An-Najah University Journal for Research – B

Humanities







# Forecasting the Success Rate of Baccalaureate Exams in Algeria Using ARIMA Model: A Statistical Analysis

Kada Becharef<sup>1,\*</sup>, Abderrezzaq Bouziani<sup>1</sup>, Keltoum Nouri<sup>2</sup>, Amel Belhamiti<sup>1</sup> & Said Chibani<sup>1</sup> Received: 27<sup>th</sup> Jun. 2024, Accepted: 17<sup>th</sup> Nov. 2024, Published: ××××

## **Accepted Manuscript, In press**

**Abstract: Objectives:** This study aims to employ the Box-Jenkins methodology to forecast high school graduation rates in Algeria from 2023 to 2030, developing an optimal ARIMA model based on 59 annual baccalaureate exam observations, and providing evidencebased predictions for educational planning. **Methodology:** The study analyzed 59 annual observations of baccalaureate exam results from 1963 to 2022, applying time series analysis techniques and conducting rigorous testing and differentiation. After extensive analysis, the ARIMA (6, 11, 1) model was implemented as the most suitable forecasting tool. **Results:** The analysis revealed that the time series achieved stationarity at the first degree, and the ARIMA (6, 11, 1) model successfully captured graduation rate fluctuations. The model demonstrated high predictive accuracy, with forecasts indicating a steady increase in graduation rates through 2030. **Conclusions:** The selected model effectively predicts Algerian high school graduation rates, with predictions suggesting a positive trend in educational outcomes. The time series analysis proves viable for educational forecasting in this context. **Recommendations:** Future research should implement these findings for educational policy planning, consider additional variables in future models, and explore the model's adaptability across diverse educational settings.

**Keywords:** ARIMA modeling; Box-Jenkins methodology; educational forecasting; time series analysis; Algerian education system; baccalaureate exam success rates.

# التنبؤ بمعدل نجاح امتحانات البكالوريا في ا<mark>ل</mark>جزائر باس<mark>تخ</mark>دام نموذج ARIMA تحليل إحصائي

بشارف قادة<sup>1،\*</sup>، وبوزياني عبد الرزاق<sup>1</sup>،ونوري كلثومة<sup>2</sup>، وبلحمي<mark>ت</mark>ي امال<sup>1</sup>، وشيباني سعيد<sup>1</sup> تاريخ التسليم: (2024/6/27)، تاريخ القبول: (2024/11/17)، تاريخ النشر: ××××

**الملخص: الأهداف:** تهدف هذه الدراسة إلى توظيف منهجية بوكس-جينكينز للتنبؤ بمعدلات نسبة النجاح في شهادة البكالوريا في الجزائر من 2023 إلى 2030، وتطوير نموذج ARIMA الأمثل بناءً على 59 ملاحظة سنوية لامتحانات البكالوريا، وتقديم تنبؤات مبنية على الأدلة للتخطيط التربوي. **المنهجية:** حللت الدراسة 59 ملاحظة سنوية لنتائج امتحانات البكالوريا من 2020، مطبقة تقنيات تحليل السلاسل الزمنية وإجراء اختبارات وتفاضلات دقيقة. الدراسة 59 ملاحظة سنوية لنتائج امتحانات البكالوريا من 2023، مطبقة تقنيات تحليل السلاسل الزمنية وإجراء اختبارات وتفاضلات دقيقة. وبعد التحليل المكثف، تم تطبيق نموذج (1, 11) (6, 11) ARIMA كأداة تنبؤ أكثر ملاءمة. **النتائج:** كشف التحليل أن السلسلة الزمنية حققت الاستقرار في الدرجة بعد التحليل المكثف، تم تطبيق نموذج (1, 11) ARIMA كأداة تنبؤ أكثر ملاءمة. **النتائج:** كشف التحليل أن السلسلة الزمنية حققت الاستقرار في الدرجة الأولى، ونجح نموذج (1, 11) ARIMA في رصد تقلبات معدلات النجاح. أظهر النموذج دقة تنبؤية عالية، مع توقعات تشير إلى زيادة مطردة في معدلات النجاح حتى عام 2030. **الاستقرار ألى 20**ال بشكر فعال بمعدلات النجاح حتى عام 2030. **الاستنتاجات: ي**نبأ النموذج المختار بشكل فعال بمعدلات النجاح في المدارس الثانوية الجزائرية، مع تنبؤات تشير إلى اتجاه إيجابي في النتائج التعليمية. يثبت تحليل السلاسل الزمنية جدواه للتنبؤ التربوي في هذا السياق. **التوصيات:** ينبغي أن يطبق البحث المستقبلي هذه النتائج في في النتائج التعليمية. والنظر في متغيرات إضادية في النتائج التعليمية. والمحال الملاسل الزمنية جدواه للتنبؤ التربوي في هذا السياق. **التوصيات:** ينبغي أن يطبق البحث المستقبلي هذه النتائج في المتائج التعليمية، والنظر في متغيرات إضادية في النتائج المتعليمية، والنظر في متغيرات إضادينية في النتائج المعامية، والنسلاس الترمنية جدواه للتنبؤ التعليمي؛ تحليل السلاسل الزمنية؛ نظام التعليم، ومخليان المحل الموضاع التعليمية. النماذ من مالتعليمة، والماليات والنائية ولي النتائج النعليمية، والنائية النموذج ومنتوا المالي معطبط السياسات التعليمية، والنظر في متغيرات إضادية النماذير المعليم؛ تحليل السلاسل الزمنية؛ نظام التعليم الموضاع الكلمات المفتاحية: مدادة المعنية بوكس-جدائية؛ التستقبلية، واستكشاف قابلية تكيف الموذج عبر مختلف الأوضاع التعليمي الرعان

1 قسم البحث في تكنولوجيا التعليم؛ المعهد الوطني للبحث في التربية العاشور، الجزائر

\* المؤلّف المراسل: becharef\_kada@yahoo.fr 2 مختبر LTC ، قسم الإلكترونيك، كلية التكنولوجيا، جامعة سعيدة الدكتور مولاي طاهر سعيدة، الجزائر.

\*Corresponding author email: becharef\_kada@yahoo.fr 2 LTC Laboratory, Department of Electronic, Faculty of Technology, University of Saida Dr. Moulay Tahar Saida, Algeria.

Research in Education Al-Achour, Algeria

1 Division of Research in Education Technology; National Institute for

## Introduction

Education plays a vital role in human progress (Alaazam, M., & Olimat, S. (2015)) (Nimer, M., & al-Jarah, A. al-Mohdi. (2015)), fundamentally impacting individuals' and communities' development (Hanushek & Woessmann, 2020). High school education, in particular, is a crucial stage that prepares students for further academic and professional pursuits (Murnane, 2013). In recent years, high school graduation rates have been of great concern to educational policymakers and researchers worldwide, serving as an essential indicator of education quality and a predictor of future social and economic trends (Rumberger & Losen, 2016; Cutler & Lleras-Muney, 2006).

Algeria has shown significant interest in improving its high school education system, implementing various policies and initiatives to enhance education quality (Kada et al, 2024). As a result, high school graduation rates in Algeria have steadily increased over the past few years (UNESCO Institute for Statistics, 2023). However, forecasting these rates remains a complex task requiring sophisticated statistical methods.

The Box-Jenkins methodology, widely used for time series forecasting, has been successfully applied in various fields (Box et al, 2015). This study aims to apply this methodology to predict Algeria's high school graduation rates from 2023-2030, focusing on constructing a standard ARIMA model based on time series analysis.

The research questions guiding this study are:

- Does the time series of annual high school baccalaureate examination rates contain a unit root?
- What level of integration can be relied upon in the time series of annual high school baccalaureate examination rates?
- What is the optimal arrangement of autoregressive and moving average models that minimize the Akaike information criterion (AIC) value and maximize forecast accuracy?

This study contributes to the current knowledge of time series forecasting in educational contexts and provides valuable insights for educational policymakers and stakeholders in Algeria.

#### Literature Review and Theoretical Framework

Previous research has demonstrated the applicability of time series analysis and forecasting methodologies in various fields, including education. Tariq et al. (2018) successfully used the Box-Jenkins methodology to forecast enrollment rates in higher education institutions in Pakistan. Similarly, Chen et al. (2017) applied this methodology to forecast electricity demand in Taiwan. In the education sector, Alghamdi (2019) and Akomolafe and Adewumi (2018) used time series analysis to forecast student enrollment and graduation rates in Saudi Arabian and Nigerian universities, respectively. The theoretical framework for this study is based on time series analysis and the Box-Jenkins methodology. Time series analysis examines and forecasts trends in data collected at consistent intervals over time (Brockwell & Davis, 2016). The Box-Jenkins methodology relies on three fundamental elements: autoregression (AR), moving average (MA), and integration (I) (Hyndman & Athanasopoulos, 2021).

The application of the Box-Jenkins methodology in this study is based on several key assumptions:

- The presence of a unit root in the time series may necessitate differencing to achieve stationarity (Dickey & Fuller, 1979).
- Determining the appropriate level of integration is critical for identifying the degree of differencing needed to stabilize the time series (Kwiatkowski et al., 1992).
- The optimal arrangement of ARIMA parameters is selected based on minimizing the Akaike Information Criterion (AIC) and ensuring forecast accuracy (Akaike, 1974).

#### Methodology

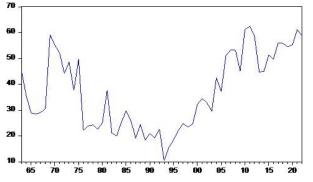
This study analyzes the annual mean results of the high school baccalaureate in Algeria, denoted by (X), spanning from 1963 to 2022 with 59 observations. We employed the following steps:

- Data Collection and Preliminary Analysis: We collected data on annual baccalaureate exam results from 1963 to 2022.
- Stationarity Testing: We used the Augmented Dickey-Fuller (ADF) test to check for stationarity in the time series.
- Model Identification: We analyzed the autocorrelation and partial autocorrelation functions to identify potential ARIMA models.
- Model Estimation and Selection: We estimated multiple models and selected the best one based on criteria such as AIC, Schwarz Criterion (SC), and R<sup>2</sup>.
- Model Validation: We conducted various diagnostic tests to ensure the selected model's appropriateness.
- Forecasting: We used the validated model to forecast graduation rates for 2023-2030.

#### **Results and Discussion**

This study's analysis is based on the annual mean results of the high school baccalaureate in Algeria, denoted by (X), spanning from 1963 to 2022 with 59 observations. The mean value of X during the study period was 36.89%, with the highest value of 62.45% recorded in 2011 and the lowest value of 10.54% in 1993.

The analysis indicates that the annual average results exhibit random fluctuations, as shown in Figure 1, and present an overall irregular trend. Thus, examining the stability of the time series is crucial.



х

Figure (1): The Annual Mean Results Ratio of the High School Baccalaureate Stage in Algeria.

The self-correlation and partial correlation functions of the time series, displayed in Figure 2, show outliers beyond the 95% confidence interval. This deviation from zero at the 5% significance level suggests that the time series is unstable and prone to trend fluctuations.

Sample: 1963 2022

ncluded observations: 6
-------------------------

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.821	0.821	42.448	0.000
·	ı 🗖 ı	2	0.728	0.167	76.413	0.000
· 🛏 🔤	i   i	3	0.632	-0.006	102.52	0.000
· 👝	ı <b> </b>   ı	4	0.597	0.141	126.23	0.000
· 📁	ı <b>d</b> ı	5	0.503	-0.129	143.37	0.000
· 👝		6	0.485	0.140	159.59	0.000
· 💻	· •	7	0.365	-0.252	168.96	0.000
· 💻	1 I 🛛 I I I	8	0.318	0.051	176.18	0.000
· 💻	l i <b>⊒</b> i	9	0.297	0.149	182.62	0.000
· 💻	1 I I I	10	0.305	0.025	189.54	0.000
· 💻	1 1 1	11		-0.033	194.38	0.000
· 🗐 ·	i 🗖 '	12		-0.342	196.11	0.000
1 <b>1</b> 1	י 🗖 י	13		-0.168	196.16	0.000
	· · • ·	14	-0.017	0.095	196.18	0.000
· 🛛 ·	1 1 1	15	-0.058		196.45	0.000
	· •		-0.129		197.86	0.000
· <b>–</b> ·	י 🗖 י	17	-0.211		201.72	0.000
'	· · · ·		-0.270	0.009	208.19	0.000
'	' <b> </b> '	19	-0.307	0.043	216.72	0.000
'	! ' <b>⊟</b> '	20	-0.329	-0.183	226.81	0.000
		21	-0.350	-0.084	238.50	0.000
		22		0.069	252.69	0.000
			-0.407	0.110	269.33	0.000
		24	-0.441	-0.116	289.42	0.000
		25	-0.394	0.055	305.92	0.000
		26	-0.374	0.031	321.20	0.000
		27	-0.369	0.053	336.54	0.000
	ים י	28	-0.392	-0.101	354.37	0.000

Figure (2): The Self-Correlation and Partial Correlation Functions for the Time Series of the Annual Mean Results Ratio of the High School Baccalaureate Stage in Algeria.

To further validate the stationarity of the series, the Dickey-Fuller (ADF) test (Dickey, DA. and WA. Fuller 1979) was performed, with results summarized in Table 1.

Table (1): Results of the Dickey-Fuller Test on Sequence (X).

The outcome		f the Dickey- er Test,	The outcome
outcome	Prop	t.Statistic	
Instability	0.6128	1.323802	Instability
Instability	0.6576	1.869199	Instability
Instability	0.6632	0.046484	Instability

The test confirmed that the time series is non-stationary, indicated by probability values exceeding 0.05. To address this, first-degree differencing was applied, yielding a stationary series as shown by the second application of the ADF test (Table 2), where the probabilities dropped to 0.0000.

Table (2): Results of the Develope	r Dickey-Fuller Test on Series (	X).
------------------------------------	----------------------------------	-----

The outcome	At the onset of the initial disparities		Variables
	t.Statistic	Prop	
Establishments	9.992164	0.0000	The relentless progression of time unfolds without any discernible direction
Establishments	9.997664	0.0000	In spite of impediments and temporal direction
Establishments	10.05335	0.0000	The relentless progression of time unfolds without any discernible direction

Subsequent model selection was based on the autocorrelation and partial autocorrelation coefficients, displayed in Figure 3, which identified models AR(1) and AR(11) as suitable candidates. After evaluating multiple models using criteria such as Akaike Information Criterion (AIC), Schwarz Criterion (SC), and R<sup>2</sup>, the ARIMA (6, 11, 1) model was selected (Table 3) as the most appropriate.

Sample: 1963 2022

Included observations: 59

=

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob
		1 -0.269	-0.269	4.5050	0.034
			-0.024	4.6669	0.097
່ 🗖 ເ		3 -0.187	-0.193	6.9028	0.075
· 🗖 ·		4 0.176	0.085	8.9300	0.063
	, <b>E</b>	5 -0.245	-0.204	12.933	0.024
· 🗖	i 🗖 🗖	6 0.299	0.196	18.996	0.004
i 🗖 i		7 -0.229	-0.120	22.628	0.002
1.		8 -0.044	-0.200	22.767	0.004
101	1 1 1 1	9 -0.074	-0.034	23.156	0.006
1 🗐 1	iĝi	10 0.105	-0.084	23.972	0.008
· 🗐 ·		11 0.161	0.322	25.927	0.007
1 <b>þ</b> 1		12 0.052	0.050	26.137	0.010
י 🗖 י	ון ו	13 -0.128	-0.053	27.416	0.011
1 <b>(</b> )		14 -0.049	-0.019	27.606	0.016
1 <b>D</b> 1		15 0.093	0.003	28.317	0.020
- i 🏚 i	ı 🗖 ı	16 0.048	0.163	28.511	0.027
1 1	<b>E</b>	17 -0.005	-0.078	28.513	0.039
1 🔲 1	( <b>1</b> )	18 -0.126	-0.090	29.904	0.038
1 <b>j</b> 1	i 🗖 i	19 0.060	0.163	30.225	0.049
1 1	1 1 1	20 -0.023	0.021	30.276	0.066
	1. 1. 1	21 0.011	-0.047	30.288	0.086
	1 <b>E</b> 1	22 -0.012	-0.154	30.303	0.111
- <b>p</b> -	<b> </b>	23 0.074	0.051	30.855	0.126
<u> </u>	י <b>ב</b> י ו	24 -0.237	-0.103	36.621	0.048

Figure (3): The Series D(X) Simple and Partial Autocorrelation Function.

Dependent Variable :D(X) Method:ARMA Conditional Least Squares (Gauss-Newton/Marquardt steps)							
Variable	Coefficient	Std Error	t-Statistic	Prob			
AR(11) MA(1) MA(6)	0.315609 -0.190551 0.759339	0.137504 0.086603 0.071482	2.295266 -2.200292 10.62287	0.0264 0.0330 0.0000			
R-squared Adjusted R-squared S.E. of regression Sum squared resid Log likelihood Durbin-Wats stat	0.402437 0.375878 6.172475 1714.475 -153.9249 2.183574	Mean depe S.D. deper Akaike info Schwarz Hannan-Qu	ndent var o criterion criterion	0.438958 7.813122 6.538538 6.655188 6.582734			
Inverted AR Roots	.90 .37+.82i 59+.68i	.7649i 1389i 8625i	.76+.49i 13+.89i 86+.25i	.3782i 5968i			
Inverted MA Roots	.86+.48i 80+.48i	.86-48i 8048i	.0395i	.03+.95i			

Validation of the estimated model: To ensure that the selected model is appropriate for representing the analysed data and to use it for future predictions, The following tests were conducted to validate the stationarity and accuracy of the ARIMA model.

**Matching test:** Comparing the original and forecasted sequences (Figure 4) shows a similarity, but residual fluctuations around.

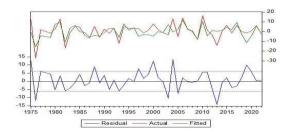


Figure (4): A comparison between the original and estimated sequences.

**zero were observed:** The Ljung-Box statistic (18.585) and its associated p-value indicated no serial correlation, as shown in Figure 5, confirming that the residual series is stationary.

Sample: 1963 2022
Included observations: 48
Q-statistic probabilities adjusted for 3 ARMA terms

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob	
. 🖬 .		1 -0.149	-0.149	1,1352		
		2 0.017	-0.006	1,1499		
. <u>.</u> .	ו בי	3 -0.104	-0.104	1.7224		
· 🗖 ·		4 0.144	0.117	2.8597	0.091	
· 🖬 ·	i i 🖬 i	5 -0.193	-0.163	4.9362	0.085	
· 🖬 ·		6 -0.184	-0.256	6.8737	0.076	
		7 -0.020	-0.070	6.8960	0.141	
1 🖬 1	i 🖬 i	8 -0.131	-0.227	7.9207	0.161	
	1 1 1 1	9 0.019	-0.056	7.9424	0.242	
· 🗖 ·		10 0.216	0.250	10.897	0.143	
· 🗐 ·		11 0.101	0.086	11.555	0.172	
		12 -0.031	-0.019	11.619	0.236	
- <b>- - -</b>	ı <b>=</b>  ı	13 -0.096	-0.174	12.251	0.269	
- <b>- </b>	🗖 '	14 -0.055	-0.315	12.467	0.330	
<b>- - -</b>		15 -0.052	-0.134	12.666	0.394	
· 🖬 ·	1 1 1 1	16 -0.100	-0.042	13.409	0.417	
· 🔲 ·	ı <b>=</b>   ı	17 -0.170	-0.158	15.638	0.336	
· 🖬 ·		18 -0.095	-0.137	16.360	0.359	
· 🗐 ·	1 1 1 1	19 0.125	-0.045	17.654	0.345	
· 🗐 ·	, <b>E</b> ,	20 0.104	-0.128	18.585	0.353	

Figure (5): Autocorrelation and Partial Autocorrelation Function of the Residual Sequence.

The Breusch-Godfrey test confirmed no autocorrelation issues within the model (Table 4).

 Table (4): Results of the Breusch-Godfrey Test

Breu	sch-Godfrey	Serial Correlation LM Test:	
F-statistic	0.553042	Prob.F(2,43)	0.5792
ObsR-squared	1.203735	Prob.Chi-Square(2)	0.547 <mark>8</mark>

The normality of residuals was examined using the Jarque-Bera test (Figure 6), and results confirmed that the residuals follow a normal distribution. Furthermore, both simple and partial autocorrelation functions (Figure 7) fell within the confidence interval, indicating stable variance.

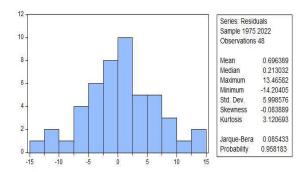


Figure (6): The normal distribution test for the remainder of the estimate.

Sample: 1	963	2022	
Included o	bse	rvations:	48

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
	· b.	1	0.155	0.155	1.2243	0.269
· 🖬 ·	i . 🖬 .	2	-0.148	-0.177	2.3729	0.305
		3	0.012	0.070	2.3809	0.497
		4	-0.022	-0.067	2.4075	0.661
· 🖬 ·	I   I	5	-0.146	-0.124	3.5990	0.608
- <b>1</b> - 1		6	0.026	0.067	3.6385	0.725
1 1		7	-0.004	-0.072	3.6395	0.820
· 🖬 ·	I   I	8	-0.107	-0.075	4.3301	0.826
· p ·		9	0.081	0.105	4.7335	0.857
1 1	- <b></b> -	10	-0.002	-0.098	4.7338	0.908
· 🔲 ·	! "티 "	11	-0.163	-0.108	6.4585	0.841
1 <b>1</b> 1		12	-0.023	0.003	6.4944	0.889
· 🗖 ·		13	0.127	0.066	7.5923	0.869
· 🗐 ·	' <b>  </b> '	14	-0.142	-0.167	9.0165	0.830
		15	0.014	0.095	9.0300	0.876
· •		16	0.074	-0.047	9.4398	0.894
· 🗖 ·	· 🖬 ·	17	-0.167	-0.163	11.604	0.823
, <b>⊟</b> , i	- E -	18	-0.176	-0.110	14.068	0.725
- <b>(</b> -	1 <b>I</b>	19	-0.064	-0.154	14.408	0.759
- <b>(</b> -	1 (1)	20	-0.061	-0.046	14.732	0.792

Figure (7): The Autocorrelation Function of the Squares of the Residuals.

Finally, the inverse roots of the ARIMA model (Figure 8) indicated that all roots lie within the unit circle, confirming the stability of the model. The BDS independence test (Table 5) further demonstrated that the time series is independent, allowing for short-term predictability.

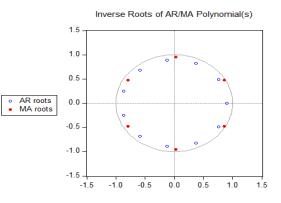


Figure (8): The results of the inverse roots of the estimated model (1, 11, 6) ARIMA

Table (5): Results of the BDS Independent Observations Test.

BDS Test for X Included observations :60					
Dimension	BDS Statistic	Std.Error	z-Statistic	Prob	
2	0.101629	0.005674	17.90981	0.0000	
3	0.165766	0.009069	18.27885	0.0000	
4	0.202838	0.010854	18.68714	0.0000	
5	0.218670	0.011370	19.23249	0.0000	
6	0.215276	0.011019	19.53648	0.0000	

The statement that the Box-Jenkins model is effective for short-term forecasting refers to its ability to produce reliable forecasts within a relatively limited time frame, typically 1-3 years ahead. In this context, 'short term' is defined by the stability of the model and the decreasing accuracy of predictions as the forecast horizon extends. The validation tests confirmed that the residuals exhibit characteristics of white noise, ensuring reliable short-term predictions but warranting caution in interpreting longer-term forecasts.

**Prediction stage:** The prediction stage, the most crucial phase of the Box-Jenkins methodology, is conducted after the model has been validated. To facilitate this phase, we used Eviews 10 software to forecast and obtain the values instantly. The table below shows Algeria's predicted percentages of official baccalaureate exams from 2023 to 2030.

After validating the model, the most critical phase of the Box-Jenkins methodology is the prediction stage. We used Eviews 10 software to make forecasts and obtain immediate results. The table below displays the predicted percentages of official baccalaureate exams in Algeria from 2023 to 2030. Table (6): Forecasted Baccalaureate Success Rates (2023-2030).

2023	55.17454
2024	52.40434
2025	60.00626
2026	66.45730
2027	66.20410
2028	68.18612
2029	68.12616
2030	67.70955

The prediction stage is crucial in the Box-Jenkins methodology, and it follows the validation of the model. We used Eviews 10 software to forecast the values and obtain instant results. The table below showcases the projected percentages of official baccalaureate exams in Algeria from 2023 to 2030.

**Thule's criterion for inequality:** Thale's criterion for inequity states that a prediction is considered reliable when the calculated statistic is equal to zero, denoted by U = 0. On the other hand, if U = 1, it indicates a failed prediction.

Table 7 indicates a failed operation as the calculated value of Thale's criterion for inequality, U, equals 1, indicating that the forecasted values deviate significantly from the actual values. This suggests better models than the ARIMA (6,11,1) model for predicting the proportion of secondary education graduates in Algeria from 2023 to 2030.

Upon examining Table 7, we can deduce that the value of Thale's criterion for inequality, U, equals 0, which is closer to 0 than 1. This implies that the ARIMA (6, 11, 1) model can predict the reality of the baccalaureate examination success rate in Algeria. Therefore, using this model to accurately forecast the baccalaureate exam success rate in Algeria is valid.

 Table (7): Forecast Accuracy and Model Validation Metrics for ARIMA (6, 11, 1) in Predicting Baccalaureate Success Rates

Forecast: XF Actual: X Forecast sample: 1963 2026 Adjusted sample: 1975 2026 Included observations: 52			
Root Mean Squared Error	15.68694		
Mean Absolute Error	14.24804		
Mean Abs. Percent Error	48.60064		
Theil Inequality Coefficient	0.209778		
Blas Proportion	0.000002		
Variance Proportion	0.855256		
Covariance Proportion	0.144742		
Theil U2 Coefficient	2.500221		
Symmetric MAPE	40.79358		

The ARIMA (6, 11, 1) model accurately captured the fluctuations in Algeria's high school graduation rates and produced reliable forecasts for the 2023–2030 period. These findings align with previous studies that applied the Box-Jenkins methodology in educational contexts (Tariq et al., 2018; Chen et al., 2017). However, the variation in graduation rates across years reflects Algeria's changing educational landscape, potentially influenced by factors such as reforms in education policies, student access to resources, and socioeconomic conditions

The forecasts provided by the model show a steady increase in graduation rates, though some fluctuations are expected. The predicted values offer critical insights for policymakers, aiding resource allocation and future planning.

#### **Conclusion, Limitations, and Recommendations**

The Box-Jenkins methodology, through the ARIMA (6,11,1) model, successfully forecasts Algeria's high school graduation rates. The model's predictions are highly reliable, supported by multiple statistical tests. The increasing trend in predicted graduation rates suggests a positive outlook for the Algerian education system.

However, several limitations of this study should be acknowledged:

- Sample Size: The data set spans only 59 years, which may limit the generalizability of the findings. While this provides sufficient data for the ARIMA model, a longer time series could potentially enhance the model's predictive accuracy.
- Exclusion of External Variables: The ARIMA model does not account for external variables such as economic conditions, educational policy changes, or socioeconomic factors that may influence graduation rates. The model's univariate nature, while effective for time series analysis, may not capture the full complexity of factors affecting graduation rates.
- Model Assumptions: The ARIMA model assumes stationarity, which may not fully capture the complexity of the data. While differencing improved stationarity in our analysis, other models like structural breaks or regime-switching could provide a more nuanced analysis of the graduation rate patterns.

Based on these findings and limitations, we recommend the following:

- 1. For Educational Policymakers:
  - Use these forecasts as a baseline for resource planning and allocation
  - Consider implementing systematic data collection for additional variables that may affect graduation rates
  - Develop contingency plans for scenarios where actual rates deviate from predictions
- 2. For Future Research:
  - Incorporate additional variables to improve the model's predictive power, such as:
    - Economic indicators
    - Educational policy changes
    - Demographic factors
    - Resource allocation metrics
  - Explore alternative modeling approaches that can account for structural breaks and regime changes
  - Conduct comparative analyses with other educational systems in similar contexts
  - Investigate the impact of specific policy interventions on graduation rates
- 3. For Educational Institutions:
  - Develop systematic approaches to data collection and analysis
  - Consider implementing early warning systems based on the predictive models
  - Use the forecasts for capacity planning and resource allocation

The insights gained from this study can significantly inform educational policy and resource planning in Algeria. While acknowledging the limitations, the model provides valuable predictions that can help stakeholders make informed decisions about the future of education in Algeria. The recommended extensions for future research will further enhance our understanding of the factors influencing graduation rates and improve the accuracy of forecasting models.

#### **Disclosure Statement**

- Ethical approval and consent to participate:

- This article does not contain any studies with human participants or animals performed by any authors.
- Availability of data and materials:
- The data analyzed in this study are available from the corresponding author upon reasonable request.

#### References

- Alaazam, M., & Olimat, S. (2015). The Level of Applying the Accreditation Standards for the Postgraduate Educational Programs in the Faculties of Education in the Jordanian Governorate Universities. An-Najah University Journal for Research - B (Humanities), 29(7), 1345–1372. <u>https://doi.org/10.35552/0247-029-007-006</u>
- Akaike, H. (1974). A new look at the statistical model identification. IEEE Transactions on Automatic Control, 19(6), 716-723.
- Akomolafe, A. J., & Adewumi, A. O. (2018). Time series modeling and forecasting of university enrollment and graduation rates in Nigeria. *Journal of Statistics Applications and Probability*, 7(1), 115-128.
- Alghamdi, A. M. (2019). Forecasting student enrollment in Saudi Arabian universities using time series analysis. *Journal of Education* and Practice, 10(20), 79-85.
- Box, G. E., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. (2015). *Time series analysis: Forecasting and control.* John Wiley & Sons.
- Box, G. E. P., Jenkins, G. M., & Reinsel, G. C. (2015). *Time series analysis: Forecasting and control* (5th ed.). Wiley.
- Brockwell, P. J., & Davis, R. A. (2016). Introduction to time series and forecasting. Springer.
- Burnham, K. P., & Anderson, D. R. (2004). Multimodel inference: Understanding AIC and BIC in model selection. Sociological Methods & Research, 33(2), 261-304.
- Chatfield, C. (2016). The analysis of time series: An introduction. CRC Press.
- Chen, W. C., Liu, C. H., & Yu, H. C. (2017). Forecasting the demand for electricity using Box-Jenkins methodology: A case study of Taiwan. *Energy Procedia*, 142, 347-352.
- Cutler, D. M., & Lleras-Muney, A. (2006). Education and health: Evaluating theories and evidence (Working Paper No. w12352). National Bureau of Economic Research.
- Dickey, D. A., & Fuller, W. A. (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association*, 74(366a), 427-431.
- Durbin, J., & Koopman, S. J. (2012). Time series analysis by state space methods. Oxford University Press.
- Nimer, M., & al-Jarah, A. al-Mohdi. (2015). The Degree of Chemistry

Teachers Practice of Educational Technology Competencies from Their Perspective and Their Students Point of View in Jordan. *An-Najah University Journal for Research - B (Humanities), 29*(5), 961– 998. https://doi.org/10.35552/0247-029-005-007

- Hamilton, J. D. (2020). *Time series analysis*. Princeton University Press.
- Hanushek, E. A., & Woessmann, L. (2015). The knowledge capital of nations: Education and the economics of growth. MIT Press.
- Hanushek, E. A., & Woessmann, L. (2020). Education, knowledge capital, and economic growth. In The economics of education (pp. 171-182). Academic Press.
- Hyndman, R. J., & Athanasopoulos, G. (2018). Forecasting: Principles and practice. OTexts. <u>https://otexts.com/fpp3/</u>
- Hyndman, R. J., & Athanasopoulos, G. (2021). Forecasting: Principles and practice (3rd ed.). OTexts. <u>https://otexts.com/fpp3</u>
- Kada, B., Amel, B., Abderrezzaq, B., & El Batoul, A. (2024). Navigating the future of education: Harnessing data mining to illuminate pathways to success. Perspectives on Global Development and Technology, 23(1-2), 109-127. https://doi.org/10.1163/15691497-12341676
- Kwiatkowski, D., Phillips, P. C., Schmidt, P., & Shin, Y. (1992). Testing the null hypothesis of stationarity against the alternative of a unit root. Journal of Econometrics, 54(1-3), 159-178.
- Ljung, G. M., & Box, G. E. (1978). On a measure of lack of fit in time series models. Biometrika, 65(2), 297-303.
- Makridakis, S., Wheelwright, S. C., & Hyndman, R. J. (2008).
   Forecasting methods and applications. John Wiley & Sons.
- Murnane, R. J. (2013). U.S. high school graduation rates: Patterns and explanations. Journal of Economic Literature, 51(2), 370-422.
- Rumberger, R. W., & Losen, D. J. (2016). The high cost of harsh discipline and its disparate impact. Civil Rights Project-Proyecto Derechos Civiles.
- Tariq, M., Rashid, A., Aslam, M., & Hussain, S. (2018). Forecasting higher education enrollment in Pakistan using Box-Jenkins methodology. Journal of Applied Research in Higher Education, 10(3), 232-245.
- Tsay, R. S. (2005). Analysis of financial time series. John Wiley & Sons.
- UNESCO Institute for Statistics. (2023). Education data summary for Algeria. UIS.
- Unterhalter, E. (2019). The many meanings of quality education: Politics of targets and indicators in SDG 4. Global Policy, 10, 39-51.
- Wei, W. W. (2006). Time series analysis. In The Oxford handbook of quantitative methods in psychology: Vol. 2.