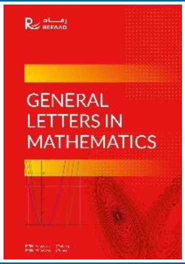
Available online at: www.refaad.com
Gen. Lett. Math., 14(1) (2024), 7-16
Research Article

General Letters in Mathematics (GLM)

Journal Homepage: https://www.refaad.com/Journal/Index/1

ISSN: 2519-9277 (Online) 2519-9269 (Print)



Determine the Best Models for Time Series by using a New Suggested Technique

Obaid Mahmmood Mohsin Alzawbaee [a*]

[a] Prof. Dr. Cihan University, Sulaymaniyah, Sulaymaniyah 46001, Kurdistan Region, Iraq

E-mails: [a] obed.muhsin@sulicihan.edu.krd

Abstract

The proposed method relies on a technique to choose the best model by giving values for the ranks of the model ARMA (p, q), where (p, q) are given the values 0, 1, 2. Every time (ACF) and (PACF) for the series of estimated errors {ât} are tested ،and the model which satisfies the two inequalities (7) and (9) is the best. It was applied to (14) annual series, representing the amounts of rainfall, and (7) daily series representing the daily Coronavirus infections. The most important conclusions and recommendations are. Applying the proposed method reduces the possibility of committing errors because its steps are clear and straightforward, giving it an application advantage. In many cases, it is impossible to know the model's rank through the behavior of (ACF) and (PACF), while the proposed method overcomes these forms to choose the best. The research recommended adopting the proposed method because it is easy to understand, learn and apply in addition to its accuracy.

Keywords: Time series analysis, ARMA (p, q) models, Rank model, Identification.

2020 MSC: 37M10.

1 Introduction

Time series analysis is one of the important statistical topics that deals with studying the behavior of phenomena and interpreting them over specific periods. The objectives of time series analysis can be summarized by obtaining an accurate description of the special features of the process from which the time series is generated, building a model to interpret the behavior of the time series, and using the results to forecast the behavior of the series in the future, as well as controlling the process from which the time series is generated by examining what might happen when some parameters change. To achieve this, the model requires a comprehensive analytical study of time series models, relying on statistical and mathematical methods. {[20], [21],[23], [16],[14]}

Our research presents a proposed method for determining the most efficient and best model from a set of possible models. It shortens some of the stages of time series analysis and gives the researcher a great deal of reassurance regarding the chosen model. A comparison with the usual method has proven the accuracy of its results.

* Corresponding author
Email addresses: obed.muhsin@sulicihan.edu.krd (Obaid Mahmmood Mohsin Alzawbaee).

doi: 10.31559/glm2024.14.1.2
Received: 20 Nov 2023 Revised: 21 Jan 2024 Accepted: 7 Feb 2024

Many researchers have addressed the issue of determining the appropriate model, and the most critical research and studies on the subject were:

In [14] offered unbiased Akaike Information Criterion (AICu) for model selection and claimed that it outperforms biased correction Akaike Information Criterion (AICc), but [17] claimed that AICc outperforms other forms of model selection tools. [1] proposed mean squared error of prediction as a criterion for selecting the model and claimed that it is better than using the residual sum of squares.

[18] proposed AIC and BIC, Predictive Least Squares (PLS), and Sequentially Normalized Least Squares (SNLS) for model selection. [11] proposed model selection method based on multiple hypothesis testing including AIC, BIC, FPE, and Minimum Description Length (MDL). [8], Only parameter consistency is required for model selection when testing predicted residuals; use final prediction Error (FPE) for model selection, which is the expected prediction error variance when Autoregressive is fitted. Use different methods for model selection, including hypothesis testing.

2 Time Series Analysis

Time series analysis consists of sequential stages, starting with the model's identification stage, then the stage of estimating the model parameters, then the stage of diagnostic checking the model's suitability, and finally, the forecasting stage. A time series consists of observations of a particular phenomenon during successive periods and within successive boundaries. They are of two types: Continuous or discrete, depending on what the values of (t) representing time take, and they can be stationary if the probabilistic characteristics are not affected by time or nonstationary. The time series model is the function that links the values of the time series to its previous values and errors.

One of the methods for analyzing time series is representing them with a general linear model, which is the mixed model, as many time series cannot be represented with an autoregressive model only or a moving average model only because the series often has the characteristics of both models. Thus, it is represented by a model. Mixed (autoregressive-moving average) and abbreviated (ARMA (p, q)), where (p) represents the rank of the autoregressive and (q) represents the rank of the moving average. {[6],[3],[19],[21],[9]}
The stationary {yt} time series is in the form ARMA(p, q)

$$\varnothing(B)\, y_t = \theta(B)\, a_t \tag{1}$$

Where:
$\varnothing(B) = 1 - \varnothing_1 B - \varnothing_2 B^2 - \ldots - \varnothing_p B^p$
$\theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \ldots - \theta_q B^q$
$B^j a_t = a_{t-j}$

$\{a_t\}$is a series of random errors called white noise, and the models (AR(p), (MA(q)) can be considered special cases of the model (ARMA (p, q)) from a mathematical standpoint. {[18],[13],[6],[5]}. The identification stage is important in time series analysis. It includes knowing the model and determining the rank for the specific model. Many procedures are applied to determine the type of model and its rank. After drawing the series, the nonstationary of the time series in its arithmetic mean (i.e., there is a general trend) is treated by taking the differences $(1-B)^d y_t$, where d>0, (and it often is, d=0,1,2). If the time series is nonstationary in its variance, it is treated using the appropriate transformation parameter for the series values (either the natural logarithmic transformation or the square root,). After the stationary of the series, we calculate and examine the autocorrelation function of the errors (ACF) and the partial autocorrelation function of the errors (PACF) for the sample to identify the model. There is a duality between the models ARMA (p,0) or AR(p) and the models ARMA (0, q) or MA(q), according to the two functions above. The problem becomes more complex in the case of mixed models ARMA (p, q) because relying on ACF PACF, to identify the model and determine its rank is not effective (especially in determining the rank) because the above functions in this case behave similarly, which is the behavior of gradual decrease {[11],[21]}.

Some models with different accuracy can be used in time series analysis to explain a set of given data. Choosing the best model takes work in many cases. Several criteria have been developed to compare models in choosing the model rank. The importance of choosing the model rank stems from the fact that choosing a rank lower than the actual rank of the model leads to inconsistency in the model parameters while choosing a rank higher than the actual rank of the model leads to an increase in the variance of the model. This leads to a loss of accuracy due to the increase in the number of parameters of the chosen model. The criteria for choosing the rank are based on the statistics of the residuals resulting from matching the unbiased model. There are several criteria for choosing the model's rank, one of which is proposed by the researcher {[3],[10],[22],[7]}.

2.1 Bayesian Criterion Information (BIC)

$$BIC= n\ln\hat{\sigma}_a^2 - (n-M)\ln(1-M/n)+M\ln(n)+M\ln[((\partial^2 y/\partial^2 a)-1)/M] \quad (2)$$

Where: P is the model rank, n is the number of observations, M is the number of parameters, $\partial^2 y$ is an estimate of variance of the series, $\hat{\sigma}_a^2 =$ The error variance estimated and calculated:

$$\hat{\sigma}_a^2 = \sum_{t=1}^{n}(y_t - \hat{y_t}) /(n-p) \quad (3)$$

The rank p corresponding to the lowest value of BIC is chosen.

2.2 Akaike Information Criterion

Its formula:

$$AIC(M) = n\ln\hat{\sigma}_a^2 + 2M \quad (4)$$

Or

$$AIC(p,q)= \ln\hat{\sigma}_a^2 + 2(p+q)/n \quad (5)$$

Where M=p+q, and p,q rank of the model, n number of observations, $\widehat{\sigma_a^2}$ Error variance estimator. The rank corresponding to the lowest value of AIC is chosen.

2.3 Suggestion Criterion (ASBC)

Its formula:

$$ASBC(p,q) = \ln\hat{\sigma}_a^2 + \frac{M}{2n}(2+\ln n) \quad (6)$$

Where M=p+q, n=number of observations, $\widehat{\sigma^2}$ Error variance estimator.
The rank corresponding to the lowest value of ASBC is chosen.
We note that choosing the correct model depends on the researcher's experience on the one hand and on testing more than one model to determine the best model, on the other hand, by relying on the criteria for choosing the rank. After knowing whether the type of model is AR(p), MA(q), or ARMA (p, q) and determining its rank, i.e., valuable knowledge p or q, or both, the process of estimating the model parameters is carried out using one of the estimation methods. {[6],[4]}. After estimating the parameters of the model that was identified, the examination process is carried out, and this is done through a set of tests, all of which agree with the goal, which is to test the autocorrelation function ACF for the series of estimated errors $\{\hat{a}_t\}$ and the partial autocorrelation function PACF for the series of estimated errors and find out whether they are not significant, which means that the model is efficient and can be adopted in forecasting. After the diagnostic checking stage, the model is used in forecasting, which is usually either a forecast of a value or a period. {[22],[21],[2]}

3 Anew Suggested Technique to Determine the Best Model

Arriving at a model that represents the carefully studied time series correctly and adequately means that the estimated error string is white-noise, and depending on that, the researcher believes that a new mechanism (technique) can be applied to access correct and efficient model relying on the fact that the estimated errors are white noise, going through the identification stage. This is because, in many cases, it is impossible to know the model's rank through the behavior of ACF and PACF in time series analysis. Especially in the case of mixed models ARMA, as shown in the figures(1a,1b,1c).

AR Process:

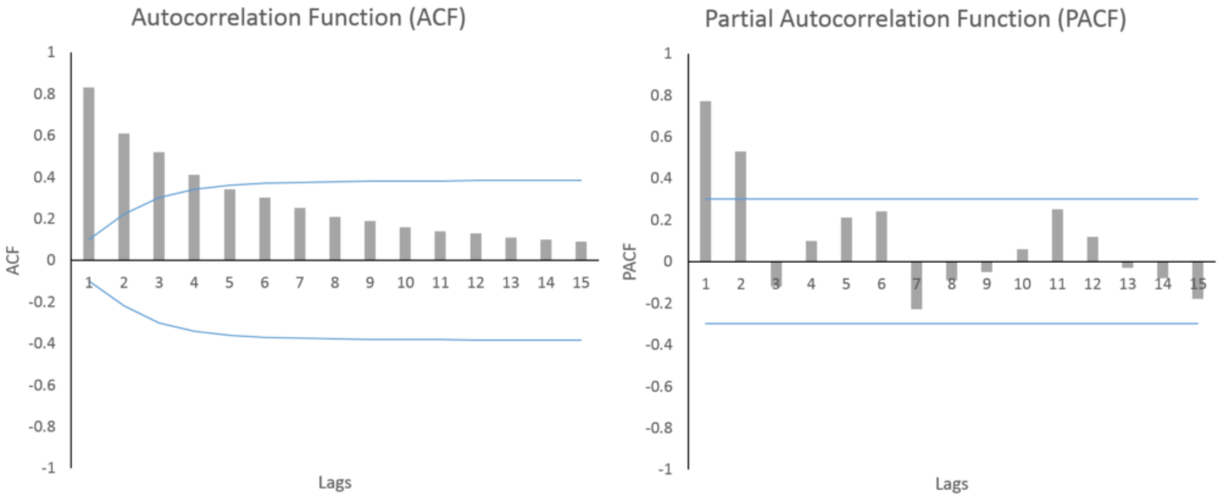


Figure 1a: behavior of ACF &PACF for autoregressive (AR (2)) models.

MA Process:

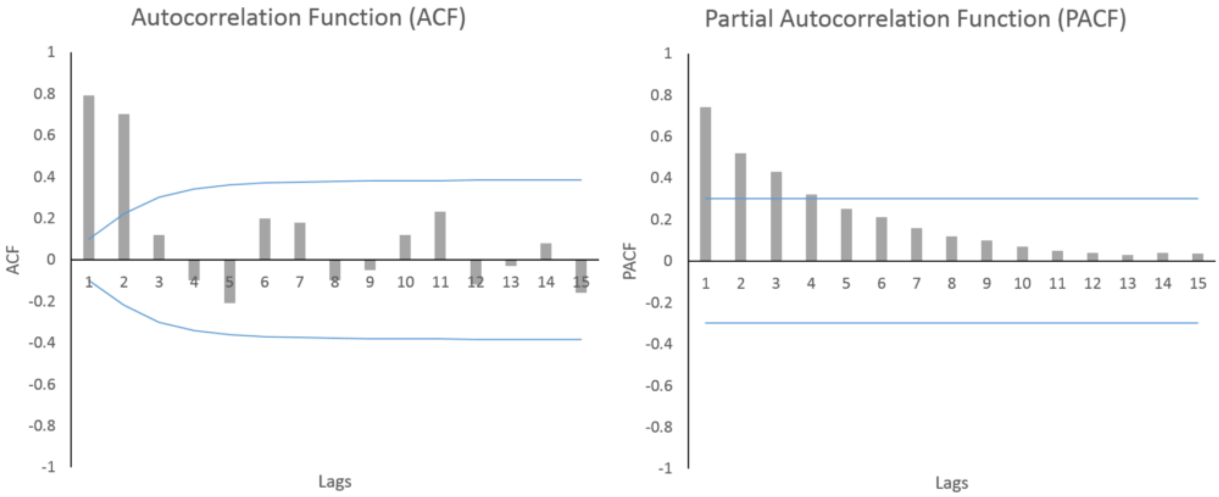


Figure 1b: behavior of ACF &PACF for Moving average (MA (2)) models.

ARMA Process:

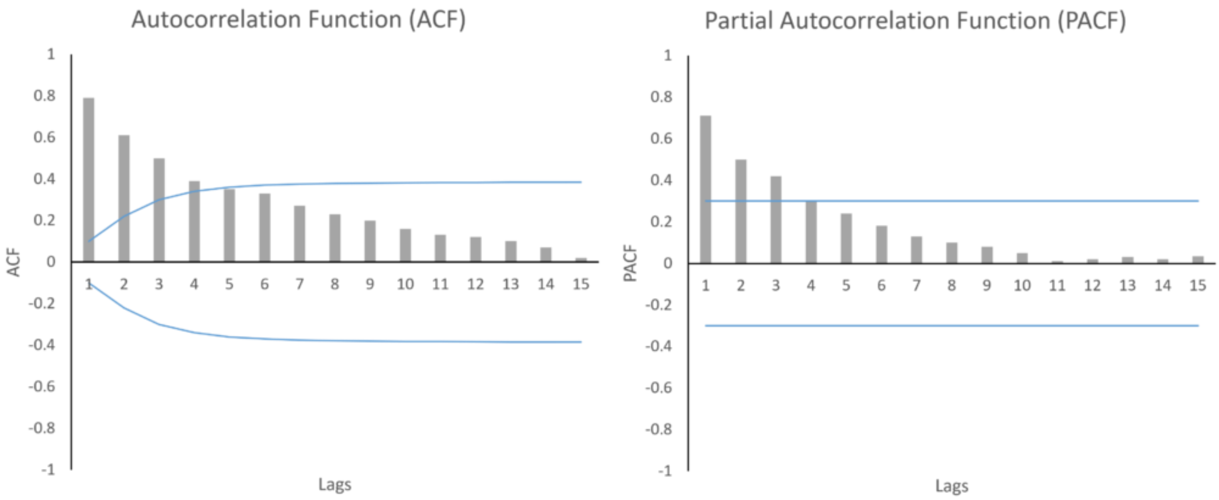


Figure 1c: behavior of ACF &PACF for mixed models (ARMA (p, q)).

Thus, the identification stage becomes useless; this proposed method also allowed for many alternatives, for example, choosing the best model by taking the possible arrangement. It is all for the model.
And compare the features that achieve the best model. This method is applied by conducting.
A test on the autocorrelation function for the estimated errors or the partial autocorrelation function of the estimated errors (or both). We will symbolize them with an abbreviation, respectively (ACFWN). (PACFWN), and by applying all possible arrangements of the
ARMA (p, q) model through bid q, p, in order of values 0,1,2. For the stationary time series $\{y_t\}$, we get Nn=32=9 models that can be ignored in the model
ARMA (0,0), so we have eight models as follows:
ARMA (p, q): (0,1) (0,2) (1,0) (1,1) (1,2) (2,0) (2,1) (2,2)
Each time we test one or both of the following relationships:

$$-2S[\rho_T(\hat{a}_t)] < \rho_T(\hat{a}_t) < 2S[\rho_T(\hat{a}_t)] \quad (7)$$

Where $S[\rho_T(\hat{a}_t)]$ = The standard deviation for ACF for estimated errors (which is calculated as follows):

$$S[\rho_T(\hat{a}_t)] = \sqrt{\frac{1}{n}[1 + 2\rho_1^2(\hat{a}_t) + \cdots + 2\rho_{T-1}^2(\hat{a}_t)]} \quad (8)$$

$\rho_T(\hat{a}_t)$ $\rho_T(\hat{a}_t)$ autocorrelation function (ACF) of the estimated error series $\{\hat{a}_t\}$
And it is (ACFWN)

$$\frac{-2}{\sqrt{n}} < \pi(\hat{a}_t) < \quad \rho_T(\hat{a}_t) \quad < \quad \frac{2}{\sqrt{n}} \quad (9)$$

Where π (a _t)` is PACF for, And it is (PACFWN)

A test is conducted each time (for each of the approved models). Inequality (7) or (9) above or both, and we stop when correct. So, the model that achieves the validity of the inequality is the correct and efficient model, and it is the best model among the models. A model that is used for forecasting purposes. What distinguishes this method is that it does not require much experience because it relies on taking all possible rankings for the mixed models, then choosing the best model based on the rank selection criteria as well as RMSE, and then testing (ACF) and (PACF) for the estimated errors series, through the inequality (7) and (9).

We did not address the theoretical proof of these inequalities because they (i.e., the two inequalities above) were taken from tests to diagnostic checking of the models (Bartlett test), which was used here for the same purpose, which is to test the estimated error correlations, but in another field or stage.[12]

4 Practical Aspect

4.1 Search data

The research was applied to two types of data: The first represents data for (14) annual time series, each of length (58) representing the period (1951-2008) for the amounts of rain falling on the regions of Derna (D) and Shahat (S) in Libya for the rainy months (January, February, March, October, November, December, and the entire year).[8]

The second represents data for (7) daily time series, each of which has a length of (76) representing the number of daily infections recorded for the Corona pandemic during the period from 1st March to 15th May of the year 2020 and for the countries (Iraq, Saudi Arabia, Emirates, Qatar, Kuwait, Oman, and Bahrain). {[6],[13]}.

4.2 statistical analysis and results:

The analysis was carried out using the usual and proposed methods for the series representing the first type of data, where the stationary of the time series was first confirmed. The nonstationary series were transformed into stationary series, and then the models were identified and their rank determined. Then, the best model was determined through the rank.

Selection criteria as well as (RMSE). After that, the parameters of the chosen model are estimated, and its accuracy is tested (diagnostic checking). Here, the model is determined in its final form and used in forecasting. Tables (1-a), (1-b), (2-a), and (2-b) represent the final results for the two usual and proposed methods for the first set of series (7series for Derna city) and the second set of series (7 series for Shahat city).

Table 1-a: The models that were tested for the first series Derna January (the usual method)

DJan	Model Fit statistics, Prediction Error Criterions						Ljung-Box Q (18)		
ARMA(p,q)	RMSE	MAE	BIC	AIC	ASBC	SBC	Statistics	DF	Sig.
ARMA (1,0)	38.934	28.64	7.46	425.750	7.358	427.662	23.778	17	0.126
ARMA (0,1)	38.931	28.63	7.46	425.742	7.358	427.654	23.748	17	0.126
ARMA (1,1)	37.665	28.42	7.47	423.543	7.338	427.367	20.550	16	0.196
ARMA (1,2)	37.830	28.67	7.55	425.992	7.398	431.728	20.939	15	0.139

Table 1-b: All possible models for the first series Derna-January (for a suggested method)

DJan	Model Fit statistics, Prediction Error Criterions						Ljung-Box Q (18)		
ARMA(p,q)	RMSE	MAE	BIC	AIC	ASBC	SBC	Statistics	DF	Sig.
ARMA (0,1)	38.931	28.63	7.46	425.742	7.358	427.654	23.748	17	0.126
ARMA (0,2)	39.081	28.45	7.54	428.167	7.418	431.991	22.972	16	0.114
ARMA (1,0)	38.934	28.64	7.46	425.750	7.358	427.662	23.778	17	0.126
ARMA (1,1)	37.665	28.42	7.47	423.543	7.338	427.367	20.550	16	0.196
ARMA (1,2)	37.830	28.67	7.55	425.992	7.398	431.728	20.939	15	0.139
ARMA (2,0)	39.086	28.44	7.54	428.184	7.418	432.008	23.305	16	0.106
ARMA (2,1)	37.804	28.67	7.55	425.915	7.397	431.651	20.662	15	0.148
ARMA (2,2)	38.605	27.91	7.66	430.708	7.497	438.356	22.982	14	0.061

It is clear that the ARMA (1 ,1) model is the best and most efficient, and its parameters are:

$$\hat{y}_t = 60.432 + 0.776y_{t-1} + 0.998a_{t-1} \quad (10)$$

Table (2-a): The model chosen for the six series (the usual method), (Derna)

	Model Fit statistics, Prediction Error Criterions						Ljung-Box Q (18)		
ARMA(p,q)	RMSE	MAE	BIC	AIC	ASBC	SBC	Statistics	DF	Sig.
DFeb_LNd1 ARIMA (0,1,1)	28.343	21.075	6.831	-10.152	-0.160	-12.365	20.978	17	0.227
DMar ARMA (1,1)	21.715	14.587	6.366	360.001	6.242	363.825	10.877	16	0.817
DOct ARMA (1,0)	44.482	31.989	7.73	441.204	7.625	443.117	7.98	17	0.967
DNov ARMA (1,0)	28.828	20.547	6.863	390.89	6.757	392.802	11.437	17	0.833
DDec_LN ARMA (2,1)	0.872	25.808	0.651	-10.9816	-0.136	-17.86	13.17	15	0.59
DTotal ARMA (2,1)	81.231	63.425	9.075	515.022	8.933	520.758	19.01	15	0.213

Table 2-b: The model chosen for the six series (the suggested method), (Derna)

	Model Fit statistics, Prediction Error Criterions						Ljung-Box Q (18)		
ARMA(p,q)	RMSE	MAE	BIC	AIC	ASBC	SBC	Statistics	DF	Sig.
DFeb_LNd1 ARIMA (0,1,1)	28.343	21.075	6.831	-10.152	-0.160	-12.365	20.978	17	0.227
DMar ARMA (1,1)	21.715	14.587	6.366	360.001	6.242	363.825	10.877	16	0.817
DOct ARMA (1,0)	44.482	31.989	7.73	441.204	7.625	443.117	7.98	17	0.967
DNov ARMA (1,0)	28.828	20.547	6.863	390.89	6.757	392.802	11.437	17	0.833
DDec_LN ARMA (2,1)	0.872	25.808	0.651	-10.9816	-0.136	-17.86	13.17	15	0.59
DTotal ARMA (2,1)	81.231	63.425	9.075	515.022	8.933	520.758	19.01	15	0.213

Table 3-a: The model chosen for the eighth series Shahat- January (the usual method)

Shahat Jan. ARMA(p,q)	Model Fit statistics Prediction Error Criterions						Ljung-Box Q(18)		
	RMSE	MAE	BIC	AIC	ASBC	SBC	Statistics	DF	Sig.
ARMA (1,0)	60.811	47.987	8.356	477.475	8.250	479.387	13.155	17	0.726
ARMA (0,1)	60.666	47.898	8.351	477.197	8.245	479.109	12.458	17	0.772
ARMA (2,1)	60.464	45.869	8.484	480.774	8.342	486.510	11.605	15	0.709
ARMA (2,2)	60.58	45.301	8.558	482.975	8.398	490.623	9.369	14	0.807

Table 3-b: All possible models for the eighth series Shahat- January (suggested Method)

Shahat Jan. ARMA(p,q)	Model Fit statistics Prediction Error Criterions						Ljung-Box Q(18)		
	RMSE	MAE	BIC	AIC	ASBC	SBC	Statistics	DF	Sig.
ARMA (0,1)	60.666	47.898	8.351	477.197	8.245	479.109	12.458	17	0.772
ARMA (0,2)	60.194	46.148	8.405	478.273	8.282	482.097	12.156	16	0.733
ARMA (1,0)	60.811	47.987	8.356	477.475	8.250	479.387	13.155	17	0.726
ARMA (1,1)	60.425	46.117	8.413	478.520	8.286	482.344	13.546	16	0.633
ARMA (1,2)	60.485	45.918	8.485	480.814	8.343	486.550	11.534	15	0.714
ARMA (2,0)	60.822	47.298	8.426	479.477	8.302	483.302	13.307	16	0.65
ARMA (2,1)	60.464	45.869	8.484	480.774	8.342	486.510	11.605	15	0.709
ARMA (2,2)	60.58	45.301	8.558	482.975	8.398	490.623	9.369	14	0.807

The ARMA (0,1) model is the best and most efficient, and its parameters are.

$$\hat{y}_t = 122.398 + 0.153\, a_{t-1} \quad (11)$$

Table 4-a: The model chosen for the six series (the usual method), (Shahat)

	Model Fit statistics, Prediction Error Criterions						Ljung-Box Q (18)		
ARMA(p,q)	RMSE	MAE	BIC	AIC	ASBC	SBC	Statistics	DF	Sig.
S Feb-d1 ARMA (0,1,1)	51.094	38.325	8.009	448.63	7.8885	454.46	13.504	17	0.7
SMar_Inv ARMA (1,2)	0.039	0.024	-6.220-	-392.6	-6.715	-419.2	10.464	15	0.79
SOct ARMA (0,1)	60.979	41.404	8.361	477.794	8.256	479.706	5.856	17	0.994
SNov_LN ARMA (1,0)	0.811	0.634	-.279-	-23.34	-0.385	-25.78	14.695	17	0.62
SDec_Inv ARMA (1,0)	0.046	0.023	-6.018-	-358.4	-6.162	-366.7	5.792	17	0.99
STotal ARMA (0,1)	131.027	104.009	9.891	566.519	9.785	568.432	8.249	17	0.961

Table 4-b: The model chosen for the six series (the suggested method), (Shahat)

	Model Fit statistics, Prediction Error Criterions						Ljung-Box Q (18)		
ARMA(p,q)	RMSE	MAE	BIC	AIC	ASBC	SBC	Statistics	DF	Sig.
SFeb_d1 ARMA (0,1,1)	51.094	38.325	8.009	448.63	7.8885	454.46	13.504	17	0.7
SMar_Inv ARMA (1,2)	0.039	0.024	-6.220-	-392.6	-6.715	-419.2	10.464	15	0.79
SOct ARMA (0,1)	60.979	41.404	8.361	477.794	8.256	479.706	5.856	17	0.994
SNov_LN ARMA (1,0)	0.811	0.634	-.279-	-23.34	-0.385	-25.78	14.695	17	0.62
SDec_Inv ARMA (1,0)	0.046	0.023	-6.018-	-358.4	-6.162	-366.7	5.792	17	0.99
STotal ARMA (0,1)	131.027	104.009	9.891	566.519	9.785	568.432	8.249	17	0.961

As for the second type of data, the proposed method was applied only after ensuring the stationary of each series and determining the models by adopting the proposed method. The model parameters were estimated, their accuracy was tested (diagnostic checking), and they were relied upon in forecasting. Its accuracy was proven, and the following tables showed the results.

Table 5: All possible models ARIMA (p, 1, q) for OMAN

ARIMA (p, 1, q)	RMSE	AIC	BIC	MAE
ARIMA (1,1,0)	31.929	521.527	7.042	19.219
ARIMA (2,1,0)	31.767	522.764	7.090	18.892
ARIMA (0,1,1)	32.277	523.1533	7.064	19.600
ARIMA (0,1,2)	31.803	522.934	7.092	18.930
ARIMA (1,1,1)	31.848	523.146	7.095	18.864
ARIMA (1,1,2)	32.022	525.963	7.163	18911
ARIMA (2,1,2)	29.743	516.889	7.073	17.892
ARIMA (2,1,1)	29.429	513.297	6.994	17.874

From the results of the above table, it was found that the model ARIMA (2, 1, 1) is the best among the other models, because it corresponds to the minimum values of (AIC, BIC, MAE, RMSE). And (Table 6) shows the result summary for the seven countries.

Table 6: The Optimal Parameters for ARIMA Models for all countries

Country	ARIMA (p, d, q)	RMSE	AIC	BIC	MAE
IRAQ	ARIMA (1, 1, 1)	13.145	390.406	5.325	8.85
QATAR	ARIMA (1, 1, 0)	122.287	722.956	9.728	80.098
AUE	ARIMA (0, 1, 1)	73.873	647.352	8.720	40.499
SAUDE	ARIMA (1, 1,0)	080.85	660.889	8.900	54.351
BAHRAYN	ARIMA (0, 1, 1)	52.384	595.790	8.032	35.710
KWAIT	ARIMA (2, 1, 2)	73.042	651.655	8.870	44.322
OMAN	ARIMA (2, 1, 1)	29.429	513.297	6.994	17.874

From the above table, and after determining the quality and rank of the models for each country, the model parameters for each of the above countries were estimated. (for example, the final model for Iraq is):

ARIMA (1, 1, 1) $$Z_t = 0.911 + 0.734Z_{t-1} + 0.998\alpha_{t-1} \tag{12}$$

Note that some criteria were not mentioned, and the presentation in some tables was limited to the most important ones for brevity, noting that their results have been verified.

After we determined the model for each country and estimated its parameters, in the step of diagnostic checking, the Ljug-Box test for each country's model was used. Which depends on the test of significance Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) for the estimated residuals. Its formula is:

$$Q = n(n+2)\sum_{k=1}^{h}\frac{\hat{\rho}_k^2}{n-k} \sim \chi^2(m-p-q) \tag{13}$$

And (Table 7) shows the results of Ljug-Box test for each model.

Table 7: Results of Ljug-Box test

Model	Ljung-Box Q (18)		
	Statistics	DF	Sig.
IRAQ-Model_1	18.801	16	.279
QTR-Model_2	12.033	17	.798
UAE-Model_3	6.844	17	.985
SUD-Model_4	17.476	17	.423
BHR-Model_5	13.903	17	.674
KUW-Model_6	21.081	14	.100
OMN-Model_7	17.031	15	.317

The results of the above table show that all the models are efficient.

5 Conclusions and Recommendations

Through a theoretical review of the proposed method and through the practical application in the usual way the proposed method connects, there are several conclusions and recommendations. The following are the most important:

- Applying the proposed method reduces the possibility of making mistakes because its steps are clear and straightforward compared to the usual method, which gives it an advantage in the application.
- In many cases of the time series analysis stages, it is impossible to know the model's rank through the behavior of ACF and PACF. Hence, the identification stage becomes less critical. The proposed method goes beyond these forms by taking all models through the possible arrangements of p and q.
- The proposed method allows all possible alternatives to be given a model to choose the most efficient and best model by applying all possible arrangements of the model and comparison of the features that if the best model is achieved, the model will likely be in the usual way not the best because not all possible models are taken into account.
- The proposed method reduces the effort and time to connect to the correct model. It is efficient, easy, and has a technique that is easy to trade and learn.
- The researcher recommends applying the proposed method, especially in the field of learning, because it is easy to understand and apply.

Reference:

[1] D. M. Allen. Mean Square Error of Prediction as a Criterion for Selecting Variables. *Journal of Technimetrics*, 13 (3), (2000), 71- 91. https://doi.org/10.2307/1267161

[2] O. M. Alzawbaee & H. M. Nozad. Building An Efficient Model by Using Panel Data. *journal of Humanities and Social Sciences (CUEJHSS)- Cihan University-Erbil*, 1 (2023), 82-87. https://doi.org/10.24086/cuejhss.v7n1y2023.pp82-87

[3] O. M. Alzawbaee, et al. Using Time Series Analysis to Forecast the Prevalence of the COVID-19 Pandemic (Iraq and GULF Cooperation Council Countries (GCC)). *Journal of Critical Reviews*, 7 (15), (2020).

[4] O. M. Alzawbaee. *Identification And Diagnostic Checking for Low Order Mixed Models*. PhD thesis- University of Baghdad- College of Administration and Economics, (2005).

[5] O. M. Alzawbaee. *Building a transfer function model to predict the numbers of students of the scientific and literary branches in Iraq*. The Second International Scientific Conference - College of Administrative and Financial Sciences - Cihan University - Sulaymaniyah - 22-23 /8/2021.

[6] J. Cao, X. Jiang, & B. Zhao, Mathematical modeling and epidemic prediction of COVID-19 and its significance to epidemic prevention and control measures. *Journal of Biomedical Research & Innovation*, 1, (2020) ,1-19.

[7] D. Nassiuma, Non-stationary Autoregressive Moving Average Processes with Infinite Variance. *Journal of time series Analysis*, 14 (3), (1993), 297-304. https://doi.org/10.1111/j.1467-9892.1993.tb00146.x

[8] G. E. DufourJ & A. Hall. Generalized Predictive Tests and Structural Change Analysis in Econometrics. *Journal of International Economic Review*, 35 (1), (1994), 199-229. https://doi.org/10.2307/2527098

[9] G. Elliott, T. Rothenberg & J. Stock. Efficient Tests for and Autoregressive Unit root. *Journal of Econometrics*, 64 (4), (1996), 813-836. https://doi.org/10.2307/2171846

[10] M. Elnazeer & O. M. Mohsin. *Using Multivariate Time Series Analysis in Constructing a Statistical Model to Forecast Traffic Accidents in Sudan*. The 7th International Scientific Conference of Arab Statisticians-(2023), Cairo, Egypt 8-9/3/2023.

[11] C. Gatean, Subset ARMA Model Identification using genetic Algorithms. *Journal of time series Analysis*, 21(5), (2000), 559-570. https://doi.org/10.1111/1467-9892.00198

[12] A. A. Hamad. *Using Time Series Analysis with Both Time and Frequency Domain to Forecast the Amounts of Rain Falling in the Regions of Derna and Shahat*. PHD thesis - Sudan University of Science and Technology (2011).

[13] H. Zhirui &T. Hongbing, Epidemiology and ARIMA model of positive rate of influenza viruses among children in Wuhan, China: A nine-year retrospective study. *International Journal of Infectious Diseases*, 74, September 2018, 61-70. https://doi.org/10.1016/j.ijid.2018.07.003

[14] H. Leeb & B. M. Pötscher. Model Selection and Inference: Facts and Fiction. *Journal of econometric*, 21, (2005), 21-59. https://doi.org/10.1017/s0266466605050036

[15] L. Lon-Mu. *Time Series Analysis and Econometrics*. Scientific Computing Associates Crop. 2nd ed, amazon (2009).

[16] H. Lukepoli. *New Introduction to Multiple Time Series Analysis*., Springer Berlin Heidelberg, New York. (2005).

[17] J. Maatta. *Model Selection Methods for Linear Regression and Phylogenetic Reconstruction*. PhD Thesis, Series of Publications A, Report A-2016-3, (2016).

[18] A. Mc. Quarrie, R. Shumway, & C-L. Tsai. The model selection criterion AIC. *Journal of Statistics & Probability Letters*, 34, (1998).

[19] H. Robert. *Shumway & David. S, Stoffer, "Time Series Analysis and Its Applications: With R Examples -Springer Texts in Statistics* (STS), 2017.

[20] S.O. Ongbali, et al. Model Selection Process in Time Series Analysis of System with Random Output. ICESW IOP Publishing. *IOP Conf. Series: Materials Science and Engineering*, 413 (2018) 012057. https://doi.org/10.1088/1757-899x/413/1/012057

[21] W.S. Wei, William. *Time Series Analysis Univariate and Multivariate Methods*. Wei-2nd Addison, Wesley- USA (2006).

[22] T. Wu. Nowcasting and forecasting the potential domestic and international spread of the COVID-19 outbreak originating in Wuhan, China: a modelling study, *PMC. Journal List-Elsevier-PMC, COVID-19, PMC*,159271(2020), pp 689-697.