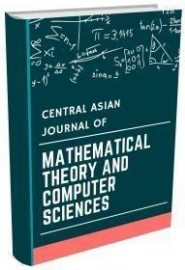

CENTRAL ASIAN JOURNAL OF MATHEMATICAL THEORY AND COMPUTER SCIENCES
https://cajmtcs.centralasianstudies.org/index.php/CAJMTCS
*Volume: 05 Issue: 02 | April 2024 ISSN: 2660-5309*

*Article*

# Forecasting the Number of Patients with Kidney Failure in Thi-Qar Governorate using Time Series

**Hassan Hopoop Razaq**[1]

1. College of Administration and Economics, Department of Economics, University of Thi –Qar
* Correspondence: hasanhopop@utq.edu.iq

**Abstract:** This paper has been used time series models for the study and analysis of monthly data to the number of patients with Kidney Failure in Thi-Qar Province for the period (2020-2023) in order forecasting by the numbers of patients with Kidney Failure for the period (2024-2025). The result of data analysis show that the proper and suitable model is Integrated Autoregressive model of order ARIMA (4, 0, 1) because it has the least mean squares error (MSE). Based on the best model, the number of people with Kidney failure was predicted monthly and for the next two years and the predictive value was consistent with the original values and this indicates the efficiency of the model.

**Keywords:** Autoregressive Model, Time Series Analysis, Mean Squares Error, ARIMA Model

## 1. Introduction

The topic of time series is one of the basic topics that has gained very wide use in various sciences, as mathematical statistics have been lost in the analysis of time series. As a result of these analyzes, important functions have begun to be used for estimation, and even for other very important points in choosing many topics and to help in the work of some research and studies. Mathematics for the intended problem. During the past thirty years, our country has gone through the disasters of wars that affected its material and human resources, destroyed its infrastructure and polluted its water and air, which requires a comprehensive renaissance in all fields and economic activities [1]. This is what happens with the intensification of the efforts of researchers in all specializations to conduct studies and research that would reduce what afflicted the country [2], [3]. Of pollution, diseases and pests that affected the health, agricultural and industrial aspects, so this study came to address the health aspect due to its importance on the developmental level because it is concerned with the human element, which is responsible for construction and reconstruction and keeping pace with progress and civilizational development [4], [5].

One of the components of building health is preventing all diseases, including kidney failure, which causes high rates of death. Given the recent increase in the number of people infected with this disease, this study came in order to reveal this phenomenon, which has increased in Thi-Qar Governorate, one of the governorates affected by bacterial and biological weapons. And the severe shortage in health and therapeutic care. The study relied on monthly data on the numbers of people infected with kidney failure for the period (2020-2023) as a time series for the purpose of analyzing it to reach the best model to

**Citation:** Razaq, H. H. Forecasting the Number of Patients with Kidney Failure in Thi-Qar Governorate using Time Series. Central Asian Journal of Mathematical Theory and Computer Sciences 2024, 5(2), 24-35.

Received: 24th Jan 2024
Revised: 28th Jan 2024
Accepted: 15th Feb 2024
Published: 27th Feb 2024

predict the numbers of people infected with this disease for subsequent periods in order to take the necessary measures to reduce this phenomenon in the future.

## 2. Materials and Methods

The following is a summary of the methods used to use time series analysis to anticipate the number of people in Thi-Qar Governorate who would have renal failure: gathering, preprocessing, analyzing time series, evaluating models, and forecasting data. The investigation and analysis of monthly data on the number of patients with kidney failure in Thi-Qar Province for the period (2020–2023) was done in this work using time series models, which were arranged in order of patient numbers for the period (2024–2025). Because it has the lowest mean square error (MSE), the integrated autoregressive model of order ARIMA (4, 0, 1) is the most appropriate and acceptable model, according to the data analysis results. The most accurate model projected the monthly number of individuals with kidney failure for the next two years and the predictive value was consistent with the original values and this indicates the efficiency of the model.

Study the time series and use one of the time series models as a statistical model for the purpose of using it to predict the numbers of people with kidney failure in Thi-Qar Governorate for the period (2020-2023) through the use of the Minitab program. This item deals with a review of some general concepts by presenting the stages of building a time series model based on the algorithm drawn up by the researchers (Box & Jenkins) in 1976 [6], which begins with the first stage, which is diagnosing the appropriate model for the data, followed by the stage of estimating the parameters of the diagnosed model, and then comes the examination stage. Fitting the diagnosed model. If the model is appropriate, the final stage comes, which is the stage of future prediction.

## 3. Results and Discussion

### *3.1. Basic Concept*

#### *3.1.1. Forecasting*

Demand forecasting is defined as an attempt to estimate a need for a specific good or service, a specific phenomenon, or a mixture of goods during a future period of time. It is also known as the art and science of anticipating events in the future.

#### *3.1.2. Definition of Time Series*

It is a set of observations of a specific phenomenon during a period of time. The time series is defined mathematically as a sequence of random variables defined within a multivariate probability space and indexed by the index t, which refers to an index set T. The time series is usually symbolized {X(t), t} or for short X(t) it consists of two variables, one of which is explanatory, which is the time variable, and the other is the response variable, which is the value of the phenomenon studied. The statistical series can be represented as follows:

$$\boldsymbol{X_t = f(t) + \alpha_t} \quad , \quad \boldsymbol{t = 0, \pm 1, \pm 2, \ldots} \qquad (\mathbf{1})$$

Where is,

$f(t)$ = Represents the regular part expressed by a mathematical function

$\boldsymbol{\alpha_t}$ = Represents the random part and may be called noise

Time series can be represented in a graphical form, and the time series can be of a deterministic type, for example:

$$X_t = Cosin2\pi f_{(t)} \quad , \quad t = 0, \pm 1, \pm 2, \ldots \qquad (2)$$

*3.1.3. Time Series Stationary*

The stationary and instability of data is important in analyzing time series as well as in finding the appropriate mathematical model for it, and drawing the time series in the period $(t,\ t\ +\ h)$ may sometimes be identical to drawing the series in another period $(s,\ s\ +\ h)$ and this It indicates that there is temporal uniformity in the behavior of the series, which is called stationary. One of the conditions for time series is that the time series $X_t$ be stable as well it is assumed that the model parameters are known, which achieves the least value of the mean squares error, this means:

$$\underline{X}_{t+m} = E(X_{t+m}), m = 1\ ,2\ , \ldots \qquad (3)$$

We can say that the time series is stable based on the graph of observations, as well as if it has an arithmetic mean and variance that is free of effects. It is stationary under the conditions below:

1) Arithmetic average is constant, that is $E(X_t) = \mu$
2) The value of the variance is constant, that is $Var(X_t) = \sigma_X^2$
3) Having the two series $Y_{t+h}$, $Y_t$ has a correlation and variance that depends on the displacement h only, that is $\gamma_h = E(X_t - \mu)(X_{t+h} - \mu)$ it depends on the absolute value of $h$ only, $h = 1,2, \ldots , m,$ and it can also be verified using autocorrelation functions by using the chi-square measure:

$$X^2_{(m-1)} = n\sum_{h=1}^{m} P_h^2 \qquad (4)$$

Where is,

$P_h$ = The autocorrelation of the Y values is represented by a shift of $h$

$m$ = The largest time regression period and is usually equal to $(\frac{n}{2})$ the calculated Chi-square value is compared with the tabulated value.

In the case that the time series is no stationary on average, that is, it does not have stationary in the general trend, we resort to taking difference operations to make the time series, which is symbolized by the symbol $\nabla$, by applying the following formula:

$$\nabla X_t = X_t - X_{t-1} = (1 - \beta)X_t \qquad (5)$$

Where $\beta$ is called the back difference indicator. Thus, the time series becomes stable after taking $(d)$ from the differences, i.e

$$X_t = \nabla^d X_t \quad , \quad d \geq 1 \qquad (6)$$

If the time series is no stationary in variance, it is treated through data transformation methods such as natural logarithmic transformation, logistic transformation, natural root transformation, and Box-Cox transformation. Time series transformations may lead us to find a stationary time series. In general, ARIMA models and time series transformations give important functions for estimation, and this case is similar to fitting models in the case of smooth.

*3.2. Time Series Models*

Non-seasonal time series models include stationary and no stationary ones. Below are the types of common non-seasonal time series model:

*3.2.1. Autoregressive model (AR)*

One of the first to study stationary time series models was the scientist Yule in 1926, where he studied the autoregressive model $AR(P)$ and continued his path to the general model of autoregressive models. The general formula for this model is of rank $(P)$ and abbreviated $AR(P)$ is:

$$X_t = C + \emptyset_1 X_{t-1} + \emptyset_2 X_{t-2} + \cdots + \emptyset_p X_{t-p} + a_t \qquad (7)$$

Where is,

$at$ = Represents random error, which has a normal distribution with a mean equal to zero and variance $\sigma_a^2$

$C$ = Represent a constant

$\emptyset_1, \emptyset_2, \ldots, \emptyset_p$ = Represent parameters of the autoregressive model where $-1 < \emptyset < 1$. The autocorrelation function decreases exponentially with increasing displacement periods $h$, while the partial autocorrelation function $(PACF)$ cats after the period $P$. For example, when $P = 1$, that in the case of $AR(1)$, then the above equation becomes as follows:

$$X_t = C + \emptyset_1 X_{t-1} + a_t \qquad (8)$$

*3.2.2. Moving Average Model (MA)*

The moving average models $MA(q)$ and developed the general formula for this model of rank $(q)$ and in short $MA(q)$ is:

$$X_t = C - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \cdots - \theta_q a_{t-q} + a_t \qquad (9)$$

Where $a_t$, $C$ are explained above. $\theta_1, \theta_2, \ldots, \theta_q$ represent the parameters of moving averages and $-1 < \theta < 1$. The above equation represents the moving average model of rank $q$ and the model rank is determined from its autocorrelation function, which cats after a period of $q$ while the partial autocorrelation function gradually decreases in a descending curve. For example, when $(q = 2)$, in the case of $MA(2)$, equation (9) is in the form of the following:

$$X_t = C - \theta_1 a_{t-1} - \theta_2 a_{t-2} + a_t \qquad (10)$$

*3.2.3. Mixed Autoregressive Moving Average Model (ARMA)*

The researcher Slutzky completed the creation of the model in a mixed form, and the researcher Wold completed the path in 1938, when he developed these two models with a series of operations in three directions in performing the estimation, and he called them the processes of autoregressive models and moving averages, as this model represents a mixture of, $AR(P)$ model with $MA(q)$ model and is symbolized by its abbreviation. $ARMA(P,q)$ is used if the data is stationary. The general formula for this model of rank $(P,q)$ is:

$$X_t = C + \emptyset_1 X_{t-1} + \emptyset_2 X_{t-2} + \cdots + \emptyset_p X_{t-p} + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \cdots - \theta_q a_{t-q} \qquad (11)$$

Autocorrelation and partial autocorrelation functions gradually decrease. For example $ARMA(1,1)$ is as follows:

$$X_t = \emptyset_1 X_{t-1} - \theta_1 a_{t-1} + C + a_t \qquad (12)$$

*3.2.4. Integrated mixed models (ARIMA)*

Box & Jenkins 1976 described the models comprehensively and put together a method or approach to related information in understanding and treating stationary in data, and they arrived at the model called the autoregressive model and integrated moving averages. This models contain of three models, first model is $AR(P)$, which is usually used in the forecasting process for time series. The second model is $MA(q)$ model and the third part $I(d)$ represents the differences that the series requires in order for it to be stationary. Therefore, it expresses Auto Regressive Integrated Moving Average Models that are not Seasonality according to the $ARIMA$ formula of the rank $(p,\ d,\ q)$ where $q$ rank of $AR$ model, $q$ rank of $MA$ model and d the number of variances that make the series stationary. $ARIMA$ is used more than other models, so are all models can be derived from them, whether autoregressive, moving averages, or mixed. The integrated mixed model is written in the form $ARIMA\ (p, d, q)$ and takes the following formula:

$$\emptyset(\beta)(1-\beta)^d Z_t = C + \theta(\beta)a_t \qquad (13)$$

Where,

$\emptyset(\beta) = 1 - \emptyset_1\beta - \emptyset_2\beta^2 - \cdots - \emptyset_p\emptyset^p$

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

Let $\nabla^d Z_t = X_t$ then the general formula of the integrated mixed model is:

$$X_t = C + \emptyset_1 X_{t-1} + \cdots + \emptyset_p X_{t-p} + \cdots + dX_{t-p-q} + a_t - \theta_1 a_{t-1} - \cdots - \theta_q a_{t-q} \quad (14)$$

*3.3. Application*

The data was collected, which consists of a time series consisting of 48 observations, dating back to the period from January 2020 until December 2023. This data represents the number of people with kidney failure in Thi-Qar Governorate and was taken from the records of Al-Hussein Teaching Hospital, as shown in Table 1.

**Table 1**. Number of people with kidney failure

| Month \ Year | 2020 | 2021 | 2022 | 2023 |
|---|---|---|---|---|
| January | 120 | 166 | 170 | 191 |
| February | 132 | 187 | 176 | 230 |
| March | 150 | 143 | 180 | 224 |
| April | 161 | 162 | 180 | 222 |
| May | 143 | 174 | 183 | 244 |
| June | 192 | 190 | 175 | 248 |
| July | 155 | 137 | 200 | 241 |
| August | 131 | 141 | 214 | 252 |
| September | 176 | 194 | 209 | 259 |
| October | 198 | 184 | 218 | 251 |
| November | 178 | 178 | 214 | 223 |
| December | 167 | 188 | 230 | 223 |

*3.4. Identification Model*

To choose the best model from the time series models for the numbers of people with kidney failure, the mean squares error (MSE) criterion was used, which is given according to the following formula:

$$MSE = \frac{\sum_{i=1}^{n} (Y_t - \hat{Y}_t)^2}{n - (k+1)}$$

Where $n$ number of observation , $k$ number of parameters.

*3.5. Data Format*

To know the pattern taken by the data shown in Table 1, it is necessary to plot the time series and the autocorrelation functions $ACF$ and partial autocorrelation $PACF$ for the numbers of people with kidney failure as in the following figures:

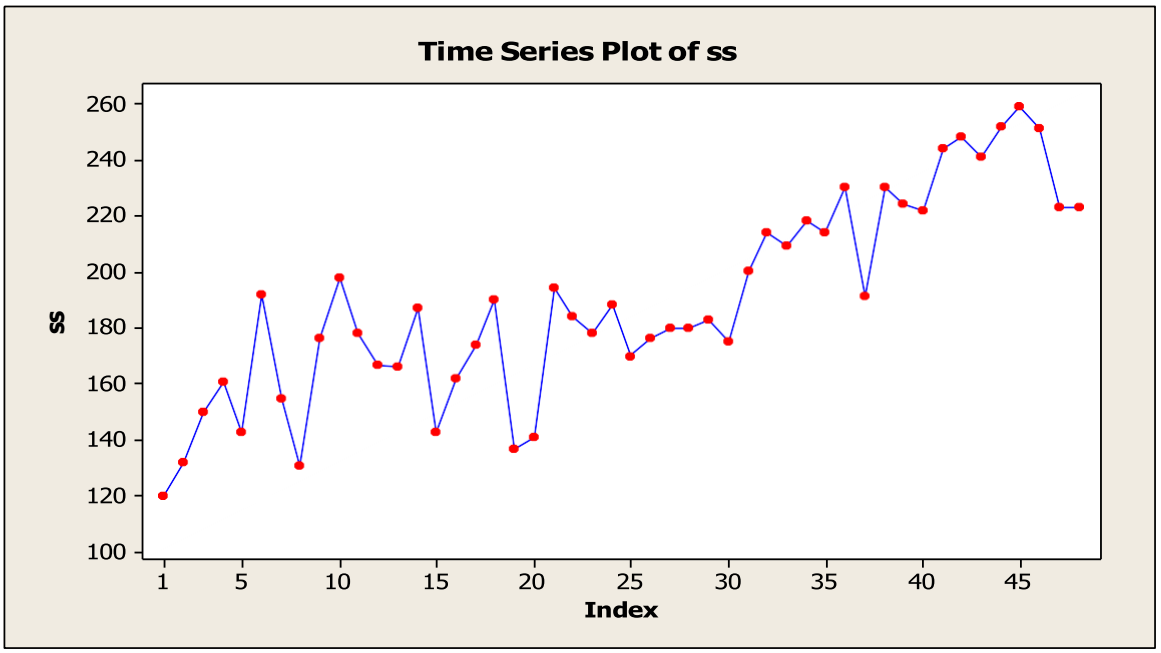


**Figure 1**. The frequency curve of the time series (the curve of the number of people with kidney failure)

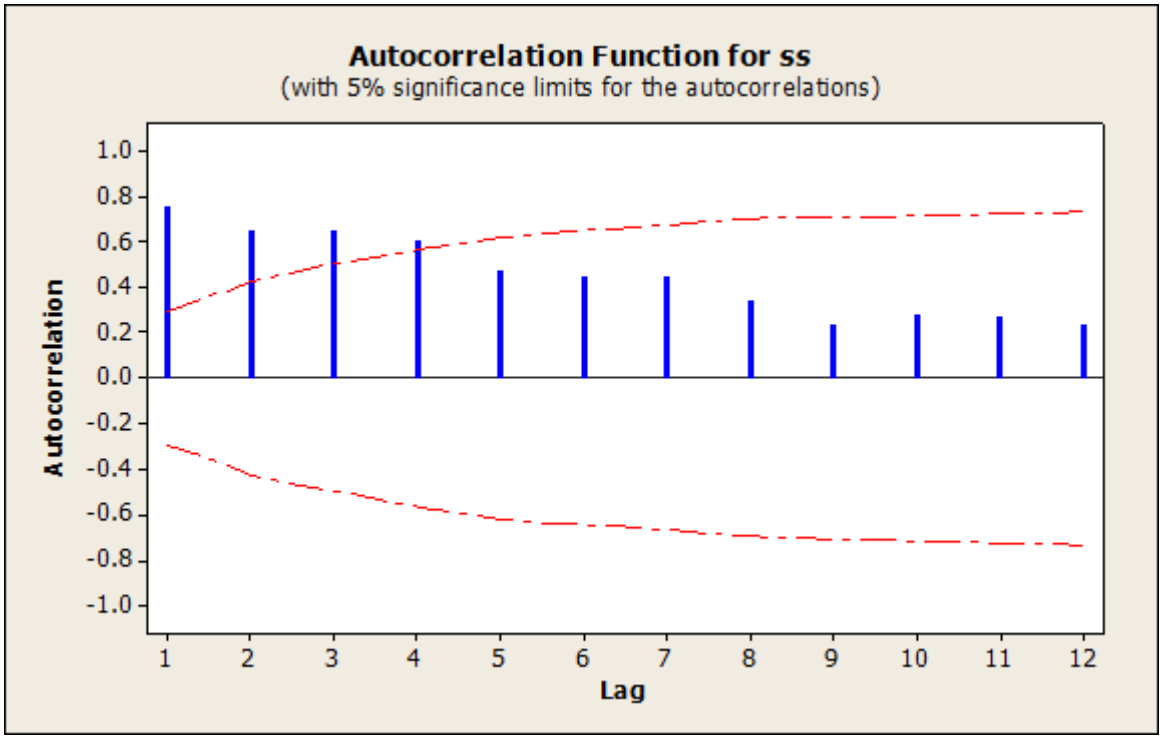


**Figure 2**. Autocorrelation function for the time series

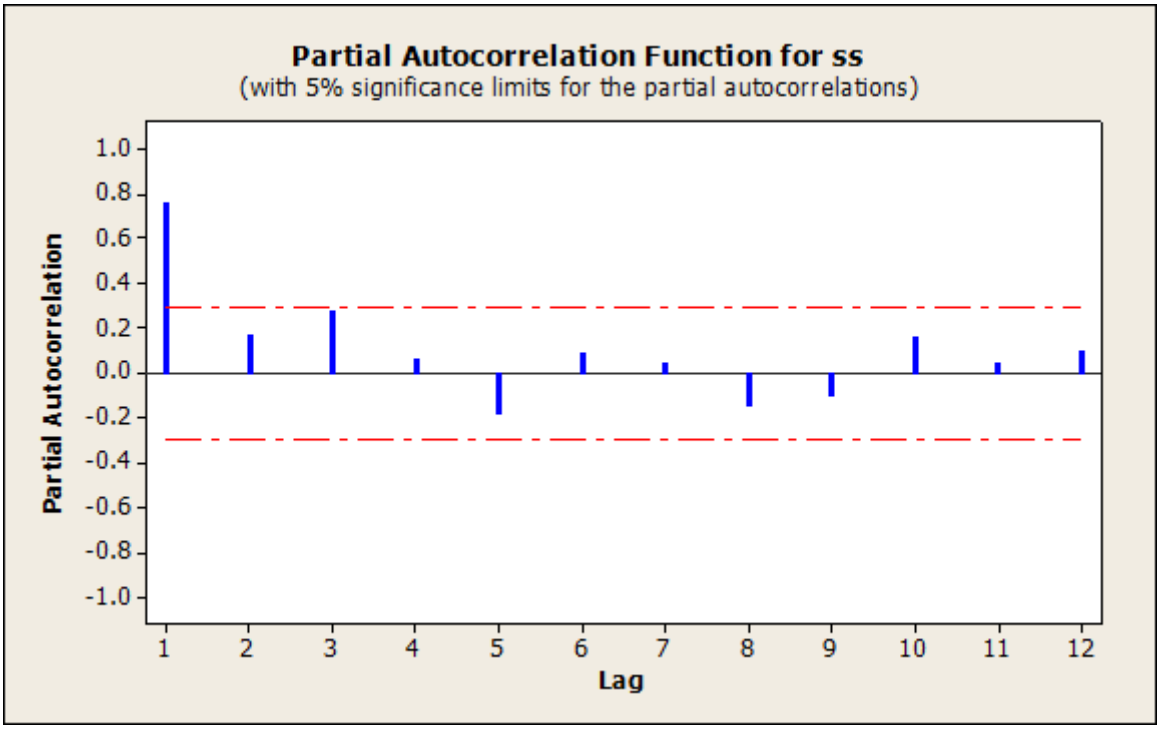


**Figure 3**. Partial autocorrelation function for the time series

By plotting the time series, $ACF$ and $PACF$ functions a group of models close to each other was chosen, which are considered the best models. By drawing the time series and the behavior of $ACF$ and $PACF$ functions, and to choose the best model by matching these coefficients with the theoretical behavior described for them, we relied on Another method for determining the appropriate model is by using a table of residual indicators

for the applied models, which suggests that the best model has the lowest $(MSE)$, as in Table 2.

**Table 2**. Mean Square Error of models

| ARIMA Models | MSE |
|---|---|
| ARIMA(1,1,1) | 403.6 |
| AR(4) | 371.3 |
| ARIMA(2,1,0) | 382.7 |
| ARIMA(3,1,0) | 371.8 |
| ARIMA(4,1,0) | 377.6 |
| ARIMA(4,0,1) | 365.7 |
| ARIMA(3,1,1) | 399.8 |
| ARIMA(3,0,1) | 378 |

We conclude that the lowest value for the comparison criterion shown above is carried by the sixth model, and therefore the $ARIMA\ (4,\ 0,\ 1)$ model is the best model and is the appropriate model for representing time series data.

*3.6. Estimation*

The results below are estimated parameters $ARIMA\ (4,\ 0,\ 1)$ model alculated using the program Minitab and applying the Ordinary Least Square method. On the time series data the following results were obtained:

| Type | Coef | SE Coef | T | P |
|---|---|---|---|---|
| AR 1 | -0.3759 | 0.1396 | -2.69 | 0.010 |
| AR 2 | 0.4771 | 0.1400 | 3.41 | 0.001 |
| AR 3 | 0.3776 | 0.1357 | 2.78 | 0.008 |
| AR 4 | 0.5265 | 0.1374 | 3.83 | 0.000 |
| MA 1 | -0.9554 | 0.1183 | -8.07 | 0.000 |

Number of observations: 48
Residuals: SS = 15726.4 (backforecasts excluded)
MS = 365.7 DF = 43

Modified Box-Pierce (Ljung-Box) Chi-Square statistic

| Lag | 12 | 24 | 36 | 48 |
|---|---|---|---|---|
| Chi-Square | 10.3 | 27.4 | 39.0 | * |
| DF | 7 | 19 | 31 | * |
| P-Value | 0.171 | 0.095 | 0.152 | |

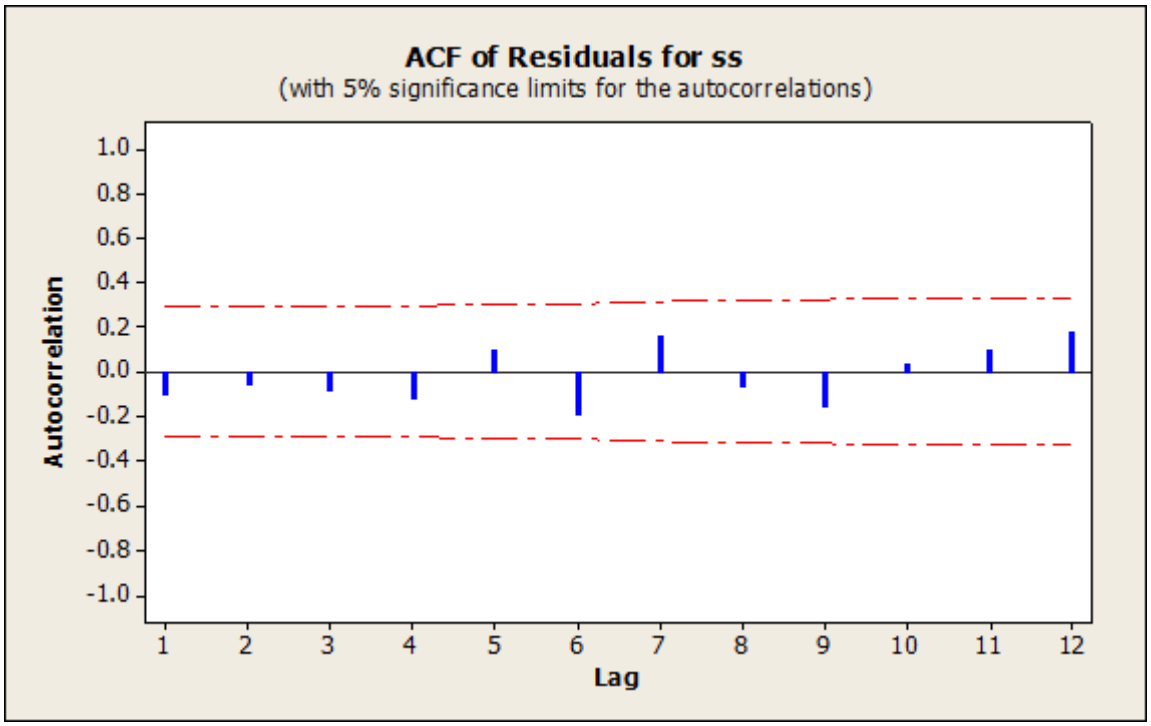


**Figure 4**. *ACF* for the residuals of the estimated model

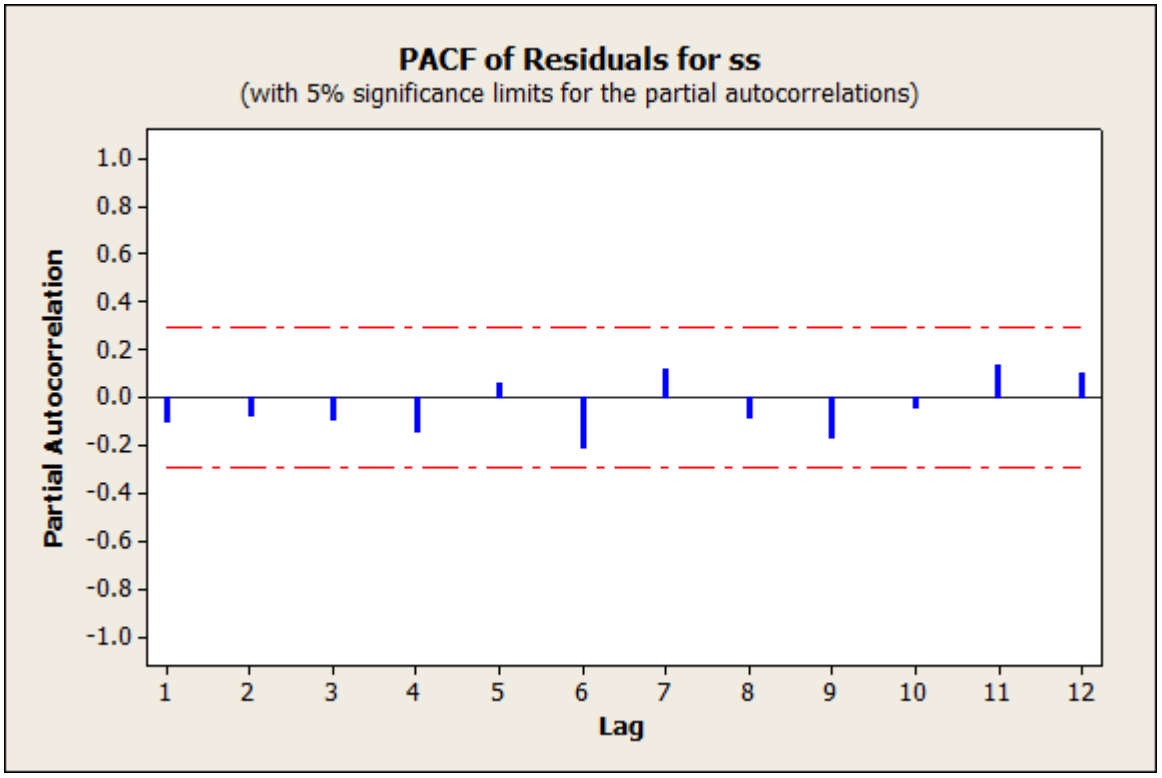


**Figure 5**. *PACF* for the residuals of the estimated model

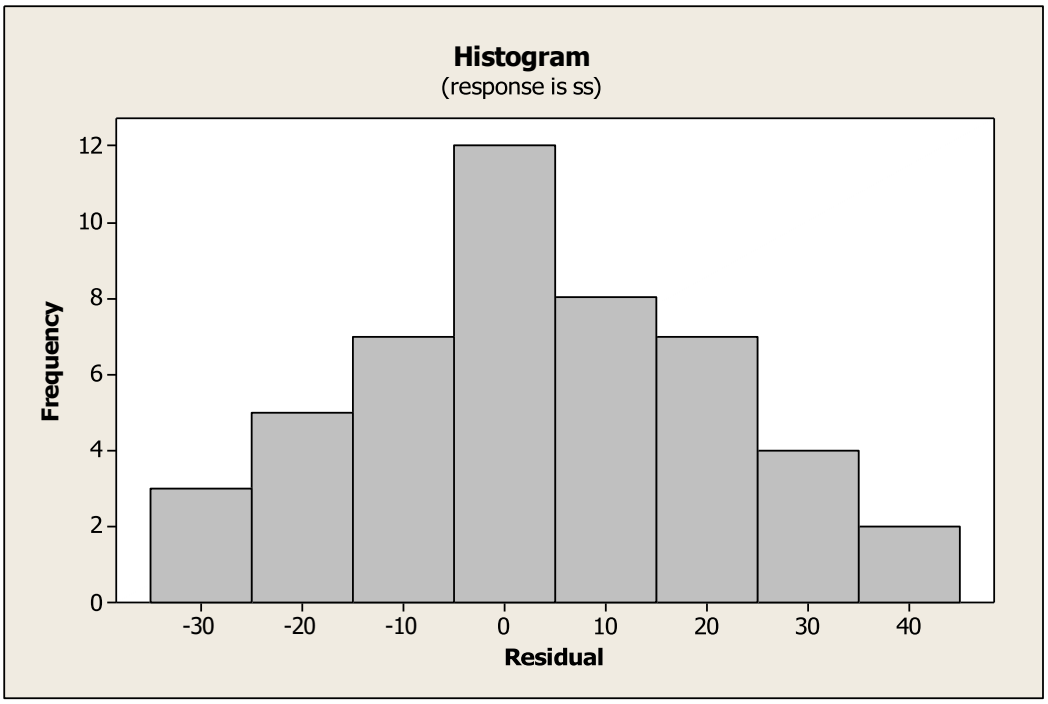


**Figure 6**. Normal distribution of the estimated model residuals

*3.6. Forecasting*

In this item, the model in Paragraph (8) is used to forecasting the number of people with kidney failure in Thi-Qar Governorate for the period (2024-2025) as follow:

**Table 3**. Number of people with kidney failure

| Month \ Year | 2024 | 2025 |
|---|---|---|
| January | 232 | 230 |
| February | 235 | 236 |
| March | 223 | 231 |
| April | 233 | 237 |
| May | 230 | 232 |
| June | 233 | 238 |
| July | 228 | 232 |
| August | 235 | 238 |
| September | 229 | 233 |
| October | 234 | 239 |
| November | 230 | 234 |
| December | 236 | 240 |

## 4. Conclusion

We conclude that the efficient and appropriate model is the integrated autoregressive model, $\boldsymbol{ARIMA}\ (\mathbf{4}, \mathbf{0}, \mathbf{1})$, it used to forecast the number of people with kidney failure in Thi-Qar Governorate for the period (2024-2025). We recommend taking into account the results of this research, which shows an increase in the number of people suffering from kidney failure over time, which requires taking the necessary measures by the competent authorities to reduce this phenomenon.

**REFERENCES**


[1] A. C. Harvey and N. Shephard, *10 Structural time series models*. Elsevier, 1993. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0169716105800458/pdf?md5=ee36af7a98c638a97873b83ba525535a&pid=1-s2.0-S0169716105800458-main.pdf

[2] A. Zellner and F. Palm, "Time series analysis and simultaneous equation econometric models," *J Econom*, 1974, [Online]. Available: https://www.academia.edu/download/86394186/zellner_palm_je74.pdf

[3] F. Palm, "Time series analysis and simultaneous equation models with macroeconomic applications," *Co-promoteur avec A. Zellner). Professeur, Universiteit …*, 1975.

[4] C. W. J. Granger and M. J. Morris, "Time series modeling and interpretation," *… , Seasonality, Nonlinearity, Methodology, and …*, 2001, doi: 10.5555/766886.766898.

[5] F. Harris and R. M. Gwier, "A receiver structure that performs simultaneous spectral analysis and time series channelization," *Proceedings of the SDR'09 Technical Conference and …*, 2009.

[6] G. E. P. Box, G. M. Jenkins, G. C. Reinsel, and G. M. Ljung, *Time series analysis: forecasting and control*. books.google.com, 2015. [Online]. Available: https://books.google.com/books?hl=en&lr=&id=rNt5CgAAQBAJ&oi=fnd&pg=PR7&dq=time+series+analysis+forecasting+and+control&ots=DL2-wPhYWF&sig=rXC9u6Y10n1MAPSPTtdJTx1rxwM

[7] M. G. Dekimpe and D. M. Hanssens, "Time-series models in marketing:: Past, present and future," *International journal of research in marketing*, 2000, [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0167811600000148

[8] B. Lim and S. Zohren, "Time-series forecasting with deep learning: a survey," *… Transactions of the Royal Society A*, 2021, doi: 10.1098/rsta.2020.0209.

[9] S. L. Zeger, R. Irizarry, and R. D. Peng, "On time series analysis of public health and biomedical data," *Annu. Rev. Public Health*, 2006, doi: 10.1146/annurev.publhealth.26.021304.144517.

[10] R. H. Sumway and D. S. Stoffer, "Time series analysis and its applications with R examples," *Time series analysis and its applications with R …*, 2006.

[11] T. W. Anderson, *The statistical analysis of time series*. books.google.com, 2011. [Online]. Available: https://books.google.com/books?hl=en&lr=&id=rCOzXIC8ZLkC&oi=fnd&pg=PR11&dq=the+statistical+analysis+of+time+series&ots=Iv-1niRFKt&sig=TqoOezBOlsoEhHS4TWMGjPDd4rs

[12] A. Lugmayr, "RePaint: Inpainting using Denoising Diffusion Probabilistic Models," *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 2022, pp. 11451–11461, 2022, doi: 10.1109/CVPR52688.2022.01117.

[13] F. Chien, "The role of renewable energy and urbanization towards greenhouse gas emission in top Asian countries: Evidence from advance panel estimations," *Renew Energy*, vol. 186, pp. 207–216, 2022, doi: 10.1016/j.renene.2021.12.118.

[14] A. Mujtaba, "Symmetric and asymmetric impact of economic growth, capital formation, renewable and non-renewable energy consumption on environment in OECD countries," *Renewable and Sustainable Energy Reviews*, vol. 160, 2022, doi: 10.1016/j.rser.2022.112300.

[15] T. S. Adebayo, "Role of country risks and renewable energy consumption on environmental quality: Evidence from MINT countries," *J Environ Manage*, vol. 327, 2023, doi: 10.1016/j.jenvman.2022.116884.

[16] S. Gu, "Vector Quantized Diffusion Model for Text-to-Image Synthesis," *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 2022, pp. 10686–10696, 2022, doi: 10.1109/CVPR52688.2022.01043.

[17] Y. Ding, "Semi-Supervised Locality Preserving Dense Graph Neural Network with ARMA Filters and Context-Aware Learning for Hyperspectral Image Classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, 2022, doi: 10.1109/TGRS.2021.3100578.

[18] Q. Wang, "Underestimated impact of the COVID-19 on carbon emission reduction in developing countries – A novel assessment based on scenario analysis," *Environ Res*, vol. 204, 2022, doi: 10.1016/j.envres.2021.111990.

[19] H. Tao, "Groundwater level prediction using machine learning models: A comprehensive review," *Neurocomputing*, vol. 489, pp. 271–308, 2022, doi: 10.1016/j.neucom.2022.03.014.

[20] G. J. Qi, "Small Data Challenges in Big Data Era: A Survey of Recent Progress on Unsupervised and Semi-Supervised Methods," *IEEE Trans Pattern Anal Mach Intell*, vol. 44, no. 4, pp. 2168–2187, 2022, doi: 10.1109/TPAMI.2020.3031898.

[21] Z. Xu, "Using econometric and machine learning models to forecast crude oil prices: Insights from economic history," *Resources Policy*, vol. 83, 2023, doi: 10.1016/j.resourpol.2023.103614.

[22] M. Sikder, "The integrated impact of GDP growth, industrialization, energy use, and urbanization on CO<inf>2</inf> emissions in developing countries: Evidence from the panel ARDL approach," *Science of the Total Environment*, vol. 837, 2022, doi: 10.1016/j.scitotenv.2022.155795.

[23] B. Wu, "A social-ecological coupling model for evaluating the human-water relationship in basins within the Budyko framework," *J Hydrol (Amst)*, vol. 619, 2023, doi: 10.1016/j.jhydrol.2023.129361.

[24] S. Kumari, "Machine learning-based time series models for effective CO<inf>2</inf> emission prediction in India," *Environmental Science and Pollution Research*, vol. 30, no. 55, pp. 116601–116616, 2023, doi: 10.1007/s11356-022-21723-8.

[25] H. M. Rasel, "Sustainable futures in agricultural heritage: Geospatial exploration and predicting groundwater-level variations in Barind tract of Bangladesh," *Science of the Total Environment*, vol. 865, 2023, doi: 10.1016/j.scitotenv.2022.161297.

[26] E. G. Kim, "Designing solar power generation output forecasting methods using time series algorithms," *Electric Power Systems Research*, vol. 216, 2023, doi: 10.1016/j.epsr.2022.109073.

[27] G. Xiaomin, "How does urbanization affect energy carbon emissions under the background of carbon neutrality?," *J Environ Manage*, vol. 327, 2023, doi: 10.1016/j.jenvman.2022.116878.

[28] J. Wei, "Ultra-short-term forecasting of wind power based on multi-task learning and LSTM," *International Journal of Electrical Power and Energy Systems*, vol. 149, 2023, doi: 10.1016/j.ijepes.2023.109073.

[29] P. Bórawski, "Perspectives of photovoltaic energy market development in the european union," *Energy*, vol. 270, 2023, doi: 10.1016/j.energy.2023.126804.

[30] J. Kaur, "Autoregressive models in environmental forecasting time series: a theoretical and application review," *Environmental Science and Pollution Research*, vol. 30, no. 8, pp. 19617–19641, 2023, doi: 10.1007/s11356-023-25148-9.

[31] Bharti, "Short-term traffic flow prediction based on optimized deep learning neural network: PSO-Bi-LSTM," *Physica A: Statistical Mechanics and its Applications*, vol. 625, 2023, doi: 10.1016/j.physa.2023.129001.

[32] M. Ali, "Prediction of Complex Stock Market Data Using an Improved Hybrid EMD-LSTM Model," *Applied Sciences (Switzerland)*, vol. 13, no. 3, 2023, doi: 10.3390/app13031429.