Al-Qadisiyah Journal for Administrative and Economic Sciences ISSN$_{Online}$: 2312-9883
QJAE, Volume 28, Issue 2 (2026) ISSN$_{Print}$: 1816-9171
DOI:/ /doi.org/10.33916/qjae.2026.02339355

Using some mathematical models to forecast the numbers of first grade primary students in Basra Governorate

zainab sami Yaseen(1)
Southern technical university / medical technical institute of Basrah
zainab.s.yassin@stu.edu.iq

Sanaa Mohammed Naeem(2)
Southern technical university/college of health& medical techniqnes in basrah
Sunamohammed70@gmail.com

Shaymaa Qasim Mohsin(3)
University of Basra/college of Administration
Southern technical university / technical institute of Basrah
Shaymaa.qassim79@gmail.com

Received: 01/03/2026
Accepted: 15/05/2026
Available online: 30/06/2026

Corresponding Author: Zainab Sami Yaseen

Abstract: The research dealt with one of the most important topics in the statistical and educational aspect, which is predicting the numbers of first-grade students, as it is important for educational, political and urban institutions, using different predictive models (Box-Jenkins models of the first order and exponential models) with the aim of conducting the forecasting process for those numbers and comparing these models using the mean absolute percentile error (MAPE) standard and selecting the best model in light of these indicators that differ from each other.tire study without repeating the introduction or including unnecessary background information.

Keywords: Forecasting , Box-Jenkins models , Autoregressive, Moving average , ARMA **.**

1. Introduction:

The introduction should clearly present the research context and define the main problem addressed in the study. It must begin with a concise background that explains the importance of the topic within the field of administrative and economic sciences. The section should identify the research gap by highlighting limitations or inconsistencies in previous studies, leading logically to the formulation of the research problem. The introduction must also state the objectives of the study in a precise and focused manner. Where appropriate, it should briefly indicate the methodological approach and the structure of the paper. All in-text citations must follow the APA style using the author–date format, presented within parentheses as (Author, Year), such as (Smith, 2020) or (Smith & Brown, 2021), and in cases of three or more authors as (Smith et al., 2022). Direct quotations must include page numbers in the format (Author, Year, p. xx). All cited sources must appear in the reference list, and no uncited references should be included. The text should be written in clear academic English, with a logical flow of ideas, avoiding unnecessary details, excessive citations, or repetition of content presented in other sections.

Autocorrelation Function (ACF): The fundamental statistic in time series analysis is the autocorrelation coefficient. The term autocorrelation can be explained as representing the relationship between successive observations of the same variable over a period of time. The essence of autocorrelation is that a random variable occurring during a specific time period is related to the random variable that precedes or follows it; that is, the series is related to itself or shifted by [1, 2, ...] periods.

The general formula for calculating the autocorrelation of a stationary series is:

$$\rho_y(k) = \frac{\sum_{t=1}^{n-k}(y_t - \bar{y})(y_{t+k} - \bar{y})}{\sum_{t=1}^{n}(y_t - \bar{y})^2}$$

$$= \frac{E(y_t - \mu_y)(y_{t+k} - \mu_y)}{\sigma_y^2}$$

$$= \frac{\gamma_{y(k)}}{\sigma_y^2} = \frac{\gamma_y(k)}{\gamma_y(0)}$$

Partial Autocorrelation Function (PACF): This function is used to measure the degree of correlation between yt and yt-k when the influence of other values is constant, i.e., the displacement values of y, and is calculated by regression with (yt-1,..., yt-p), i.e.

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \ldots + \phi_p y_{t-p} + a_t$$

Box-Jenkins models: When studying the Box-Jenkins models, we will take the non-seasonal models, as these models are divided into two sections, which are the stationary models (Stationary Models) and the non-stationary models (Nonstationary Models).

What is meant by stable models is that they are suitable for representing time series, and time series are a set of observations of a specific phenomenon measured in time periods (t= 0, 1, 2, ….) If the observations of this set are continuous over time, the time series is called continuous, and if it is discontinuous, i.e. taken at specific and limited times, it is called discrete. As for the non-stationary series, the differences must be taken to make them stable, and then they are suitable for representing time series.

Observations in time series are expressed as Xt1, Xt2,…, Xtn at time intervals t1, t2,…, tn where n is the number of consecutive values. As we mentioned, stable models are those that are suitable for representing time series that have the property of stability before taking any number of differences for them, meaning that they are originally stable.

Therefore, time series are divided into:

1- Stationary Time Series

2- Non-Stationary Time Series.

The time series is stable when the observations fluctuate around a fixed mean and has a constant variance, meaning the behavior of the random variables does not change over time and does not take on other characteristics. Time series are unstable if the series does not have a fixed mean. These models include four types:

1) The Autoregressive Models.
2) The Moving average Models.
3) The Mixed Autoregressive-Moving average Model.
4) The Mixed Autoregressive-Identify-Moving average Models.

Autoregressive Model: It is the autoregressive of order (P) and the model is written as (P) MA as follows:

$$X_t = \mu^{'} + \varphi_1 X_{t-1} + \varphi_2 X_{t-2} + \ldots + \varphi_p X_{t-p} + a_t \quad (1)$$

Where:

$\mu^{'}$: represents the constant term

φ: represents the autoregressive parameter of degree (P)

at: represents the error for the period (t).

When (p = 1) we obtain a first-order autoregressive model (MA (1)), so the model is as follows:

$$X_t = \mu^{'} + \varphi_1 X_{t-1} + a_t \qquad (2)$$

Using the backscatter factor, equation (2) is as follows:

$$\left.\begin{array}{c} X_t = \mu^{'} + \phi_1 X_{t-1} = \mu^{'} + a_t \\ (1 - \varphi B) X_t = \mu^{'} + a_t \end{array}\right] \qquad (3)$$

The moving average model is of order (q) and the model (q) AR is written as follows:

$$X_t = \mu + a_{t - \theta_1 a_{t-1} - \theta_2 a_{t-2 - \cdots - \theta q} a_{t-q}} \quad (4)$$

Where

θ_1, θ_2 : represents the parameters of the moving average.

at-q: represents the error term for the period (t-q) μ: represents the constant.

When (q=1) we get the first-order moving average model (AR(1)), the model is as follows:

$$X_t = \mu + a_{t-\theta_1 a_{t-1}} \qquad (5)$$

Using the backscatter factor, equation (5) is as follows:

$$X_t = \mu +_{(1-\theta_1\beta)a_t} \qquad (6)$$

Where β: represents the backscatter factor.

Mixed model (autoregressive - moving average): The general model of autoregressive moving averages is written as (ARMA(p,q)) as follows:

$$X_t = \mu' + \varphi_1 X_{t-1} + \varphi_2 X_{t-2} + \ldots + \varphi_p X_{t-p+} a_{t-\theta_1 a_t - \theta_2 a_{t-1} - \theta_q a_{t-q}} \qquad (7)$$

When the rank is (P=1) and (q=1) we get an (autoregressive moving average) model (ARMA (1.1) so the model is as follows:

$$\left.\begin{aligned} X_t &= \mu' + \varphi_1 X_{t-1} + a_{\theta_2 a_{t-1}} \\ (1-\varphi_1\beta)X_t &= \mu' + (1-\varphi_1\beta)t - \theta_1 a_t \end{aligned}\right\} \qquad (8)$$

Autoregressive – Integrate-Moving Average: In this case, the problem of Non stationary will be added to the mixed (ARMA) process, and the general model becomes (p, d, q) ARIMA. When each of (P=1), (d=1) and (q=1), we get the simple state equation (1,1,1) ARIMA, which is as follows:

$$(1-\beta)(1-\varphi\beta)X_t = \mu' + (1-\theta_1\beta)at \qquad (9)$$

Stationary and Non stationary: There is an important symbolic tool for the back-displacement process, symbolized by (β), which is also the back bounce factor and is used as follows:

$$\beta X_t = X_{t-1} \qquad (10)$$

The operation (β) on (Xt) means shifting the data backwards for one period, and the application of (β) on (Xt) means shifting the data backwards for two periods, as follows:

$$B(BX_t) = B^2 X_t = X_{t-2} \qquad (11)$$

The back-shift factor (B) is suitable for describing the process of differences. For example, if the time series is not stationary, it can be made stationary by taking the first differences of the series, as follows:

- **First Difference:** $X_t' = X_t X_{t-1}$ (12)

  The above equation can be rewritten using (B) the back-displacement factor:

  $$X_t' = X_t - BX_t = (1 - B)X_t \qquad (13)$$

  We note that the first differences are represented by (1-B), but if we want to calculate the second-order differences (which are the differences for the first differences), they will be as follows:

- **Second – order Difference:** $X_t'' = {X_t}' - {X_{t-1}}' = (1 - \beta)^2 X_t$ (14)

  Where $(1 - \beta)^2$ means that the differences are of the second degree.

  It is necessary to distinguish between second-order differences and secondary differences, where the first is represented by$(1 - \beta)^2$ and the second by $(1 - B^2)$.

  The second difference:

  $$X_t - X_{t-2} = X_t - B^2 X_t = (1 - B^2)X_t \quad (15)$$

  The purpose of taking the differences from degree (d) to obtain stability is that we will represent it as follows:

  $$(1 - \beta)^2 X_t = \alpha_t \qquad (16)$$

Model Estimation: After diagnosing the model and determining its score, its parameters are estimated. There are several methods for estimation, including:

- **Yule-Walker equations:** This method relies on the autocorrelation function of the model. To estimate the parameters of an AR model, its autocorrelation function takes the following form.

  $\rho_j = \phi_1 \rho_{j-1} + \phi_2 \rho_{j-2} + \ldots + \phi_p \rho_{j-p}$

  To obtain Yule-Walker estimates for the parameters ρ_j, we substitute the autocorrelation coefficient ρ_j with the estimated autocorrelation coefficient r_j, as follows:

$$\phi_j = R^{-1} r_j \qquad R = \begin{bmatrix} 1 & r_1 & \cdots & r_{p-1} \\ \vdots & \ddots & & \\ r_{p-1} & & & 1 \end{bmatrix}, \quad r_j = \begin{bmatrix} r_1 \\ \vdots \\ r_p \end{bmatrix}$$

- **least squares method:** In this method, we obtain the sum of squared errors as little as possible $\sum \hat{a}_t^2$, and by partial differentiating the parameters and setting them equal to zero, we obtain the parameter estimates.
- **Maximum likelihood method**: To estimate the parameters of the ARMA model using the maximum likelihood method, maximizing the function minimizes the sum of squared errors, and then partially differentiating the parameters yields the estimates.

Diagnostic Check: The model's accuracy and efficiency in representing the time series must be verified before use, and this is done by:

- **Using autocorrelation coefficients for estimated errors (residues) where:**

$$r_k(\hat{a}) = \frac{\sum_{t=1}^{n-k} \hat{a}_t \hat{a}_{t+k}}{\sum_{t=1}^{n} \hat{a}_t^2}$$

Since $r_k(\hat{a})$ is normally distributed with a mean of zero and a variance $\frac{1}{n}$, then:

$$\phi = n \sum_{k=1}^{m} r_k^2(\hat{a})$$

Where n is the number of observations, d is the number of calculated shifts when the autocorrelation coefficient of errors is given, where φ is approximately distributed in a χ^2 distribution with degrees of freedom (n-p-q). If the calculated value is less than the tabulated value, this means that the model is good and suitable.

- By plotting the estimated autocorrelation coefficients of the residuals, if they fall within the 95% confidence interval, then the model is suitable, i.e., that

$$-1.96 \frac{1}{\sqrt{n}} \leq r_k(a) \leq +1.96 \frac{1}{\sqrt{n}}$$

Forecasting methods:

Exponential Smoothing Method: The exponential smoothing method is one of the most common methods for smoothing discrete time series and forecasting for the near future. due to the simplicity and efficiency of this method and its suitability to the changes that occur in the time series to be forecast, this method has become more widely used in various statistical fields, and therefore the method can be divided into the following:

- **Exponential smoothing method for fixed processes:** This method is used when the time series does not change with time or may fluctuate slightly. In this case, the model for the stationary process can be formulated according to the following equation: $X_t = b + a_{t...}$ $t = 1,2,3, \ldots$ (17)

$$a_t \sim N(0, \sigma_a^2)$$

$$\hat{X}_t(T) = \hat{b}(t) \quad (18)$$

Where $\hat{b}(t)$ represents the expected value of the average of the time series and the model for the stationary process can be obtained by applying the weighted least squares method, which can be represented by the following equation:

$$\hat{b}(t) = S_t(X) = \alpha X_t + (1 - \alpha) S_{t-1}(X) \quad (19)$$

The process of choosing the value of the smoothing constant (α) is of great importance in determining the practical characteristics of exponential smoothing. Sources differed in determining the value of this constant theoretically and practically. The scientist Montgomery showed that this value can be equal to:

$$\alpha = \frac{2}{N + 1} \quad (20)$$

It also depends on the nature of the data. By applying the principle of trial and error, it was found that the value of the constant (α) is limited to the range $(0.01 \leq \alpha \leq 0.30)$. Therefore, the ideal determination of this value (α) is by giving different values to it and finding the mean square errors (MSE) at each chosen value of the constant (α), and the ideal value then is the one that corresponds to the lowest (MSE).

- **Exponential smoothing method for processes with a linear direction:** The idea of exponential smoothing can be extended to other situations when the process changes with time. Among these cases is the case where the process includes a linear change at certain time periods according to the following equation:

$$X_t = b_1 + b_2 + \alpha_t \quad (21)$$

$$\hat{X}_{t+L} = \hat{b}_1(t) + L\hat{b}_2(t) \quad (22)$$

- **Exponential smoothing method for higher order operations:** In general, exponential smoothing can be used to estimate the parameters of polynomial models to any degree. to clarify the idea of this method, we assume the second-order model, which is formulated as follows:

$$X_t = b_1 + b_2 t + ½\, b_3 t^2 + a_t \qquad (23)$$

The forecast equation for a future period of time (t + L) is as follows:

$$X'_{t+L} = \hat{b}_1(t) + L\hat{b}_2(t) + 1/2L^2\hat{b}_3(t) \qquad (24)$$

The estimation of the model parameters in Equation (24) is according to the equations below:

$$\left.\begin{aligned} \hat{b}1(t) &= 3S'_t(X) - 3S''_t(X) + S'''_t(X) \\ \hat{b}2(t) &= \frac{\alpha}{2(1-\alpha)^2}\left[(6-5\alpha)S'_t(X) - 2(5-4\alpha)S''_t(X) + (4-3\alpha)S'''_t(X)\right] \\ \hat{b}3(t) &= \left(\frac{\alpha}{1-\alpha}\right)^2\left[S'_t(X) + S''_t(X) + S'''_t(X)\right] \end{aligned}\right] \qquad (25)$$

Model Evaluation Measures: Model evaluation measures are fundamental tools in statistical modeling, used to assess the quality of a model's performance and its efficiency in interpreting and accurately representing data. These measures also help verify a model's predictive and generalizable capabilities when applied to new, unused data. evaluation measures vary depending on the nature of the model and the aspect of performance being measured, whether it relates to prediction accuracy, fit quality, or error magnitude. The measures adopted in this study are as follows:

- **Akaike Information Criterion (AIC):** This criterion is used to balance the quality of fit of the model with its complexity, as it penalizes models that contain a larger number of parameters, with the aim of selecting the most efficient model that best explains the data with the least amount of complexity.

$$\mathrm{AIC} = \ln(\sigma^2) + \frac{2\,K}{T} \qquad (26)$$

- **Bayesian Information Criterion (BIC):** The Bayesian Information Criterion (AIC) is similar to the Akaiki Criterion (AIC), except that it imposes a greater penalty for model complexity and an increased number of parameters, making it more stringent when comparing models and selecting the most suitable one.

$$\mathrm{BIC} = \ln(\sigma 2) + \frac{K\ln(T)}{T} \qquad (27)$$

Use of conventional and unconventional models for forecasting:

To analyze and forecasting the number of first-grade students in Basra Governorate, a sample size of (37) years was used from 1987 to 2024 from the Basra Education Directorate.

- **Forecast model: AR (1):** The model (ARIMA(0,0,1)) was chosen from among the selected models to forecast the number of first-grade primary school students. The process included the following conclusion:

Parameters	Estimation	Standard Error	Test statistic (t)	P-value
AR (1)	0.6835	0.1644	7.2872	0.000

Average	5.2285	1.6487	6.3973	0.003
Constant	0.7825	–	–	–

Criteria	**Result**
Mean absolute Error (MAE)	0.5496
Mean absolute percent Error(MAPE)	0.4863
Mean Error (ME)	1.1859
Mean percent Error(MPE)	0.2754
Mean squares Error(MSE)	0.3728

The value of the variance of the white noise estimates (residuals) in this model reached (0.2549) at a degree of freedom (19) and a standard deviation of (0.0735).

By reviewing the results of the probability value of the model, it becomes clear that it is not significant at the level of significance (0.05). On the other hand, the average value was recorded at a high level of significance, as the probability value to examine its significance exceeded $P < 0.1$.

The first three statistics shown above measure the importance of the residuals resulting from the differences recorded between the observed value at time t and the predicted value at time t-1, where the best model indicates the smallest value of those statistics, while the other two statistics measure the amount of bias, where the best model indicates when the values of the two statistics approach zero.

- **Forecast model: MA (1)*:*** The model (ARIMA(1,0,0)) was chosen from among the selected models to forecast the number of first-grade primary school students. The process included the following conclusion:

Parameters	**Estimation**	**Standard Error**	**Test statistic (t)**	**P-value**
AR (1)	0.7387	0.2331	8.0013	0.002
Average	4.9924	0.9446	7.5572	0.001
Constant	0.8116	–	–	–

Criteria	**Result**
Mean absolute Error (MAE)	0.6003
Mean absolute percent Error(MAPE)	0.5280
Mean Error (ME)	0.9223
Mean percent Error(MPE)	0.3047

Mean squares Error(MSE)	0.4627

The value of the variance of the white noise estimates (residuals) in this model reached (0.3188) at a degree of freedom (19) and a standard deviation of (0.0832).

By reviewing the results of the probability value of the model, it becomes clear that it is not significant at the level of significance (0.05). On the other hand, the average value was recorded at a high level of significance, as the probability value to examine its significance exceeded $P < 0.1$.

The first three statistics shown above measure the importance of the residuals resulting from the differences recorded between the observed value at time t and the predicted value at time t-1, where the best model indicates the smallest value of those statistics, while the other two statistics measure the amount of bias, where the best model indicates when the values of the two statistics approach zero.

- **Forecast model ARMA (1,1):** The model (ARMA (1,1)) was chosen from among the selected models to forecast the number of first-grade primary school students. The process included the following conclusion:

Parameters	Estimation	Standard Error	Test statistic (t)	P-value
AR (1)	0.7802	0.1775	8.4998	0.001
Average	6.3951	1.2105	8.0038	0.007
Constant	0.9075	–	–	–

Criteria	Result
Mean absolute Error (MAE)	0.4887
Mean absolute percent Error(MAPE)	0.3811
Mean Error (ME)	1.0265
Mean percent Error(MPE)	0.2881
Mean squares Error(MSE)	0.3383

The value of the variance of the white noise estimates (residuals) in this model reached (0.4002) at a degree of freedom (19) and a standard deviation of (0.08227).

By reviewing the results of the probability value of the model, it becomes clear that it is not significant at the level of significance (0.05). On the other hand, the average value was recorded at a high level of significance, as the probability value to examine its significance exceeded $P < 0.1$.

The first three statistics shown above measure the importance of the residuals resulting from the differences recorded between the observed value at time t and the predicted value at time t-1, where the best model indicates the smallest value of those statistics, while the other two statistics measure the amount of bias, where the best model indicates when the values of the two statistics approach zero.

- **Forecast model ARIMA (1,1,1):** The model (ARIMA(1,1,1)) was chosen from among the selected models to forecast the number of first-grade primary school students. The process included the following conclusion:

Parameters	Estimation	Standard Error	Test statistic (t)	P-value
AR (1)	0.5572	0.1162	6.9022	0.002
Average	4.9033	0.8946	7.2840	0.001
Constant	0.6072	–	–	–

Criteria	Result
Mean absolute Error (MAE)	0.6092
Mean absolute percent Error(MAPE)	0.5496
Mean Error (ME)	1.202
Mean percent Error(MPE)	0.3028
Mean squares Error(MSE)	0.4428

The value of the variance of the white noise estimates (residuals) in this model reached (0.3074) at a degree of freedom (19) and a standard deviation of (0.0893).

By reviewing the results of the probability value of the model, it becomes clear that it is not significant at the level of significance (0.05). On the other hand, the average value was recorded at a high level of significance, as the probability value to examine its significance exceeded $P < 0.1$.

The first three statistics shown above measure the importance of the residuals resulting from the differences recorded between the observed value at time t and the predicted value at time t-1, where the best model indicates the smallest value of those statistics, while the other two statistics measure the amount of bias, where the best model indicates when the values of the two statistics approach zero.

- **Forecast model Constant mean:** The Constant mean model was chosen from among the selected models to forecast the number of first grade primary school students to generate forecasting values. The process included the following summary:

Parameters	Estimation	Standard Error	Test statistic (t)	P-value
AR (1)	0.6047	0.2004	7.5998	0.003

Average	5.6744	0.7236	6.6299	0.004
Constant	0.7926	–	–	–

Criteria	Result
Mean absolute Error (MAE)	0.7851
Mean absolute percent Error(MAPE)	0.6839
Mean Error (ME)	1.3861
Mean percent Error(MPE)	0.4024
Mean squares Error(MSE)	0.5028

The value of the variance of the white noise estimates (residuals) in this model reached (0.4008) at a degree of freedom (19) and a standard deviation of (0.0963).

By reviewing the results of the probability value of the model, it becomes clear that it is not significant at the level of significance (0.05). On the other hand, the average value was recorded at a high level of significance, as the probability value to examine its significance exceeded $P < 0.1$.

The first three statistics shown above measure the importance of the residuals resulting from the differences recorded between the observed value at time t and the predicted value at time t-1, where the best model indicates the smallest value of those statistics, while the other two statistics measure the amount of bias, where the best model indicates when the values of the two statistics approach zero.

- **Forecast model Simple Exponential Smoothing:** The Simple Exp. Smoothing model was chosen from among the selected models to forecast the number of first grade students to generate forecasting values. The process included the following summary:

Parameters	Estimation	Standard Error	Test statistic (t)	P-value
AR (1)	0.6183	0.1547	8.4372	0.006
Average	5.3819	1.5886	5.7760	0.005
Constant	0.9033	–	–	–

Criteria	Result
Mean absolute Error (MAE)	0.5927
Mean absolute percent Error(MAPE)	0.4481
Mean Error (ME)	1.0052
Mean percent Error(MPE)	0.3036
Mean squares Error(MSE)	0.4818

The value of the variance of the white noise estimates (residuals) in this model reached (0.2886) at a degree of freedom (19) and a standard deviation of (0.0749).
By reviewing the results of the probability value of the model, it becomes clear that it is not significant at the level of significance (0.05). On the other hand, the average value was recorded at a high level of significance, as the probability value to examine its significance exceeded $P < 0.1$.
The first three statistics shown above measure the importance of the residuals resulting from the differences recorded between the observed value at time t and the predicted value at time t-1, where the best model indicates the smallest value of those statistics, while the other two statistics measure the amount of bias, where the best model indicates when the values of the two statistics approach zero.
The following table is for the indicator of forecasting the number of first-grade primary school students.

Models	**Model set**			**Bias**	
	MSE	**MAE**	**MAPE**	**ME**	**MPE**
AR (1)	0.2956	0.2592	1.3879	0.0504	0.3189
MA (1)	0.6924	0.6994	5.6982	0.0406	0.2892
ARMA (1,1)	0.4952	0.3738	1.6861	0.0334	0.2004
ARIMA(1,1,1)	0.5025	0.4935	2.4901	0.0104	0.4026
Constant mean	1.2729	1.1083	7.6994	0.0406	0.2893
Simple Exponential Smoothing	0.3821	0.2206	1.1665	0.0334	0.3008

For the first column, column (MSE), the best model was AR(1), and the model was Simple exponential smoothing. For the second column, column (MAE), and the third column, column (MAPE), the best model was Simple exponential smoothing. as for the fourth column, column (ME), the best model was ARIMA(1,1,1), and as for the fifth column, column (MPE), the best model was ARMA(1, 1).

Table of forecasting values for the number of first grade students to generate by using different models:

Years	**Models**	**forecasting**
	AR (1)	106773

2025	**MA (1)**	107509
	ARMA (1,1)	109826
	ARIMA(1,1,1)	103917
	Constant mean	105582
	Simple Exponential Smoothing	106003
2026	**AR (1)**	107217
	MA (1)	108114
	ARMA (1,1)	108745
	ARIMA(1,1,1)	104684
	Constant mean	106289
	Simple Exponential Smoothing	107300
2027	**AR (1)**	107964
	MA (1)	108765
	ARMA (1,1)	109276
	ARIMA(1,1,1)	105187
	Constant mean	106823
	Simple Exponential Smoothing	107848
2028	**AR (1)**	108104
	MA (1)	109176
	ARMA (1,1)	109985
	ARIMA(1,1,1)	105710
	Constant mean	107228

	Simple Exponential Smoothing	108201
2029	**AR (1)**	108538
	MA (1)	109693
	ARMA (1,1)	110617
	ARIMA(1,1,1)	105710
	Constant mean	107695
	Simple Exponential Smoothing	108602
2030	**AR (1)**	109154
	MA (1)	112074
	ARMA (1,1)	111072
	ARIMA(1,1,1)	109946
	Constant mean	108392
	Simple Exponential Smoothing	109207

Conclusions and Recommendations

As for the table for the indicator of the number of first-grade primary school students, the best model for the indicators was the mean square error (MSE), the autoregressive model AR(1), and for the mean absolute errors (MAE) and the mean absolute percentage errors (MAPE). The rates of the number of first-grade students during the series studied are simple but increasing.

It was noted the importance of using a wide network of statistical models in light of several indicators for the purpose of making a comparison in selecting the best model, as the variation in the results of the selection process may provide more efficient inference in selecting the best model.

Therefore, we concluded that there was a state of disturbance in the selection of the model on the one hand and in the adoption of the degree of bias on the other hand.

Educational and leadership bodies must pay attention to the forecasting student numbers for the coming years and what preparations are needed in terms of buildings, schools, and teachers. This is one of the most important objectives that researchers have reached.

We recommend that researchers, when making any comparisons or predictions of models, use a large number of indicators to provide wider options, as the large number of indicators leads to more positive results, not negative ones, to know the best model.

References:

1. Bibi, A. & Oyet, A.J. (2014) "Estimation of Some Bilinear Time Series Models With Time Varying Coefficients" Journal of Stochastic Analysis & Application ,Vol. 22 ,No.2, PP.355-376.
2. Box G, E.P & Jenkins, G. M., (2016) "Time series analysis forecasting and control), sanfrancisco Helden-day".
3. Box, G. E. and Tiaa, (2012)"Distribution of Residual Autocorrelation in Multiple Autoregressive", JASA, Vol. (74), pp. (928-934).
4. Brockwell, P. & Davis, R. (2000) "Introduction to Time Series and Forecasting" Springer-Verlag Co.
5. De Gooijer, J.G. & Hyndman, R.J. (2018) "25 Years of Time Series Forecasting" International Journal of Forecasting, Vol.22, PP.443-473.
6. Hamilton, J.D. (2014) "Time Series Analysis" Princeton University Press Princeton, New Jersey.
7. Li, W.K. (2000) "A Simple One Degree of Freedom Test for Time Series Model Discrimination" Statistica Sinica, Vol. 3, PP.245-254.
8. Makridakis, S. & Wheelwright, S.C. & Hyndman, R.J. (2002) "Forecasting: Methods and Application" 3rd edition, John-Wiley & Sons Inc.
9. Makridakis, S., Wheelwright ,S.C.& McGee, V.E. (2008), "Forecasting Methods and Application" New York, John Wiley & Sons .
10. Pankratz , Alan , (2013) : " Forecasting with Dynamic Regression Models " , DePauw University , Greencastle , Indiana .
11. Tsay, R.S. (2012) "Analysis of Financial Time Series" John-Wiley & Sons Inc.
12. Wegmon. Edward. J ,(1998) : " Time Series Analysis Theory " , Data Analysis and Computation .

13. Wei, W.W.S. (2000) "Time Series Analysis: Univariate and Multivariate Methods" Addison-Wesley.
14. Wei, William W.S. (2009), "Time Series Analysis, Univariate& Multivariate Methods", Addison-Wesley Publishing Company Inc.