hypothetical equation of the rank (q), then we need (q) from the inputs only to find the current value of the output, that this past (q) outputs summarize the past for the system whenever we studied the current output for the system, so there is no need for detailed details on the past of the system that affects its present and future. That part of the system's past that affects its present and future influence on the system's behavior is called the state of the system, and the variables that represent the state of the system are called state variables. Moreover, we can develop the state space by relying on the relationship between inputs and outputs or by initial knowledge of the natural law that governs the system [7].

This includes distinguishing the model from the state variable in all input and output variables and the currently unlimited natural statistical knowledge, and in the case of space models, the term of Estimation is usually used to refer to:

- Estimate the unknown (non-known) parameters of the state-space model.
- Estimate states.

To form the state-space model, it is necessary to prepare at least three variables, namely:

- Inputs
- Outputs
- State variables

Likewise, the dimension of the state space vector must be at least equal to the order of the system [8],

The representation of the state space of the system is a basic idea in the modern control theory, as the state defines the system as a studied minimum (specific) of present and future information and can be described as the future behavior of the system and depends on Markov properties that require that the state space is what gave the future to the system in a way. It is independent of its past, and accordingly, the representation of the system's state space is also called the Markov representation of the system.

The system can be a linear time constant if both are linear and constant over time. Fixed linear time systems that contain fixed properties such as fixed states of a fixed and linear time system can represent many physical characteristics. The representation of the state space is described by the state and output equations and can be described as follows [8]:

$$Yt+1 = AYt+GXt+1 \tag{1}$$

and the output equation:

$$Zt=HYt \tag{2}$$

Y_t :represent a specific vector with a dimension (k×1), A: represent transitional matrix dimension (k×k), G: represent input matrix dimension (k×m), Xt: represent the input vector of the system has a rank (m×1), H: represent output views matrix(m×k), Z_t: represent output vector rank(m×1).

If the input X_t and the output Z_t are stochastic operations, then the representation of the state space is given as follows:

$$Y_{t+1} = AY_t + Ga_{t+1} \tag{3}$$

$$Z_t = HY_t + b_t \tag{4}$$

As $a_{t+1} = X_{t+1} - E(X_{t+1}|X_t, t \leq n)$ It is the vector (n × 1) for one forward step to predict the error of the X_t process, and b_t is the vector (m × 1) for the noise and assumed to be independent of a_t.

If $Z_t = X_t$ fades from equation (4) and the representation of the state space for the stochastic process of the constant Z_t becomes as follows:

$$Z_t = HY_t \tag{5}$$

Thus, Zt's process is the output of a linear stochastic system with a fixed time indicating the white noise input a_t. As for Y_t, it is defined as the state of the system, the state equation is also known as the system equation or the transformation equation, and the resulting equation also indicates the measurement equation or the observed equation.

The representation of the system's state space is related to a Kalman filter, and initially, the idea was clarified in the engineering applications of the first to apply the concept of state space to ARMA model analysis [9].

A. The Relationship between the State Space Model and the ARMA Model

To represent the ARMA model as the state-space model in the state of a multivariate and a single variable, we assume the mean is constant and is zero and the ARMA model (p, q) is a vector with a dimension of (m) as follows:

$$\varphi(B)Z_t = \theta(B)a_t \tag{6}$$

$$Z_t = \varphi_1 Z_{t-1} + \ldots + \varphi_p Z_{t-p} + \theta_1 a_{t-1} \tag{7}$$

As:

$\varphi(B) = (1 - \theta(B) = (1 - \theta_1 B - \ldots - \theta_q B^q)$
$\varphi_1 B - \ldots - \varphi_p B^p)$

And a_t: multivariate with a dimension (m) and a zero mean for the white noise process of trapping (ϕ, θ) is used for both single-variable or multivariate states, and there should be no noise because the state is clear from the context of the speech, and the equation can be rewritten (6) in the form of a moving average and by multiplying by$\emptyset^{-1}(B)$, we get the following:

$$Z_t = \sum_{j=0}^{\infty} \Psi_j a_{t-j} \tag{8}$$

$$Z_{t+1} = \sum_{j=0}^{\infty} \psi_j a_{t+1-j} \tag{9}$$

Assume that:

$$Z_{t+1|t} = E[Z_{t+1}|Z_k, k \leq t] \tag{10}$$

Then:

$$= Z_{t+i|t} + \psi_{i-1} a_{t+1}$$

We assume and without loss of generality, that (p > q) plus $\Phi_i = 0$ if necessary, from equation (7) we obtain:

$$Z_{t+p|t} = \Phi_1 Z_{t+p-1|t} + \ldots + \Phi_p Z_t$$
$$Z_{t+p+1|t} = \Phi_1 Z_{t+p|t} + \ldots + \Phi_p Z_{t+1|t}$$
$$= f(Z_t, Z_{t+1|t}, \ldots, Z_{t+p-1|t})$$

It is clear that $Z_{t+p+i|t}$ (i>0) is a function of $(Z_t, Z_{t+1|t}, \ldots, Z_{t+p-1|t})$ the representation of the state space of the ARMA (p, q) model vector is given as follows:

$$\begin{bmatrix} Z_{t+1|t+1} \\ Z_{t+2|t+1} \\ \vdots \\ Z_{t+p|t+1} \end{bmatrix} = \begin{bmatrix} 0 & I & 0 & 0 & \ldots & 0 \\ 0 & 0 & I & 0 & \ldots & 0 \\ \vdots & & \ddots & & & \\ \varphi_p & \varphi_{p-1} & & & \ldots & \varphi_1 \end{bmatrix} \begin{bmatrix} Z_t \\ Z_{t+1|t} \\ \vdots \\ Z_{t+p-1|t} \end{bmatrix} + \begin{bmatrix} I \\ \Psi_1 \\ \vdots \\ \Psi_{p-1} \end{bmatrix} a_{t+1} \tag{11}$$

$$Z_t = [1 \quad 0 \ldots. \quad 0] \begin{bmatrix} Z_t \\ Z_{t+1|t} \\ \vdots \\ Z_{t+p-1|t} \end{bmatrix} \tag{12}$$

B. Kalman Filter

The recursive strategy [10] was developed based on a state-space representation called a Kalman filter, a general formula that handles stationary and non- stationary time series. The candidate is characterized by the recursive property, which allows us to take advantage of new data and information as times change. A Kalman filter can be defined as a recursive algorithm to find an optimal estimator for the state or parameter variable, i.e., the estimation error variance is minimal [11]. A Kalman filter has been used widely in many fields of application, especially those that have evolved and originated after the scientific revolution that occurred in communications engineering, computer, and space science, and from these applied fields: [12].

- Engineering applications include signal processing by satellites or radar tracking and navigation systems for spacecraft, aircraft, and cars.
- Computer applications: such as image processing, real-time graphics identification, and audio processing.
- Economic and statistical applications: such as forecasting economic indicators and in the field of statistical control and control theory (Control Theory).

The mathematical model of the filtering issue can be expressed by the state-space model, through which the system can be represented by the state (system) and viewing equations as follows [13]:

$$Z_t = \varphi_t Z_{t-1} + w_t$$
$$y_t = M_t Z_t + v_t$$

Our goal is to obtain an estimate of the state vector Z_t and estimate the covariance matrix P_t based on available information. Likewise, observation $y_0, y_1 \ldots, y_n$.
To facilitate the estimation process, we assume the following assumptions [14].

- The Z_t operation is a dimensional vector (p * 1) and the y_t operation observation is a dimensional vector (q * 1).
- The transformation matrix ϕ_t is not anomalous and has a dimension (p * p) and the matrix Mt is the observation matrix dimension (q * p).
- The noise process wt is a dimensional vector (p * 1) representing a strong white noise process so that:

$$E(w_t) = 0$$
$$E(w_t w_t') = Q$$

- The observation noise v_t is a dimensional vector (q * 1) representing a strong white noise operation so that:

$$E(v_t) = 0$$
$$E(v_t v_t') = R$$

- The Pt covariance matrix has a dimension (p * p) that is positive, defined, and written as:

$$P_t = E(\hat{Z}_t - Z_t)(\hat{Z}_t - Z_t)'$$

III. Results and Discussion

In this part of the research, data are presented in the amount of daily electrical energy consumption for the city of Mosul (mW / hour). During the period from 15/6/2003 to 25/9/2003, the description of the time series was traced by drawing what is known as the time series plot. We also investigate and learn about the nature of the oscillations in it and whether it includes a trend or not. The general trend is beneficial in confirming the stationary of the time series and its close relationship to forecasting.

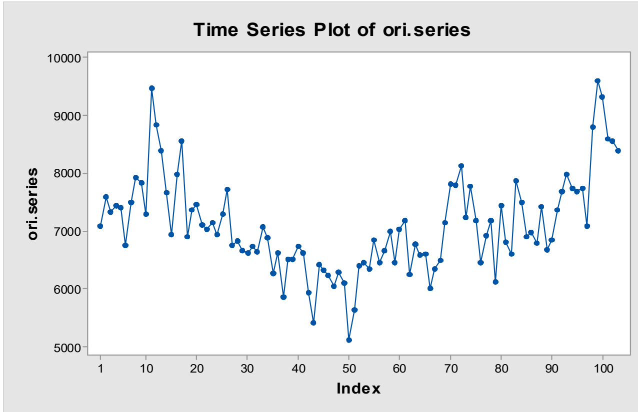

Fig.1 The timeline of the electric energy consumption in Mosul

Figure 1 shows the drawing of the time series that we are dealing with, and it is noted that the fluctuation of the series begins with a gradual decline, then it takes a gradual rise and a parabolic pattern of the second degree, which indicates the non- stationery of this series. On the other hand, we noted a clear dispersion around the general path of oscillation, which confirms the presence of clear and influential random effects in this series.

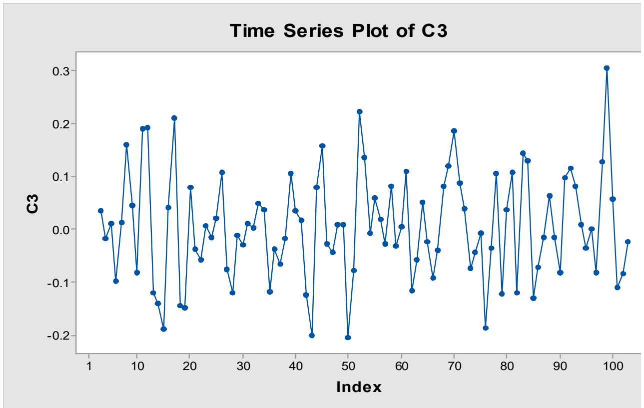

Fig. 2 Observational behavior after converting the time series to a stationary series.

To convert the time series to a stationary series, predicting the classical methods requires that the series be stationary. The logarithmic transformation of the original series observations was performed as they are not stationary by variance, then Differences are taken, and they start with the first difference (ΔX_t) and then the second difference ($\Delta^2 X_t$). After obtaining the stationary in the behavior of the series, as shown in Figure 2.

In order to determine the appropriate model and its rank, the auto-correlation function (ACF) and the partial autocorrelation function (PACF) had been calculated shown in Figure 3 as follows:

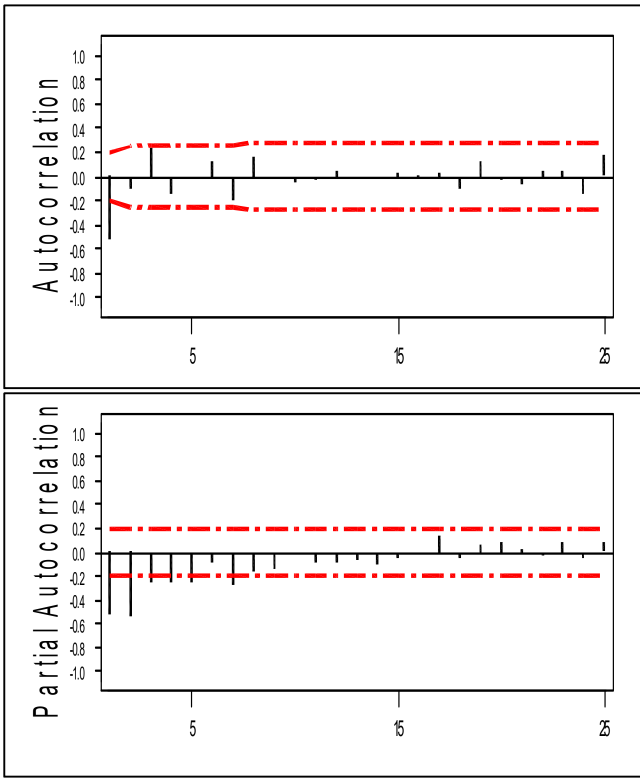

Fig. 3 The autocorrelation function and the partial correlation function of the stationary series.

Figure 3 above represents the auto-correlation function and the partial correlation function of the stationary series. The appropriate series model is the ARMA(2,1) model also the AIC standard was used to choose the model, and the following table shows the appropriate results for a group of ARMA models (p, q) and different values from (p) and (q) as well as the AIC standard value in each state.

TABLE I

FITTING ARMA (P, Q) MODEL OF TRANSFERRED DATA

P	Q	MSE	AIC	$\hat{\sigma}_{\varepsilon}^{2}$
0	1	0.007314	-495.982	0.007223
0	2	0.007179	-496.686	0.007032
0	3	-	-	-
1	0	0.01348	-433.976	0.013345
2	0	0.009625	-467.028	0.009432
3	0	0.009045	-472.337	0.008774
1	1	0.006782	-502.682	0.006627
1	2	0.007292	-494.322	0.007057
1	3	0.006360	-507.372	0.00608
2	1	0.006243**	-510.221*	0.006029
2	2	0.006354	-507.262	0.006087
2	3	0.006344	-506.641	0.006004
3	1	0.006366	-508.852	0.005992
3	2	0.006419	-505.406	0.006078
3	3	0.006444	-504.075	0.006038

* The lowest value for a standard AIC.

** The lowest value for the squared error mean.

When testing a set of ARMA (p, q) models and for different values of (p) and (q) it was found that the minimum AIC value at the ARMA (2,1) model and the estimated final model is as follows:

$$X_t = -0.3882X_{t-1} - 0.3369X_{t-2} + \varepsilon_t + 0.9876\varepsilon_t$$

To confirm the ARMA (2,1) model's validity, the auto-correlation function and the partial autocorrelation function of the residuals shown in Figure 4 and found to fall within the confidence limits, indicating that the residues of this model are not correlated with each other.

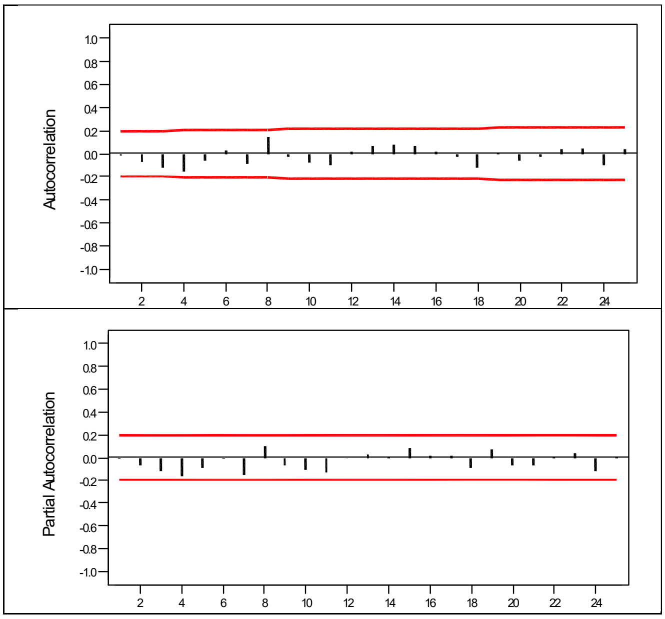

Fig. 4 The autocorrelation function and the partial correlation function for the ARMA residual (2,1).

Using equations (11) and (12) that represent the state space of the ARMA model can be transformed ARMA(2,1) model in the above equation to represent the state space as in the following equation

$$\begin{bmatrix} Z_{t|t} \\ Z_{t+1|t} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ \varphi_2 & \varphi_1 \end{bmatrix} \begin{bmatrix} Z_{t-1} \\ Z_{t-2|t} \end{bmatrix} + \begin{bmatrix} 1 \\ \psi_1 \end{bmatrix} a_t$$

$$\begin{bmatrix} Z_{t|t} \\ Z_{t+1|t} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0.3369 & 0.3882 \end{bmatrix} \begin{bmatrix} Z_{t-1} \\ Z_{t-2|t} \end{bmatrix} + \begin{bmatrix} 1 \\ 0.5994 \end{bmatrix} a_t$$

$$\psi_1 = \phi_1 - \theta_1$$

$$= -0.3882 + 0.9876 = 0.5994$$

The following equations show the mathematical model of the filtering issue obtained through the state space model:

State equation:

$$\begin{bmatrix} Z_{t|t} \\ Z_{t+1|t} \end{bmatrix} = \begin{bmatrix} 0 & 0.6285 \\ 0.3369 & 0.0672 \end{bmatrix} \begin{bmatrix} Z_{t-1} \\ Z_{t-2|t} \end{bmatrix} + \begin{bmatrix} 0.3715 \\ 0.321 \end{bmatrix} w_t$$

Observation equation:

$$y_t = [0.5248n \quad n0.3715] \begin{bmatrix} Z_{t|t} \\ Z_{t+1|t} \end{bmatrix} + v_t$$

The last transformation matrix was obtained after several iterations, and the state covariance and observation covariance were as follows:

$$\hat{Q} = \begin{bmatrix} 0.00622 & 0.00045 \\ 0.00045 & 0.00621 \end{bmatrix}$$

$$\hat{R} = [0.0058]$$

Figure 5 below shows the time series of electrical energy consumption before filtration and after filtration, and it is clear from the figure how efficient a filter such as security is in purifying the chain of noise.

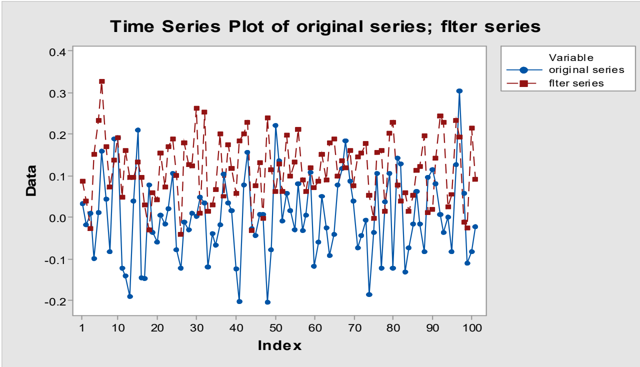

Fig. 5 Time series of electric power consumption before and after filtration.

It is possible to clarify some statistical measures of localization and dispersion to compare the filtering efficiency as shown in Table 2, as we note from the table that the filtering process significantly reduced the dispersion of data, which indicates the efficiency of a Kalman filter such as security in the filtering process for the series of noise data.

TABLE II
COMPARING THE ORIGINAL AND FILTERING DATA

Data	Mean	SE Mean	Standard deviation	Coefficient of variation
Original Data	0.0029	0.0100	0.1010	34.828
Candidate Data	0.1142	0.0078	0.0784	0.6865

IV. CONCLUSION

The time series for the consumption of electrical energy is non-stationary, so the logarithm and the second difference of data were taken to obtain a time-series of stationary by variance and arithmetic mean, and then drawing the auto-correlation and partial auto-correlation functions. Hence, the best model representing the data is ARMA (2,1).

The state-space model is characterized by being an efficient scale in all states that are not observed or controlled, and for this, the state-space model can be used to estimate states that cannot be observed. It can also express the state-space model simply for complex operations and is characterized by the flexible model.

A Kalman filter enjoy despite his complex calculations, has important and very useful features in different applications due to his dependence on the representation of the state space, we can immediately filter the time series and rely on the information.

REFERENCES

[1] Chen, K., & Yu, J. (2014). Short-term wind speed prediction using an unsented Kalman filter based state-space support vector regression approach. Applied Energy, 113, 690-705.

[2] Morley,J. C. (2002). A state–space approach to calculating the Beveride–Nelson decomposition. Economics Letters, 75(1), 123-127.

[3] Koopman, S. J., & Durbin, J. (2003). Filtering and smoothing of state vector for diffuse state-space models. Journal of Time Series Analysis, 24(1), 85-98.

[4] Kim, C. J. (1994). Dynamic linear models with Markov-switching. Journal of Econometrics, 60(1-2), 1-22.

[5] Sharaf, H. K., Ishak, M. R., Sapuan, S. M., & Yidris, N. (2020). Conceptual design of the cross-arm for the application in the transmission towers by using TRIZ–morphological chart–ANP methods. Journal of Materials Research and Technology, 9(4), 9182-9188.

[6] Salmi, Jussi; Richter, Andreas; Koivunen, Visa. Detection and tracking of MIMO propagation path parameters using state-space approach. IEEE Transactions on Signal Processing, 2008, 57.4: 1538-1550.

[7] Kitagawa, G., & Gersch, W. (1984). A smoothness priors–state space modeling of time series with trend and seasonality. Journal of the American Statistical Association, 79(386), 378-389.

[8] Sharaf, H. K., Ishak, M. R., Sapuan, S. M., Yidris, N., & Fattahi, A. (2020). Experimental and numerical investigation of the mechanical behavior of full-scale wooden cross arm in the transmission towers in terms of load-deflection test. Journal of Materials Research and Technology, 9(4), 7937-7946.

[9] Ojeda, L. L., Kibangou, A. Y., & De Wit, C. C. (2013, June). Adaptive Kalman filtering for multi-step ahead traffic flow prediction. In 2013 American Control Conference (pp. 4724-4729). IEEE.

[10] Shukur, O. B., & Lee, M. H. (2015). Daily wind speed forecasting through hybrid KF-ANN model based on ARIMA. Renewable Energy, 76, 637-647.

[11] Sharaf, H. K., Salman, S., Dindarloo, M. H., Kondrashchenko, V. I., Davidyants, A. A., & Kuznetsov, S. V. (2021). The effects of the viscosity and density on the natural frequency of the cylindrical nanoshells conveying viscous fluid. The European Physical Journal Plus, 136(1), 1-19.

[12] Watson, M. W., & Engle, R. F. (1983). Alternative algorithms for the estimation of dynamic factor, mimic and varying coefficient regression models. Journal of Econometrics, 23(3), 385-400.

[13] Sharaf, H. K., Salman, S., Abdulateef, M. H., Magizov, R. R., Troitskii, V. I., Mahmoud, Z. H., ... & Mohanty, H. (2021). Role of initial stored energy on hydrogen microalloying of ZrCoAl (Nb) bulk metallic glasses. Applied Physics A, 127(1), 1-7.

[14] Pfeffermann, D., & Tiller, R. (2005). Bootstrap Approximation to Prediction MSE for State–Space Models with Estimated Parameters. Journal of Time Series Analysis, 26(6), 893-916