$

Then the MA(∞) representation for the VAR(1) is:

$$z_t = w_t + \phi^1 w_{t-1} + \phi^2 w_{t-2} + \ldots\ldots + \phi^j w_{t-j} + \ldots \quad (2\text{-}18)$$

The MA representation's coefficient quantifies the impulse response

$$\phi^j = \frac{dz_t}{dw_t} \quad (2\text{-}19)$$

Where ϕ^j is a 2×2 matrix for a bivariate system. More generally, the $(m,n) - th$ component of ϕ^j by $\phi^j = (m,n)$. Then $\phi^j = (m,n)$ measures the response of $m - th$ variable to the $n - th$ error after j periods. There are four impulse responses plots for a bivariate system. In general u_{1t} and u_{2t} are contemporaneously correlated (not-orthogonal), i.e., $\sigma_{12} \neq 0$.

However, we can always find a lower triangular matrix A so that:

$$\Omega = AA' \quad \text{(Cholesky Decomposition)} \quad (2\text{-}20)$$

Then define a new error vector $\tilde{w}_t$ as (linear transformation of old error vector w_t)

$$\tilde{w}_t = A^{-1} w_t \quad (2\text{-}21)$$

The new error is by design orthogonal since the variance-covariance matrix is diagonal:

$$var(\widetilde{w}_t) = A^{-1}(w_t)A^{-1'} = A^{-1}\Omega A^{-1'} = A^{-1}AA'A^{-1'} = I$$

2-7 Cholesky Decomposition:

Let $A = \begin{pmatrix} a & 0 \\ b & c \end{pmatrix}$. The Cholesky decomposition tries to solve:

$$\begin{pmatrix} a & 0 \\ b & c \end{pmatrix}\begin{pmatrix} a & b \\ 0 & c \end{pmatrix} = \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{pmatrix}$$

The solutions for a, b, c always exist and they are:

$$a = \sqrt{\sigma_1^2} \qquad (2\text{-}22)$$

$$b = \frac{\sigma_{12}}{\sqrt{\sigma_1^2}} \qquad (2\text{-}23)$$

$$c = \sqrt{\sigma_2^2 - \frac{\sigma_{12}^2}{\sigma_1^2}} \qquad (2\text{-}24)$$

Where c is always a real number since Ω is a variance-covariance matrix, and so is positive definite (i.e., $\sigma_2^2 - \frac{\sigma_{12}^2}{\sigma_1^2}$ is always positive because the determinant of Ω, or the second leading principal minor, is positive).

2-8 Orthogonal Errors and Impulse Response:

The MA(∞) representation should be rewritten as:

$$z_t = w_t + \phi w_{t-1} + \cdots + \phi^j w_{t-j} + \cdots \qquad (2\text{-}25)$$

$$= AA^{-1}w_t + \phi AA^{-1}w_{t-1} + \cdots + \phi^j AA^{-1}w_{t-J} + \cdots \qquad (2\text{-}26)$$

$$= A\widetilde{w}_t + \phi A\widetilde{w}_{t-1} + \cdots + \phi^j A\widetilde{w}_{t-j} + \cdots \qquad (2\text{-}27)$$

The impulse response to the orthogonal error $\widetilde{w}_t$ after j periods $j = (1, 2, 3, \dots)$ is implied by this:

$$j - th\ orthogonal\ impulse\ response = \phi^j A \qquad (2\text{-}28)$$

where A satisfies (2-20).

2-9 Fitting model:

The selection of a particular model from among a variety of models is an important element of the analysis of the data since it helps us determine which model is best. The following statistical criteria are used:

- AKaike Information Criterion (AIC): The concept of penalizing the addition of regressors to the model has been expanded upon in the AIC criteria, which is defined as:

$$AIC = e^{2k/n}\frac{\sum \hat{u}_i^2}{n} = e^{2k/n}\frac{RSS}{n} \qquad (2\text{-}29)$$

Where
k: The intercept is included in the total number of regressors.
n: The total number of observations.
Eq. (2-29) is conveniently represented mathematically as follows:

$$ln\,AIC = \left(\frac{2k}{n}\right) + ln\left(\frac{RSS}{n}\right) \quad (2\text{-}30)$$

Where
$ln\,AIC$: is the natural log of AIC.
$2k/n$: is penalty factor.
The best model is the one with the lowest AIC.

- Schwarz Information Criterion (SIC): The AIC and SIC are identical in spirit, according to the definition of the SIC criteria.

$$SIC = n^{k/n}\frac{\sum \hat{u}^2}{n} = n^{k/n}\frac{RSS}{n} \quad (2\text{-}31)$$

or using a log:

$$ln\,SIC = \frac{k}{n}\,ln\,n + ln\left(\frac{RSS}{n}\right) \quad (2\text{-}32)$$

When the penalty factor $[(k/n)\ ln\ n]$ is present. It is clear from comparing Eqs. (2-32) and (2-30) that SIC imposes a higher penalty than AIC. Similar to AIC, the better the model, the lower the value of SIC. SIC may be used to compare a model's in-sample or out-of-sample predicting ability, just as AIC.

- Hannan - Quinn Criterion (HQC): The Hannan-Quinn Criterion, denoted by the sign (HQC), was developed by Hannan and Quinn to evaluate the rank of the model under study. Its formula is as follows:

$$H - Q = ln\,\hat{\sigma}^2 + 2MC\,ln\,(ln\,n)/n;\ c > 2 \quad (2\text{-}33)$$

Where
M: The quantity of model parameters.
C: Is constant.

When the rank is constant, the repeating logarithm and the suitable model that yields the lowest value of the criteria HQC cause the second term of the aforementioned formula to decline as rapidly as feasible.

3. Practical aspect:

we take the data of weather in Kirkuk Governorate for the period between (January 1990–December 2022) which are (maximum temperatures(MAT), minimum temperatures(MIT)) are used to illustrate the vector autoregressive of time series data in the Figure(1) for the original data below. Through statistical analysis and visual representation, time series data can self-regress to reveal their fundamental characteristics. The data were assessed, and the model's parameters were estimated, using the Eviews program.

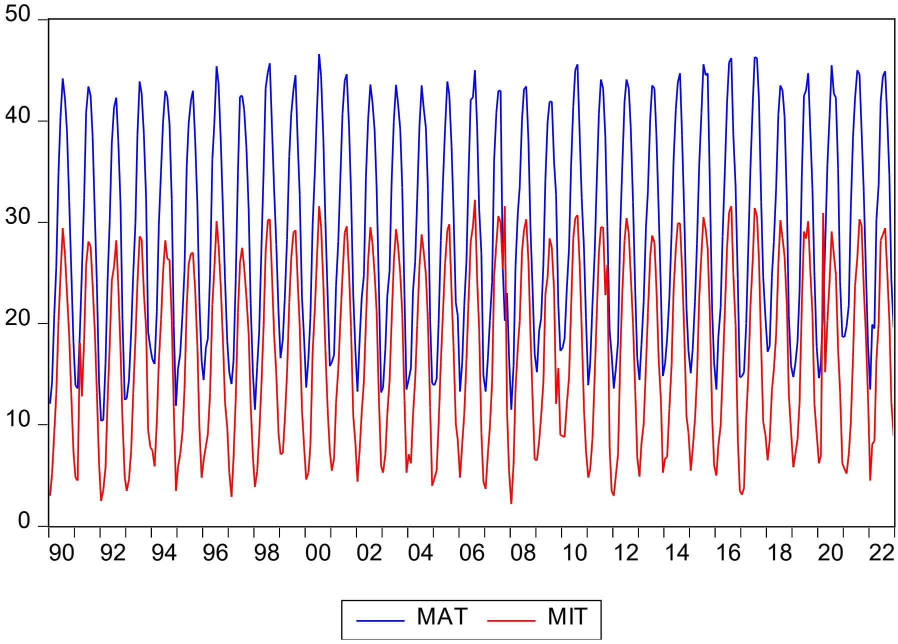


Figure (1) The original time series of weather phenomena data

3-1 Time Series Stationary Tests:

We use the Dickey - Fuller test to check if a chain is stable. Shows from the tables below the result of testing the null hypothesis that the series is not constant and then computed for the purpose of testing the original constant data.

3-1-1 Max Temperature and Min Temperature:

To explain and illustrate the maximum temperature and minimum temperature data in order to achieve the research's goal, we must build a timeline for the time series in order to determine if it is stationary or non-stationary. For this reason, a graph showing the highest temperatures in Kirkuk Governorate from January 1990 to December 2022 was created. As seen in figure (1) above, the time series suffers from a seasonal component, indicating that it is non-stationary. As a result, we conduct stationary tests to confirm that the time series is non-stationary. And a graph of the lowest temperatures in Kirkuk Governorate . As seen in figure (1) above, the series is non-stationary because of the time series' increasing seasonal component and general trend component. Therefore, we conduct stationary tests on the time series to ensure that they are not stationary.

Table (1) : Dickey-Fuller test result (ADF) for MAT

ADF	t-Statistics	P-value	Log likelihood	AIC	SIC	HQC
Level	-3.560411	.0070	-782.7633	.192983	4.358560	4.258678
1st difference	-16.78767	0.0000	-801.3157	4.236019	4.359477	4.284988

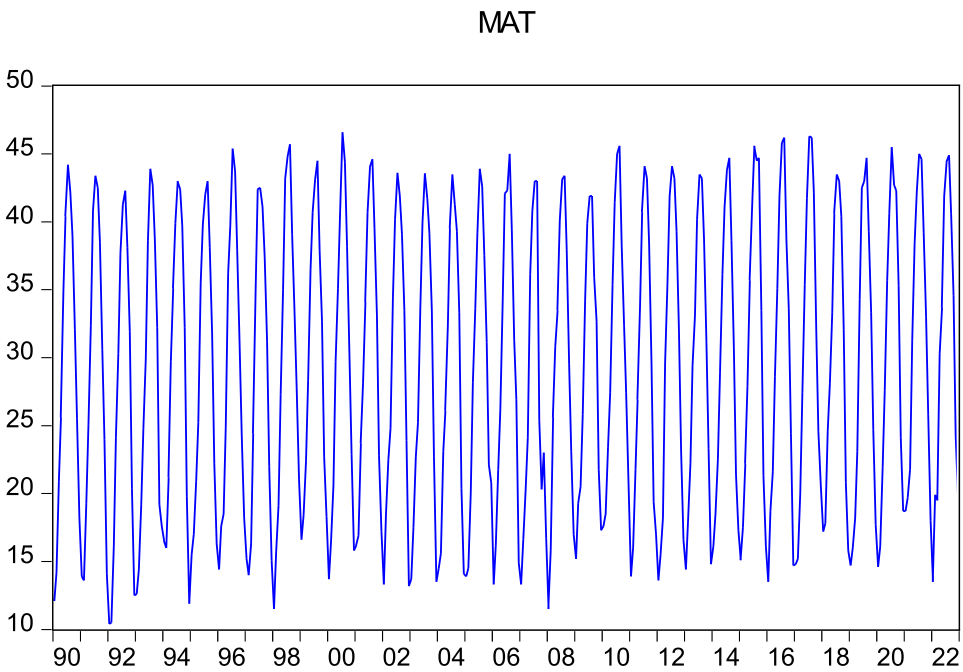


Figure (2) the series after taking the differences for MAT

Table (2) : Dickey-Fuller test result (ADF) for MIT

ADF	t-Statistics	P-value	Log likelihood	AIC	SIC	HQC
Level	-4.140423	0.0009	-863.9297	4.567342	4.701088	4.620392
1st difference	-17.11953	0.0000	-872.6027	4.607306	4.730764	4.656275

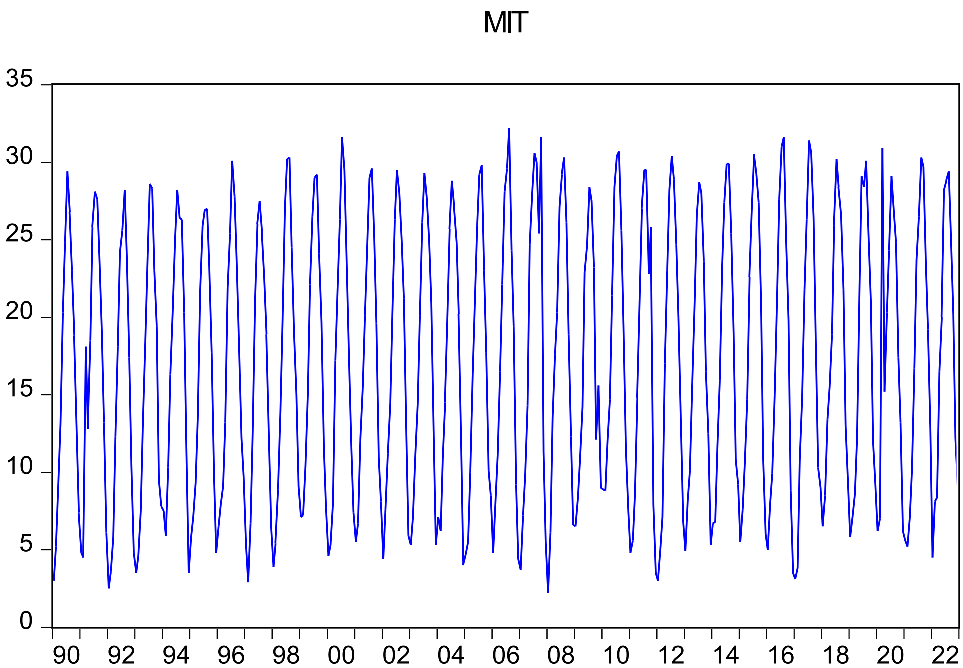


Figure (3) the series after taking the differences for MIT

We note from the above table(1) and table(2) that the p-value for each level and the first difference of the test result is smaller than the level of importance, and therefore we reject the null hypothesis that the series is constant, and then accept the alternative hypothesis that the time series is stationary.

3-2 Model Estimation (VAR):

After the series has stationary, we have to choose the degree of the model and in such a step to initially know the diagnosed model and then a number of models close to the diagnosed model are tested to choose the best ones based on some statistical criteria and it has been shown. The diagnostic model is VAR(MAT), and the table (3) shows the comparison between these models according to the criteria for choosing the model (Note the bold number indicates the best value for the criterion and thus the best model for the time series).

Table (3) : Comparison of VAR models estimated to represent a series of weather phenomena

Model	Log likelihood	AIC	SIC
VAR(MAT)	-984.5437	**5.023064**	**5.073526**
VAR(MIT)	-995.6602	5.079493	5.129955

By the criteria (AIC), and (SIC) Which have the smallest value. We find that the model VAR(MAT) is the best among the proposed models for the above criteria.

3-3 Determine the degree of delay using the VAR model:
We select the lowest value of the four criteria because it is the criterion for the model's applicability to the given data.

Table (4) : VAR Lag Order Selection Criteria

Lag	Log L	LR	FPE	AIC	SIC	HQ
0	-2364.794	NA	583.4734	12.04475	12.06498	12.05277
1	-2092.636	540.1601	149.0572	10.68008	10.74075	10.70413
2	-1837.095	504.5812	41.43895	9.399973	9.501087	9.440043
3	-1763.254	145.0512	**29.04371***	**9.044549***	**9.186109***	**9.100647***

Given that we selected the lowest value of the four criteria, we will utilize lag3 in the estimation of the VAR model. In the following table (5).

Table (5) : Comparison of VAR models estimated to represent a series of weather phenomena

Model	AIC	SIC
VAR(MAT)	**4.701477**	**4.772258**
VAR(MIT)	4.810265	4.881045

By the criteria (AIC), and (SIC) Which have the smallest value. We find that the model VAR(MAT) is the best among the proposed models for the above criteria.

3-3-1 Estimate MAT model:
Thus, it was concluded that the appropriate model for weather series data is VAR (MAT), as table (6) shows the model's estimated coefficients. The above model is a two-variable model that we can describe in the other model with a system of equations, using the method of least squares.
The estimated model used for weather forecasts (series) is:
MAT = C(1)*MAT(-1) + C(2)*MAT(-2) + C(3)*MAT(-3) + C(4)*MIT(-1) + C(5)*MIT(-2) + C(6)*MIT(-3) + C(7)

Table (6) : The estimated values of the model parameters for the weather phenomena for MAT

Models	Coefficient	Std. error	t-Statistic	P-value

C(1)	0.948935	0.061367	15.46317	0.0000
C(2)	-0.046943	0.078299	-0.599535	0.5490
C(3)	-0.271967	0.059941	-4.537246	0.0000
C(4)	0.227934	0.059627	3.822668	0.0001
C(5)	0.010446	0.063291	0.165054	0.8689
C(6)	-0.304723	0.059925	-5.085086	0.0000
C(7)	11.98762	0.736188	16.28337	0.0000

By substituting the values of the coefficients into the above formula, we get the following:

MAT = 0.948935124298*MAT(-1) - 0.0469431304173*MAT(-2) - 0.271967323898*MAT(-3) + 0.227934032183*MIT(-1) + 0.0104464424505*MIT(-2) - 0.30472346821*MIT(-3) + 11.9876244205

3-3-2 Estimate MIT model:

The estimated model used for weather forecasts (series) is:

MIT = C(8)*MAT(-1) + C(9)*MAT(-2) + C(10)*MAT(-3) + C(11)*MIT(-1) + C(12)*MIT(-2) + C(13)*MIT(-3) + C(14)

Table (7) : The estimated values of the model parameters for the weather phenomena for MiT

Models	Coefficient	Std. error	t-Statistic	P-value
C(8)	0.451053	0.064798	6.960915	0.0000
C(9)	0.076554	0.082676	0.925945	0.3548
C(10)	-0.315789	0.063292	-4.989417	0.0000
C(11)	0.465435	0.062960	7.392540	0.0000
C(12)	0.068446	0.066829	1.024195	0.3061
C(13)	-0.234950	0.063275	-3.713169	0.0002
C(14)	5.781733	0.777341	7.437832	0.0000

By substituting the values of the coefficients into the above formula, we get the following:

MIT = 0.451052649571*MAT(-1) + 0.0765536222174*MAT(-2) - 0.315789126751*MAT(-3) + 0.465435139147*MIT(-1) + 0.0684459357695*MIT(-2) - 0.23494988664*MIT(-3) + 5.78173304166

3-4 Model Diagnoses:

At this stage, errors (residuals) are examined and diagnosed to find out the extent of the preference of the proposed model that has been identified and its parameters estimated.

The residual distribution has a normal distribution that matches with assumptions that $\varepsilon_t \sim IID(0, \sigma_\varepsilon^2)$. This can be known from the residual drawing by using the histogram drawing of model errors closer to the normal distribution which indicates its randomness and this is confirmation of the quality of the model which is shown by the figure (4).

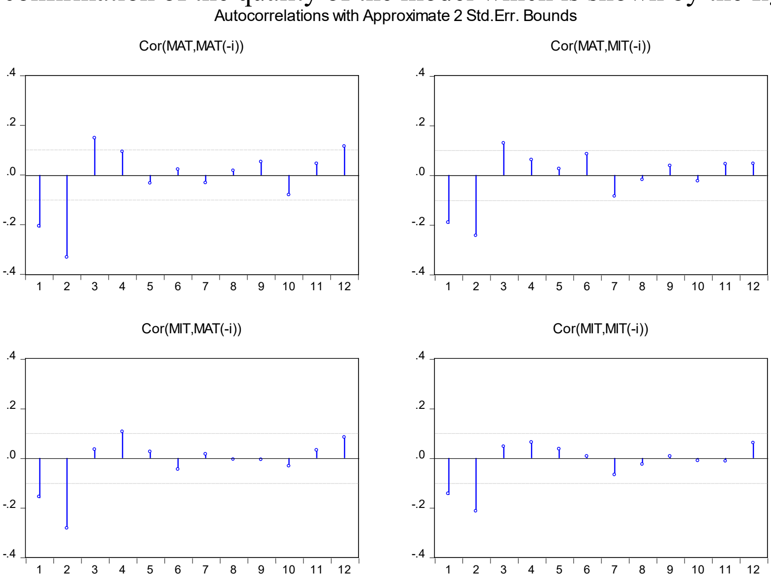


Figure (4) Test for the residuals' correlograms for a number of weather phenomena variables

With the exception of variables that are outside the confidence domains, we observe that the all of the residuals we have are not inside them. The confidence contain all of the residuals.

Table (8) : VAR Residual Normality Tests

Component	Jarque-Bera	df	P-value
1	46.98730	2	0.0000
2	20775.78	2	0.0000

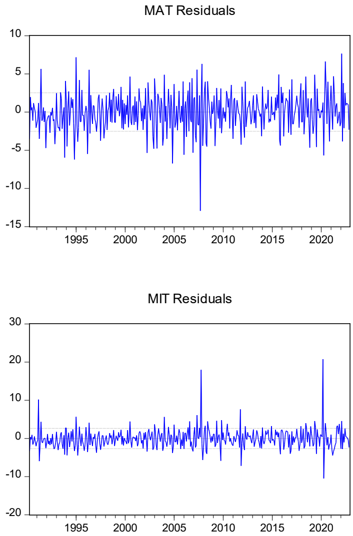


Figure (5) VAR Residuals

3-5 Analysis of Impulse Response:

After reviewing the steps for determining the appropriate model for the atmospheric series data, estimating its parameters, and examining the model, we use impulse response analysis. Always (VAR) is associated with the impulse response function. Here, we need research on the influence of other factors on (MAT), as shown in table (9).

Table (9) : Effect of cholesky (d.f adjusted) MAT innovation

Period	MIT
1	1.623531
2	1.890842
3	2.427997
4	1.623806
5	0.762018
6	-0.477585
7	-1.408035
8	-1.993230
9	-1.962820
10	-1.431194

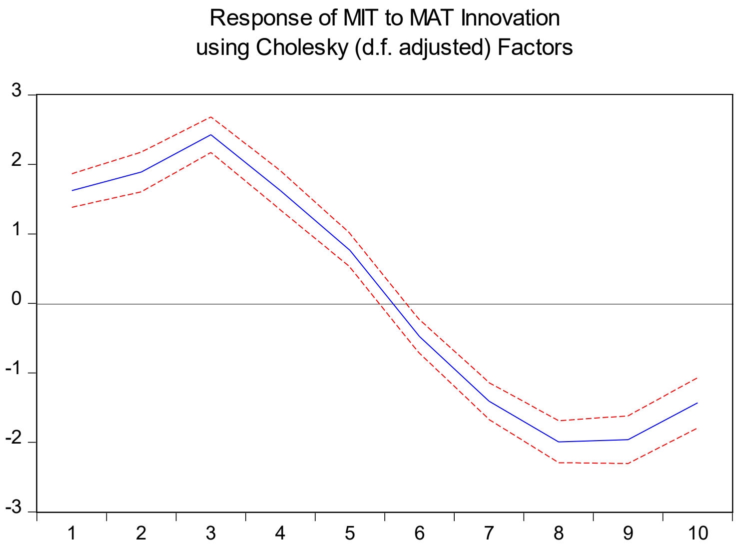


Figure (6) The impact of other factors on (MAT)

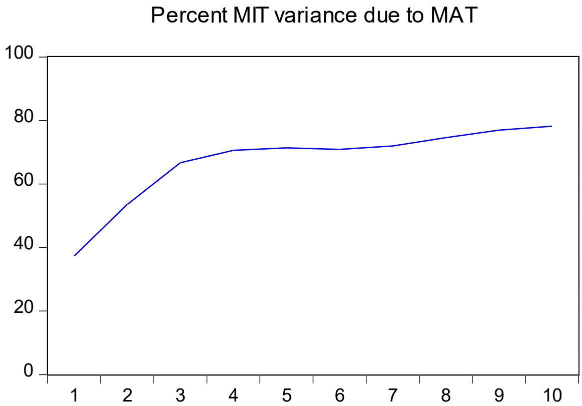
3-6 Analysis of Variance:

After reviewing the steps to determine the appropriate model for the data of the weather conditions chain, estimating its coefficients, examining the model, and analysis of the impulse response, we use the analysis of variance. Here, we need a research on the influence of the factors for analysis (MIT) on (MAT, MIT), as shown in table (10).

Table (10) : Analysis of Variance Using Cholesky (d.f adjusted) Factors

Analysis of Variance MIT			
Period	S.E	MAT	MIT
1	2.516766	37.32414	62.67586
2	3.764614	53.56187	46.43813
3	4.832075	66.67456	33.32544
4	5.257044	70.58436	29.41564
5	5.334932	71.37438	28.62562
6	5.407383	70.90754	29.09246
7	5.754756	71.98205	28.01795
8	6.316612	74.59446	25.40554
9	6.788203	76.95705	23.04295
10	7.007844	78.20154	21.79846

Variance Decomposition using Cholesky (d.f. adjusted) Factors



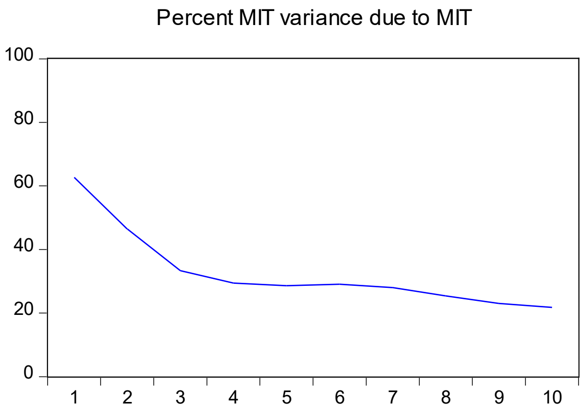


Figure (7) of the factors for analysis (MIT) on (MAT, MIT)

3-7 Forecasting:

After going through the steps of identifying the appropriate model for the data of the weather phenomena series estimating its parameters and examining the model we use the model to predict future values of the weather conditions rates for the coming period from (2023 - 2027), as in table (11) showing the results of the monthly rates of weather conditions where the table includes new forecasts for five years and they were compared with the original values and build 95% confidence limits for these predictions. Figure (8) illustrates the time series drawing of the real data, the limits of confidence and new predictions.

Table (11) : Monthly averages of the predicted weather phenomena series with 95% confidence limits for 5 years

Period	MAT-Forecast	IIT- Forecast t
2023M01	16.0	5.8
2023M02	17.4	7.3

2023M03	21.4	10.4
2023M04	27.8	15.7
2023M05	34.1	20.8
2023M06	38.9	24.8
2023M07	40.9	26.5
2023M08	39.7	25.6
2023M09	35.7	22.4
2023M10	30.2	17.9
2023M11	24.7	13.5
2023M12	20.7	10.2
2024M01	19.3	8.9
2024M02	20.6	9.9
2024M03	24.3	12.9
2024M04	29.2	16.8
2024M05	33.8	20.6
2024M06	37.1	23.3
2024M07	38.1	24.2
2024M08	36.6	23.1
2024M09	33.3	20.5
2024M10	29.0	16.9
2024M11	25.1	13.7
2024M12	22.4	11.6
2025M01	21.8	10.9
2025M02	23.2	12.1
2025M03	26.3	14.5
2025M04	29.9	17.5
2025M05	33.4	20.3
2025M06	35.5	22.1
2025M07	35.8	22.4
2025M08	34.4	21.3
2025M09	31.7	19.1
2025M10	28.4	16.5
2025M11	25.6	14.1
2025M12	23.8	12.7
2026M01	23.8	12.6
2026M02	25.1	13.6
2026M03	27.6	15.6
2026M04	30.4	17.9
2026M05	32.8	19.9
2026M06	34.1	20.9
2026M07	34.1	21.0
2026M08	32.8	19.9

2026M09	30.6	18.2
2026M10	28.2	16.2
2026M11	26.2	14.6
2026M12	25.1	13.7
2027M01	25.3	13.8
2027M02	26.5	14.8
2027M03	28.5	16.3
2027M04	30.6	18.1
2027M05	32.2	19.4
2027M06	33.0	20.1
2027M07	32.8	19.9
2027M08	31.6	19.0
2027M09	29.9	17.6
2027M10	28.1	16.2
2027M11	26.7	15.0
2027M12	26.1	14.5

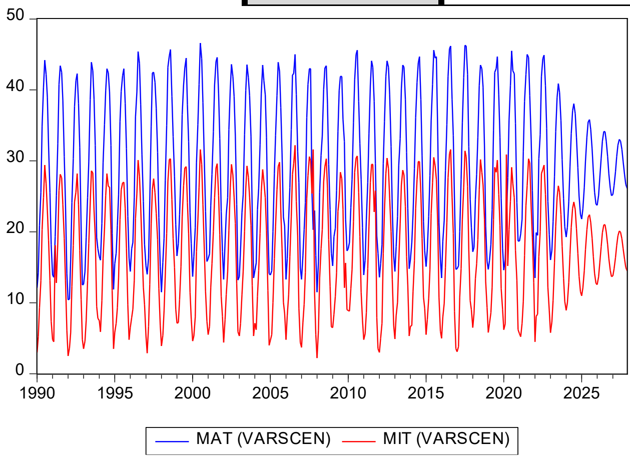


Figure (8) the original values of weather phenomena time series with predicted for the next five years

Therefore in Figure(8) appear that the forecasting values will decreasing for MAT and MIT per month during the year (2023-2027) and by using the model VAR, as it's shown in table(11) and Figure(8) fitting the model by original data and forecasting for years (2023-2027).

4. <u>Conclusions:</u>

1. For the application results for the maximum temperature data indicated that the information criteria, AIC=(4.701477) and SIC=(4.772258), and for the minimum temperature data indicated that the information criteria, AIC= (4.810265), and SIC= (4.881045), and we choose VAR (MAT) because it is the lowest value between AIC and SIC.
2. In this study, utilizing the model VAR, maximum temperature is found to be the best model for predicting compared to other models.

3. In general, maximum temperature and minimum temperature forecasts show a decline for the upcoming months in 2023 and 2027.

5. References:

1. Al-Birmani, Fatima Abdul-Hamid Jawad, and Arsheed, Ahmad Judeh (2019). "The Use of Time Series to Predict Indices Number of Rents of Residential Homes in Iraq for the Years 2018-2021", Al-Rafidain University College of Science Journal, No. (45), (pp. 82-84)
2. Al-Zaidi, Najlaa Khaled Ahmed Abbawi(2010)" A comparative study of some information criteria in selecting time series models from the lower ranks" , Master Thesis, College of Computer Science and Mathematics, University of Mosul (pp. 40)
3. Cavanaugh, J. E., & Neath, A. A. (2019). The Akaike information criterion: Background, derivation, properties, application, interpretation, and refinements. Wiley Interdisciplinary Reviews: Computational Statistics, 11(3), e1460.
4. Cho, H., Maeng, H., Eckley, I. A., & Fearnhead, P. (2023). High-dimensional time series segmentation via factor-adjusted vector autoregressive modeling. Journal of the American Statistical Association, 1-13.
5. Gredenhoff, M., & Karlsson, S. (1999). Lag-length selection in VAR-models using equal and unequal lag-length procedures. Computational Statistics, 14, 171-187.
6. Kearney, C., & Monadjemi, M. (1990). Fiscal policy and current account performance: International evidence on the twin deficits. Journal of Macroeconomics, 12(2), 197-219.
7. Kinal, T., & Ratner, J. B. (1982). Regional forecasting models with vector autoregression: The case of New York State. Department of Economics, State University of New York at Albany.
8. Krishnan, J., Money, R., Beferull-Lozano, B., & Isufi, E. (2023, June). Simplicial vector autoregressive model for streaming edge flows. In ICASSP 2023-2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (pp. 1-5). IEEE.
9. Lütkepohl, H. (2005). New introduction to multiple time series analysis. Springer Science & Business Media.
10. Portet, S. (2020). A primer on model selection using the Akaike Information Criterion. Infectious Disease Modelling, 5, 111-128.
11. Shibata, R. (1976). Selection of the order of an autoregressive model by Akaike's information criterion. Biometrika, 63(1), 117-126.
12. Sims, C. A. (1980). Macroeconomics and reality. Econometrica: journal of the Econometric Society, 1-48.
13. Watson, P. K., & Teelucksingh, S. S. (2002). A practical introduction to econometric methods: Classical and modern. University of West Indies Press.