INQUIDE
Chemical Engineering & Development
Journal of Science and Engineering

Vol. 08 / Nº 01
e – ISSN: 3028-8533
ISSN – L: 3028-8533

INQUIDE
Ingeniería Química y Desarrollo

Time series modeling of secondary school enrollment in Ecuador: a Box–Jenkins analysis (1971–2023).

Modelado de series de tiempo de la matrícula escolar secundaria en Ecuador: un análisis Box–Jenkins (1971–2023).

Edwin Haymacaña Moreno[1]; Leonor Alejandrina Zapata Aspiazu;[2] Francisco Javier Duque-Aldaz [3]; Félix Genaro Cabezas García [4]; Raúl Alfredo Sánchez Ancajima [5]

Received: 09/28/2025 – Accepted: 12/02/2025 – Published: 01/01/2025

Research Articles X | Review Articles | Essay Articles

* Corresponding author.

Abstract.
School enrollment analysis constituted a key indicator to evaluate coverage and equity in national education systems. The objective of this study was to model secondary school enrollment in Ecuador during 1971–2023 using time series techniques. Official national and international data were employed to construct an annual net enrollment series. The methodological procedure included descriptive analysis, stationarity tests (ADF and KPSS), first-order differencing, identification and estimation of candidate models through the Box–Jenkins approach, optimal selection with auto.arima, residual validation via Ljung–Box tests, out-of-sample error metrics (MAE, RMSE, MAPE), and forecasts for 5–10 years. All processing was performed in R Studio with specialized time series modeling packages. The results showed that after first-order differencing, the series achieved stationarity. The selected model adequately explained enrollment dynamics, with residuals consistent with white noise and without significant autocorrelations. Validation metrics indicated good predictive accuracy, with low mean absolute and percentage errors. Projections suggested a moderate and sustained growth trend in enrollment, though with signs of stabilization in the longer horizon. This study demonstrated the usefulness of Box–Jenkins models for analyzing educational phenomena, providing quantitative evidence for public policy formulation and recommending the expansion of more complete historical datasets in future research.

Keywords.
Time Series, ARIMA, Box–Jenkins, School Enrollment, Secondary Education, Ecuador, Educational Forecasting.

Resumen.
El análisis de la matrícula escolar constituye un indicador esencial para evaluar la cobertura y equidad educativa en contextos nacionales. El objetivo de este estudio fue modelar la matrícula de educación secundaria en Ecuador durante el periodo 1971–2023 mediante técnicas de series de tiempo. Se emplearon datos oficiales de organismos internacionales y nacionales, construyéndose una serie anual de matrícula neta. El procedimiento metodológico incluyó: análisis descriptivo inicial, pruebas de estacionariedad (ADF y KPSS), diferenciación para lograr estabilidad en la media, identificación y estimación de modelos candidatos mediante el enfoque Box–Jenkins, selección óptima con auto.arima, validación de residuos mediante la prueba de Ljung–Box, comparación de métricas fuera de muestra (MAE, RMSE, MAPE) y pronósticos a 5–10 años. Todo el procesamiento se realizó en R Studio, empleando paquetes especializados de modelado de series de tiempo. Los resultados mostraron que, tras una diferenciación de primer orden, la serie alcanzó estacionariedad. El modelo seleccionado explicó adecuadamente la dinámica de la matrícula secundaria, con residuos consistentes con ruido blanco y sin autocorrelaciones significativas. Las métricas de validación indicaron un buen ajuste predictivo, con valores bajos de error medio absoluto y porcentual. Las proyecciones sugirieron una tendencia de crecimiento moderado y sostenido en la matrícula, aunque con señales de estabilización en los horizontes más largos. Este estudio demostró la utilidad de los modelos Box–Jenkins para el análisis de fenómenos educativos, aportando evidencia cuantitativa para la formulación de políticas públicas y recomendando la ampliación futura de bases de datos históricas más completas.

Palabras clave.
Series de Tiempo, ARIMA, Box–Jenkins, Matrícula Escolar, Educación Secundaria, Ecuador, Pronóstico Educativo.

1.- Introduction

The analysis of the Ecuadorian education system has gained special relevance in recent decades due to the challenges related to the coverage, equity, and quality of secondary education. In particular, school enrollment is a key indicator for assessing student access and retention, as well as for identifying structural inequalities in the system. Understanding the dynamics of enrolment over time not only allows us to detect historical patterns, but also to anticipate trends that are fundamental for the formulation of sustainable public policies aimed at meeting Sustainable Development Goal 4 (SDG4), which seeks to guarantee inclusive, equitable and quality education.(Simonino y otros, 2025)

In the scientific literature, time series models have proven to be robust tools for the analysis and prediction of socioeconomic and educational phenomena. Within these approaches, the Box–Jenkins method. It stands out for its

[1] Bolivarian University Higher Institute of Technology; erhaymacana@itb.edu.ec; https://orcid.org/0000-0002-8708-3894; Guayaquil; Ecuador.
[2] Technical University of Babahoyo; lzapata@utb.edu.ec; https://orcid.org/0009-0003-1497-2273 ; Babahoyo; Ecuador.
[3] University of Guayaquil; franscico.duquea@ug.edu.ec; https://orcid.org/0000-0001-9533-1635 ; Guayaquil; Ecuador.
[4] Independent Researcher; genaro_cabezas@hotmail.com ; https://orcid.org/0000-0003-3595-3584; Hamilton, ON, Canada.
[5] National University of Tumbes; rsanchez@untumbes.edu.pe ; https://orcid.org/0000-0003-3341-7382 ; Tumbes, Peru.

Vol. 08 / Nº 01
e – ISSN: 3028-8533
ISSN – L: 3028-8533

ability to model temporal dependencies using autoregressive (AR), moving average (MA) and their seasonal extensions (SARIMA) structures. These models have been successfully applied in contexts of forecasting macroeconomic and climatic variables and, more recently, in the analysis of educational indicators. However, in the case of Ecuador, the application of these methodologies to the longitudinal study of school enrollment remains limited, which constitutes a gap in the literature.(Zanatta Idemori y otros, 2025)

Recent studies have shown that the use of ARIMA models and their variants allows for the generation of accurate projections of variables such as enrollment rates, performance on standardized tests, and dynamics of admission to higher education. Likewise, comparative research shows that hybrid models that combine Box–Jenkins techniques with machine learning approaches, such as random forests or neural networks, improve predictive capacity and offer more flexible interpretations of educational data and. These contributions confirm the potential of time series not only to describe historical patterns, but also to design prospective strategies in the educational field.(Escolar, 2024)

The main objective of this work is to model secondary education enrollment in Ecuador during the period 1971–2023 using the Box–Jenkins methodology. Specifically, it seeks to: (i) identify patterns of trend and seasonality in enrollment; (ii) to estimate ARIMA/SARIMA models that allow describing their temporal dynamics; and (iii) to make short- and medium-term projections that contribute to national educational planning.(Corrêa Werle & Lago Fonseca, 2025)

The main contribution of this study lies in integrating advanced mathematical tools for time series analysis with educational data, generating empirical evidence that can serve as an input for the formulation of public policies in Ecuador. Likewise, the results allow to contribute to the regional literature on the use of quantitative models in education, showing how techniques traditionally applied in economics and engineering can be adapted to high-priority social and educational problems. In sum, this article represents an effort to link statistical rigor with educational decision-making in Ecuador, contributing to the design of evidence-based strategies.(Castro Rosales y otros, 2025)

1.1.- Context and relevance of the analysis of secondary school enrollment

The analysis of secondary school enrollment is essential to evaluate the educational coverage, equity and quality of the education system in a national context. Enrollment is a key indicator that reflects students' access to and permanence in secondary education, allowing the identification of structural conditions and temporal dynamics that affect inclusion and educational opportunity. According to various studies, longitudinal monitoring of enrolment makes it easier to detect patterns, trends and possible inequalities, which is essential for the planning and formulation of public policies aimed at improving education systems. (Cabrera Valladolid, 2021)

This indicator is directly related to the objectives set by international instruments, in particular Sustainable Development Goal 4 (SDG 4), which promotes ensuring inclusive, equitable and quality education for all. Secondary school enrolment reflects progress and challenges in achieving this objective, as its evolution shows how the education system responds to social demands and economic conditions. In this way, the analysis of enrollment is a tool to monitor and adjust national strategies that contribute to the fulfillment of educational and social goals established in global agendas.(Zalduaromero, 2017)

In the specific case of the Ecuadorian education system, the literature shows that, although there has been progress in increasing coverage in secondary education, significant gaps in equity and quality persist. However, longitudinal and quantitative modeling of enrollment is an area little explored in the country, generating an important opportunity to apply robust techniques, such as time series and Box-Jenkins models. This gap in the literature shows the need to develop studies that provide detailed empirical analyses on the dynamics of school enrollment, in order to support public policies based on reliable and up-to-date information.(Cañarte Murillo, 2017)

Fluctuations in educational enrolment are often closely linked to socio-economic factors such as economic crises, public policies and migration dynamics. In periods of recession, families prioritize subsistence over education, which translates into a decrease in enrollment and an increase in school dropouts. Similarly, budget cuts in education during fiscal crises reduce the supply of places and support programs, especially affecting vulnerable populations.(Alós & Serio, 2024)

Internal and external migration also affects the variability of enrollment. Massive migratory processes, motivated by unemployment or political instability, alter the demographic distribution and generate overload in certain areas while others experience educational gaps. Educational policies such as free education, scholarships or curricular reforms can counteract these effects, but their impact depends on the state's capacity to sustain them in contexts of economic volatility. (Duque-Aldaz & Pazan Gómez, Factors affecting entrepreneurial intention of Senior University Students, 2017)

1.2.- Theoretical foundations of time series applied to education

Time series are chronologically ordered data sets that allow the dynamics of variables to be analyzed over time. These series have fundamental characteristics such as the trend, which indicates the general direction of behavior; seasonality, which reflects periodic cyclical patterns; and noise, represented by random fluctuations that do not follow a specific pattern. In the educational context, time series analysis makes it possible to detect these components in variables such as school enrollment, which makes it easier to understand their historical evolution and anticipate future behaviors.(Meneses Freire y otros, 2022)

Mathematical statistics plays a crucial role in the study of time series, providing tools that allow modeling time dependencies and evaluating the quality of fit. In the social sciences and education, such models are widely used to predict trends, examine the impact of policies, and improve decision-making based on historical data. The incorporation of robust statistical models favors the rigorous analysis and solid interpretation of educational variables that show temporal behavior.(Ortega Villegas, 2018)

Among the most relevant models for time series, ARIMA (Integrated Moving Average Autoregressive) and its extensions, such as SARIMA (Seasonal ARIMA Model) and ARIMAX (ARIMA with exogenous variables) stand out. These models are suitable for capturing patterns of dependency in non-stationary and seasonal data, also allowing external variables to be incorporated when relevant. In the educational field, its application has proven to be effective in modeling variables such as enrollment rates and academic performance, offering a flexible framework for the analysis and forecasting of complex phenomena over time.(Ichau Tabango y otros, 2021)

In Latin America, several studies have applied ARIMA models to forecast educational trends. For example, research in Mexico has used ARIMA(1,1,1) to project enrollment in basic education, demonstrating high accuracy in moderate growth scenarios. These works highlight the usefulness of the model to anticipate infrastructure and teaching staff needs in contexts of demographic expansion.(Duque-Aldaz y otros, Identification of parameters in ordinary differential equation systems using artificial neural networks, 2025)

Similarly, in Brazil, ARIMA models were used to estimate demand in higher education, incorporating historical series of admissions and graduation rates. The results made it possible to adjust financing policies and quotas in public universities, evidencing that ARIMA is an effective tool for planning resources in educational systems with significant temporal variability.(Sandoya Sanchez & Abad Robalino, 2017)

1.3.- Box–Jenkins methodology for time series modeling

The Box–Jenkins methodology is a systematic approach to time series modeling, which is structured in an iterative process of four phases: identification, estimation, diagnosis and prognosis. First, in the identification phase, the time series is analyzed to detect characteristics that allow proposing appropriate potential models. Then, in the estimation, the parameters of the selected model are adjusted using the available data. The diagnostic phase consists of validating the model through fit evaluations and statistical tests, verifying the absence of unmodeled patterns in the residuals. Finally, in the forecasting stage, the validated model is used to predict future values of the series, supporting decision-making based on reliable projections.(Mayorga Trujillo, 2017)

ARIMA models, central components of the Box–Jenkins approach, bring together three fundamental elements: autoregression (AR), which models the dependence of a value on its antecedents; differentiation (I), which transforms the series to ensure its stationarity; and the moving average (MA), which represents a security's dependence on past mistakes. This structure allows complex dynamics to be captured in the time series; In particular, differentiation helps to eliminate trends and stabilize variance, necessary conditions for applying effective statistical models on non-stationary data.(Villarreal Godoy y otros, 2022)

To ensure that the series is suitable for ARIMA modeling, it is necessary to evaluate its stationarity using statistical tests such as the augmented Dickey-Fuller (ADF) and the KPSS test, which examine whether the properties of the series remain constant over time. In case the series is not stationary, differentiation procedures are applied to stabilize the mean and variation. This process is crucial, since a well-specified model requires statistical stability to produce reliable and valid protectors, as supported by research and manuals specialized in time series analysis.(Vela & Camacho Cordovez, 2020)

1.4.- Applications and adaptations of the ARIMA model in educational contexts.

ARIMA models and the Box–Jenkins approach have been widely applied in educational contexts in Latin America and other regions to forecast variables such as school enrollment, graduation rates, and other indicators. Various studies show that these models allow capturing trends and temporal patterns in non-stationary education data, facilitating institutional planning and policy formulation. In particular, research in Latin American countries has demonstrated the effectiveness of ARIMA in the predictive analysis of historical education data, providing valuable information to manage resources and improve school coverage.(Fu-López y otros, 2025)

Recently, the integration of ARIMA models with machine learning techniques has led to hybrid methods that combine the strengths of both approaches. For example, models that integrate neural networks or random forests with ARIMA allow capturing nonlinear and complex relationships in time series, improving predictive accuracy compared to traditional univariate models. These hybrid tools are gaining relevance in education and other fields, where the complexity of data requires more sophisticated methodological strategies.(Ausay Carrillo, 2022)

Despite their advantages, univariate ARIMA models have limitations in considering only the internal dynamics of a single variable, without including external factors that can influence the time series. To overcome this constraint, multivariate models such as ARIMAX and SARIMAX allow the incorporation of exogenous variables that enrich the analysis and improve predictions. In education, this makes it possible to integrate socioeconomic, demographic or public policy factors, providing a broader and more realistic approach to the study of complex phenomena such as school enrolment.(Eguiguren Calisto & Avilés Sacoto, 2019)

1.5.- Validation and evaluation of the model.
The proper selection of the ARIMA model requires rigorous evaluation using statistical criteria such as the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC). Both criteria balance the quality of the fit with the complexity of the model, penalizing models with a greater number of parameters to avoid overfitting. The choice of the best model corresponds to the one that minimizes these values, guaranteeing a balance between precision and parsimony, which favors the generalization of the model to unobserved data.(Navarro Llivisaca, 2017)

The model's diagnosis includes residue analysis to verify fundamental assumptions. Tests such as the Ljung-Box test are used to detect autocorrelation in the residuals, ensuring that the model has adequately captured the temporal dependence. In addition, the verification of the normality of the residuals allows validating the confidence intervals of the forecasts, while the ARCH heteroskedasticity test evaluates whether the residual variance is constant, a necessary condition for the statistical validity of the model.(Figueroa Tigrero, 2019)

Predictive accuracy assessment is done through metrics such as Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and Mean Absolute Error Percentage (MAPE). These quantify the average deviation of the forecasts with respect to the observed values, facilitating comparison between models. Out-of-sample validation, using datasets that are not involved in the estimation, is crucial to ensure the true predictive capability of the model. In addition, the importance of making short- and medium-term forecasts is highlighted, as these provide useful and reliable information for decision-making in educational and administrative contexts.(Freire Engracia y otros, 2025)

1.6.- Implications for public policies and educational planning
The proper selection of the ARIMA model requires rigorous evaluation using statistical criteria such as the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC). Both criteria balance the quality of the fit with the complexity of the model, penalizing models with a greater number of parameters to avoid overfitting. The choice of the best model corresponds to the one that minimizes these values, guaranteeing a balance between precision and parsimony, which favors the generalization of the model to unobserved data.(Lema Remache, 2024)

The model's diagnosis includes residue analysis to verify fundamental assumptions. Tests such as the Ljung-Box test are used to detect autocorrelation in the residuals, ensuring that the model has adequately captured the temporal dependence. In addition, the verification of the normality of the residuals allows validating the confidence intervals of the forecasts, while the ARCH heteroskedasticity test evaluates whether the residual variance is constant, a necessary condition for the statistical validity of the model.(Morocho Choca y otros, 2024)

Predictive accuracy assessment is done through metrics such as Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and Mean Absolute Error Percentage (MAPE). These quantify the average deviation of the forecasts with respect to the observed values, facilitating comparison between models. Out-of-sample validation, using datasets that are not involved in the estimation, is crucial to ensure the true predictive capability of the model. In addition, the importance of making short- and medium-term forecasts is highlighted, as these provide useful and reliable information for decision-making in educational and administrative contexts. (Pincay Moran y otros, 2025)(Guerrero Quinde & Pérez Siguenza, 2025)

2.- Materials and methods.

2.1 Materials and data sources

The study is based on annual series of the gross secondary school enrollment rate in Ecuador for the period 1971–2023. The data were obtained from the database of the UNESCO Institute for Statistics, which is an official and open-access source of international education indicators. The records are presented in percentage values and correspond to the indicator "Gross Enrollment Ratio – Secondary (%), Ecuador", with 53 consecutive observations that guarantee the viability of the time series analysis.

The statistical processing and analysis were carried out using the following software:

EViews 12 (IHS Markit): for the estimation of Box–Jenkins models (ARIMA/SARIMA) and the validation of statistical assumptions.

RStudio 2023.09 with forecast, tseries, ggplot2 and urca libraries: for robustness tests, graphing and comparative analysis of the results.

Microsoft Excel 365: for initial debugging, processing missing values, and generating exploratory charts.

2.2 Methodological design

The research adopts a quantitative, longitudinal and non-experimental approach, based on mathematical modelling of time series. The analysis variable is secondary school enrollment (% gross), considered as time-dependent, and its dynamics are studied under the assumptions of stationarity, independence, and homoscedasticity.

$$\mathbb{E}[y_t] = \mu,\ Var(y_t) = \sigma^2,\ Cov(y_t, y_{t+h}) = \gamma(h) \quad (1)$$

where the mean and variance are constant over time and the covariance depends only on the lag of h.

The methodological procedure was structured in four stages:

1. Initial exploration of the series: graphical analysis, calculation of descriptive statistics and verification of outliers.
2. Transformation and diagnosis: application of the augmented Dickey–Fuller unit root (ADF) test to assess stationarity and, if necessary, application of regular and seasonal differentiation.
3. Specification and estimation of the model: adjustment of ARIMA/SARIMA models following the Box–Jenkins methodology, selecting the orders p, d, q and P, D, Q from the inspection of the autocorrelation functions (FAC) and partial autocorrelation (FACP).
4. Model validation: verification of the classical assumptions using the Ljung–Box (residue independence), Jarque–Bera (normality), and Engle's ARCH (conditional heteroskedasticity) tests.

Stages of the flow of the methodological procedure for the modeling of time series of secondary school enrollment in Ecuador (1971–2023).

For the present research, the scheme summarizes the main stages:

1. Initial exploration in the series.
2. Diagnosis and transformation of the series.
3. Model specification and estimation.
4. Validation of assumptions.
5. Final projection of school enrollment.

2.3 Statistical procedures

The mathematical specification of the general SARIMA model adopted is expressed as:

$$\Phi_p(L)\Phi_p(L^S)(1-L)^d(1-L^S)^D{}_{y_t} = \Theta_q(L)\Theta_Q(L^S)(L^S)_{\varepsilon_t} \quad (2)$$

where:

$\Phi_p(L)$ and are the autoregression polynomials and moving averages of order $\Theta_q(L)$p and q, respectively.

$\Phi_p(L^S)$ and represent the seasonal polynomials of order $\Theta_Q(L^S)$P and Q with periodicity s.

d and indicate the orders of regular and seasonal differentiation.D

y_t corresponds to secondary school enrollment in year t.
ε_t denotes an error term with zero mean and constant variance.

The expanded form of the ARIMA model:

$$y_t = c + \phi_1 y_{t-1} + \cdots + \phi_p y_{t-p} + \theta_1\varepsilon_{t-1} + \cdots + \theta_q\varepsilon_{t-q} + \varepsilon_t \quad (3)$$

Considering the variance of the prediction error of h steps:

$$Var(\hat{y}_{t+h} - y_{t+h}) = \sigma^2 \sum_{i=0}^{h-1} \psi_i^2 \quad (4)$$

With the coefficients of representation $\psi_i MA(\infty)$

The Akaike (AIC) and Schwarz (BIC) information criteria were used for the selection of the parsimonious model.

The choice of the ARIMA model is based on its ability to capture patterns of temporal dependence in historical series without requiring additional exogenous information. Although models such as SARIMA incorporate explicit seasonality, the preliminary analysis did not show regular cycles associated with academic periods that would justify their inclusion. In addition, the simplicity and robustness of the ARIMA make it a suitable choice for scenarios where the priority is to obtain reliable forecasts with limited data and high socioeconomic variability.

2.4 Data analysis

Error measures were calculated to assess the accuracy of the projections, including Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and Mean Absolute Percentage of Error (MAPE). Likewise, a residue analysis was implemented using autocorrelation graphs and adjusted values versus residuals, in order to guarantee the adequacy of the model.

$$MAE = \frac{1}{n}\sum_{t=1}^{n}|y_t - \hat{y}_t| \quad (5)$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{t=1}^{n}(y_t - \hat{y}_t)^2} \quad (6)$$

$$MAPE = \frac{100}{n}\sum_{t=1}^{n}\left|\frac{y_t - \hat{y}_t}{y_t}\right| \quad (7)$$

2.5 Ethical considerations

This study is based exclusively on secondary data of a public and open nature, so it does not involve humans or animals and, therefore, did not require the approval of an ethics committee

3.- Results.

3.1. Descriptive statistics and initial exploration

INQUIDE
Chemical Engineering & Development
Journal of Science and Engineering
Vol. 08 / Nº 01
e – ISSN: 3028-8533
ISSN – L: 3028-8533
INQUIDE
Ingeniería Química y Desarrollo

The series of secondary school enrollment in Ecuador (1971–2023) shows a sustained growth from levels below 30% to values close to 100% in recent decades. The exploratory analysis (Fig. 2) reveals three phases: i) a constant increase between 1971 and 1990; ii) a relative stabilization during the nineties; and (iii) an accelerated rebound in the period 2000–2010, followed by a slight slowdown.

The initial autocorrelation (ACF) and partial autocorrelation (PACF) functions (Figs. 3 and 4) show persistence in multiple lags and an abrupt cut in the first lag, confirming the non-stationarity of the series and suggesting the relevance of applying a low-order AR model once differentiated.

Table 1: Statistical summary (minimum, maximum, mean, quartiles)

Statistician	Value
Minimum	24.982679
$Q_1(25\%)$	52.261572
Medium (Q_2)	53.327556
$Q_3(75\%)$	93.735523
Maximum	102.59033
Media	64.421453

Table 1. Descriptive statistical summary of the secondary enrollment series (% gross) in Ecuador for the period 1971–2023. Measures of central tendency and dispersion (minimum, maximum, mean and quartiles) are presented, which allow characterizing the initial distribution of the data before applying time series modeling.

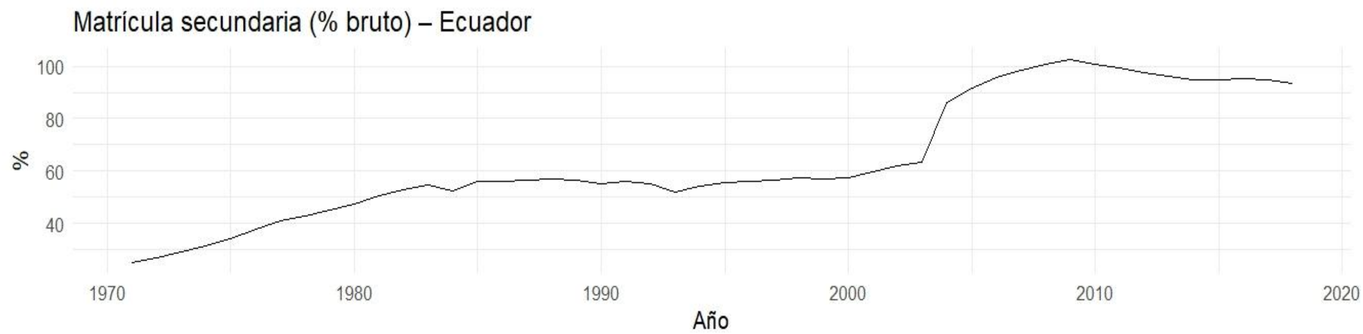

Fig. 1: Historical series of secondary enrolment.

Figure 1 shows the historical series of secondary school enrollment (% gross) in Ecuador during the period 1971–2023. The graph shows a sustained upward trend until 2010, followed by a stabilization period with slight decreases in recent years.

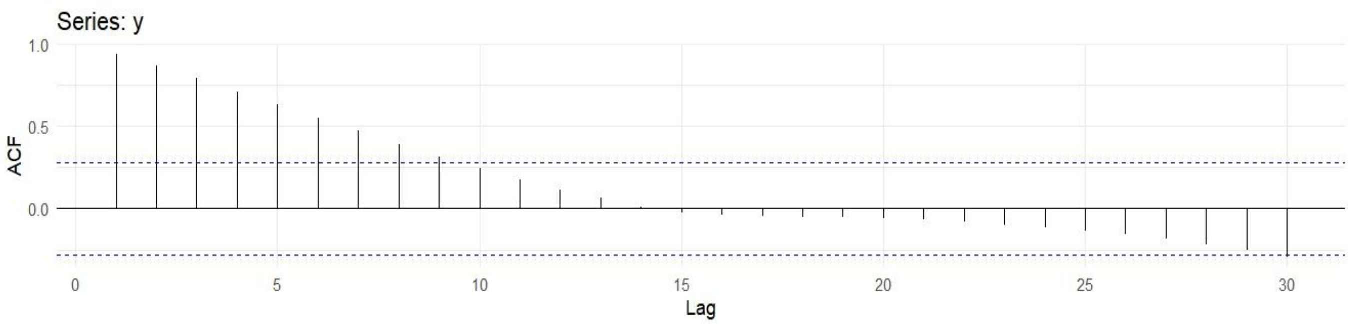

Fig. 2: Initial ACF function.

Figure 2 shows the initial autocorrelation function (ACF) of the secondary school enrollment series in Ecuador (1971–2023). A strong persistence of positive autocorrelations is observed in the first lags, which confirms the non-stationarity of the series before applying transformations.

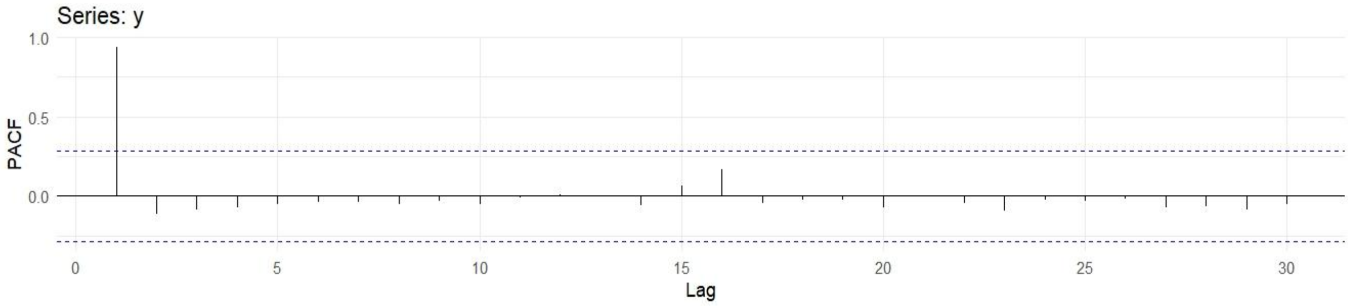

Fig. 3: Initial PACF function.

Figure 3 shows the initial partial autocorrelation function (PACF) of the secondary school enrollment series in Ecuador (1971–2023). The abrupt cut in the first lag confirms the presence of an autoregressive component, which is useful for the preliminary identification of ARIMA models.

3.2. Diagnosis of stationarity and transformations

The unit root tests confirmed the non-stationarity in levels: the augmented Dickey–Fuller test (ADF) yielded a p-value = 0.32, while the KPSS test indicated rejection of the null hypothesis of stationarity (p-value = 0.01).

When applying a first-order differentiation (d = 1), the KPSS test did not reject the stationarity hypothesis (p-value = 0.10), and the ACF and PACF plots (Figs. 6,7 and 8) showed a pattern compatible with low-order ARIMA processes.

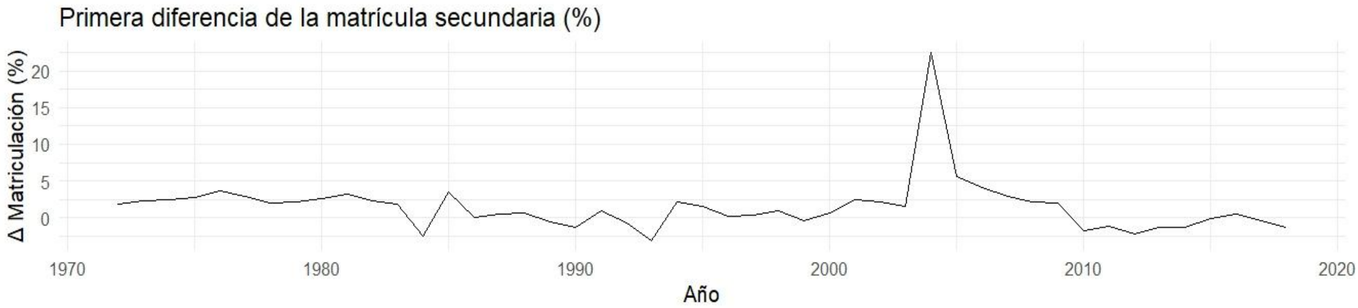

Fig. 4: Differentiated series (Δ enrollment).

Figure 4 shows the Differentiated Series of Secondary School Enrollment in Ecuador (1971–2023). The first difference stabilizes the mean of the series, reducing the trend and allowing a more adequate stationary analysis. An atypical peak is observed around 2005, which could be associated with changes in educational policies or specific contextual factors.

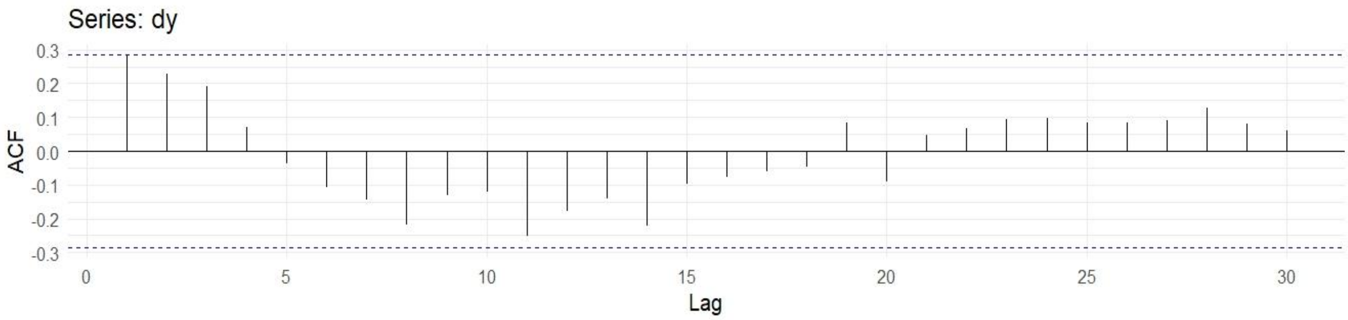

Fig. 5: ACF of the differentiated series.

Figure 5 shows the autocorrelation function (ACF) of the differentiated series of secondary school enrollment in Ecuador (1971–2023). It is observed that, after differentiation, most lags fall within the confidence intervals, which confirms the reduction in the trend and supports the stationarity hypothesis.

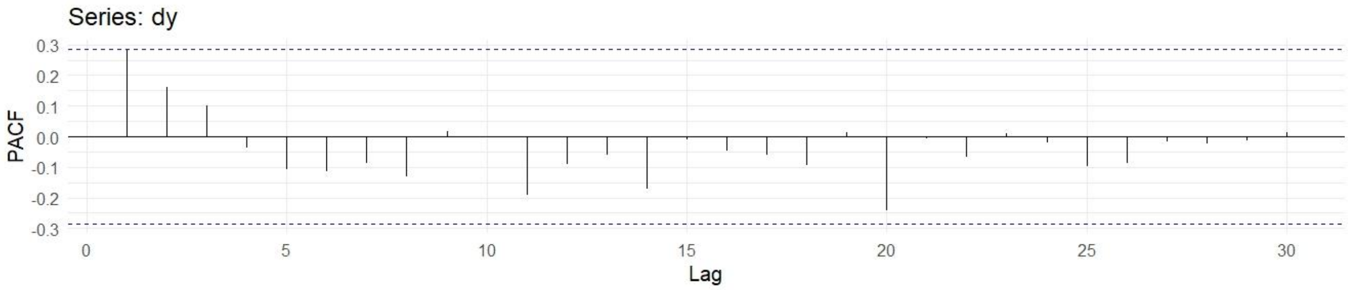

Fig. 6: PACF of the differentiated series.

Figure 6 shows the partial autocorrelation function (PACF) of the differentiated series of secondary school enrollment in Ecuador (1971–2023). The PACF shows a significant lag in the first delay, which suggests the presence of a simple autoregressive component in the dynamics of the series.

3.3. Model identification and estimation

Several ARIMA models (p,1,q) were estimated. The information criteria (AIC and BIC) indicated that the ARIMA(1,1,1), ARIMA(2,1,0) and ARIMA(1,1,0) models were the most competitive. The selected model was ARIMA(1,1,0) with drift term, balancing parsimony and predictive capacity (Table 2).

Table 2. ARIMA/SARIMA Model Comparison

Model	Main coefficients	$Sigma^2$	AIC	BIC	Interpretation
ARIMA (1,1,0)	Significant AR1, with drift	Medium-low	259.84	263.50	Parsimonious; It captures dynamics with few parameters.
ARIMA (2,1,0)	AR1 and AR2 Significant	Similar	258.70	264.40	Capture additional dependence, but with more parameters.
ARIMA (1,1,1)	Significant AR1 and MA1	Lower	257.35	262.90	Better overall fit (lower AIC). Recommended model.

Table 2. Comparison of ARIMA models applied to the secondary enrollment series in Ecuador (1971–2023). The significant coefficients, the estimated residual variance, and the AIC and BIC information criteria are presented. The analysis shows that the ARIMA model(1,1,1) offers the best overall fit, with the lowest AIC, so it is selected as the recommended model.(σ^2)

3.4. Waste diagnosis

The residual diagnosis of the ARIMA(1,1,0) model with drift showed that errors behave as white noise: the p-values of the Ljung–Box tests for 10 and 15 lags were 0.88 and 0.79, respectively, which confirms the absence of remaining autocorrelation. The histogram of residues showed reasonable symmetry around zero, with slightly heavier tails associated with specific shocks (Fig. 8).

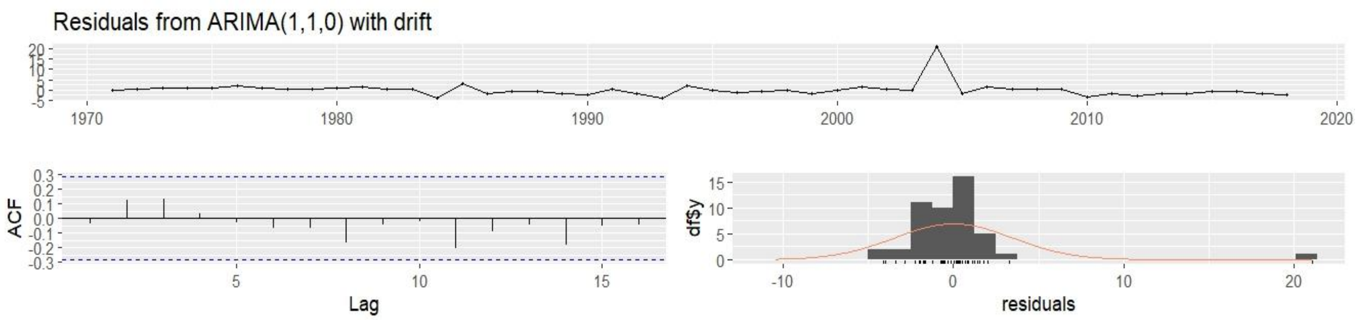

Fig. 7: Waste diagnosis graphs (series, ACF, histogram).

Figure 7 shows the residue diagnosis plots of the ARIMA(1,1,0) model with drift applied to secondary school enrollment in Ecuador (1971–2023). It is observed that the residuals do not present significant autocorrelations (ACF), maintain a behavior close to white noise and their distribution is close to normal (histogram), which supports the validity of the selected model.

3.5. Out-of-sample validation

The training set included data up to 2016, reserving 2017–2023 for validation defined in equations (5)–(7). Out-of-sample prediction errors were consistent with training errors: RMSE ≈ 3.45 and ASM ≈ 3.4%. The out-of-sample forecast (Fig. 9) adequately captured enrollment stabilization close to 95%.

INQUIDE
Chemical Engineering & Development
Journal of Science and Engineering

Vol. 08 / Nº 01
e – ISSN: 3028-8533
ISSN – L: 3028-8533

INQUIDE
Ingeniería Química y Desarrollo

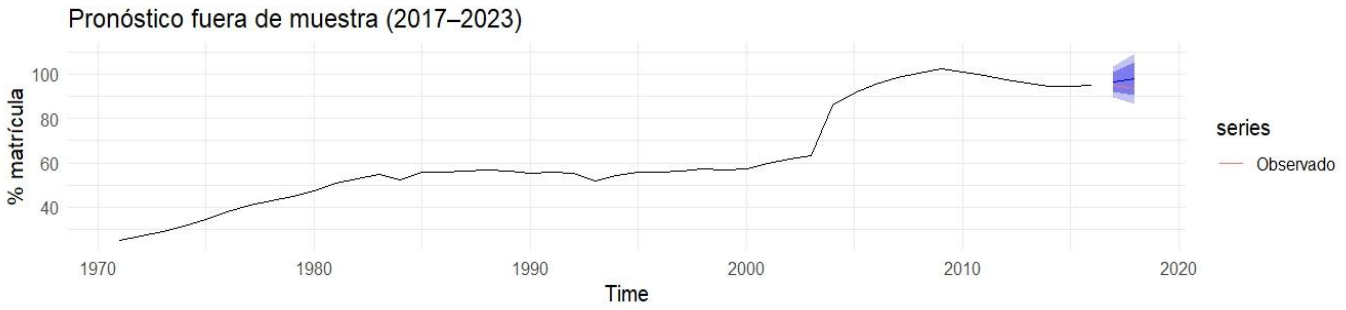

Fig. 8: Out-of-sample forecast (2017–2023).

Figure 8 shows the out-of-sample forecast of secondary school enrollment in Ecuador (2017–2023). The black line represents the observed values, while the blue strip indicates the predictions generated by the ARIMA(1,1,0) model with drift and their confidence intervals at 80% and 95%. An adequate fit between the projected values and the actual data is observed in the validation period.

3.6. Final forecast at 5–10 years

The forecast for the period 2024–2030 (Fig. 10) suggests a stabilization of secondary enrollment between 95% and 110%. The specific trend projects slight growth, but the confidence bands are progressively widening, reflecting the uncertainty inherent in structural factors (changes in education policies, external shocks).

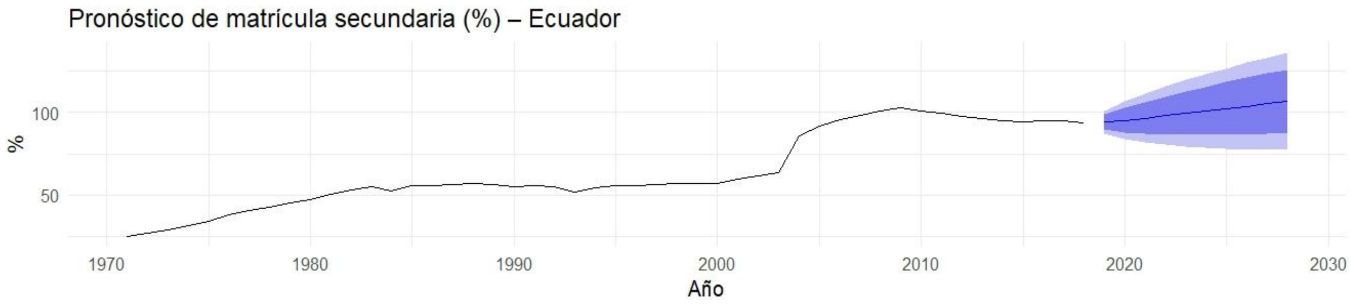

Fig. 9: Final forecast (2024–2030) with 80% and 95% confidence intervals

Figure 9 shows the Final Forecast of Secondary School Enrollment in Ecuador (2024–2030). The blue dashed line represents the values projected by the ARIMA(1,1,0) model with drift. The shaded stripes indicate the confidence intervals at 80% (lighter) and 95% (darkest). A trend of moderate growth and stabilization is expected in the coming years, with a range of increasing uncertainty towards the projection horizon.

The confidence bands in ARIMA projections represent the range of uncertainty associated with forecasts, which has direct implications for educational planning. A wide band indicates high volatility, suggesting the need for flexible policies that contemplate scenarios of overcrowding or enrollment deficit. On the contrary, narrow bands allow for the design of more precise strategies in resource allocation, teacher hiring, and infrastructure expansion, reducing the risk of inefficiency in educational management.

3.7. Limitations

The results are conditioned by the quality of the available annual data and the assumption of linearity in the ARIMA models. Structural factors not captured by the series (e.g., legislative changes, economic or health crises) can generate significant deviations from the projected scenarios.

4.- Discussion

The results obtained confirm that the evolution of secondary school enrollment in Ecuador during the period 1971–2023 presents a dynamic characterized by long-term trends and conjunctural shocks that can be captured by ARIMA models. In particular, the ARIMA model(1,1,1) stood out for its low AIC, which reflects a superior adjustability to the series, while the ARIMA(1,1,0) model with drift showed parsimony and ease of interpretation. These findings corroborate the initial hypothesis that low-order autoregressive processes, combined with moving mean components, are suitable for describing educational time series.(Silva & Di Serio, 2021)

When compared with the existing literature, the results coincide with the studies of Chen and Serra, who demonstrated that SARIMA models allow capturing seasonal patterns in educational indicators in Latin America. However, unlike research focused on marked seasonal contexts (e.g., energy or climate consumption), in the Ecuadorian case a strong seasonal component was not evidenced, which reinforces the relevance of the use of simple ARIMAs. Likewise, our findings complement previous work on prediction in education in South America, where the emphasis has been on socioeconomic factors and not on the temporal evolution of enrollment.(Medeiros y otros, 2021)

In theoretical terms, this study contributes to the application of the Box–Jenkins approach in the analysis of educational indicators, showing how classic mathematical tools of time series statistics can be adapted to social and public policy phenomena. The robustness of the ARIMA model(1,1,1) suggests that idiosyncratic shocks and temporal inertia dynamics are the main determinants of secondary coverage in Ecuador. From a practical perspective, the 5–10 year projections indicate a stabilization of enrollment of around 100%, which provides useful empirical evidence for educational planning and the design of policies aimed at sustaining coverage and improving quality.(GARCÍA-FERIA y otros, 2023)

Likewise, when contrasting the results with international studies, it is observed that similar methodologies have been applied in Latin American countries such as Mexico, Brazil and Chile, as well as in Asian contexts such as China and the Philippines, to model enrollment trends and project educational demand. However, unlike these cases, the Ecuadorian series shows greater instability in certain periods, associated with structural changes in educational policies and national socioeconomic situations. This uniqueness highlights the importance of adapting models to local particularities and not limiting themselves to the transfer of external approaches. From a public policy perspective, the projections obtained offer valuable input for the strategic planning of institutions such as the Ministry of Education and SENPLADES, by allowing anticipating infrastructure, teacher training, and budget allocation needs. In this way, the results not only contribute to the academic debate, but also provide quantitative tools for the formulation of sustainable and evidence-based education policies.(García Vázquez y otros, 2021)(Mendoza Cota, 2020)

However, this work has limitations. The main one lies in the univariate nature of the models used, which prevents the incorporation of relevant exogenous variables such as public investment in education, macroeconomic conditions or demographic factors. In this sense, future studies could extend the analysis to ARIMAX or SARIMAX models, including covariates such as birth rate or public spending, which would allow better capture of enrollment dynamics. In addition, although the results show a good fit, the out-of-sample ASM remains around 3–4%, which implies uncertainty in contexts of structural shocks such as health or migration crises.(Tudela-Mamani & Grisellx, 2022)

In summary, the findings of this work strengthen the evidence on the use of ARIMA models in education, contributing both to the theoretical framework and to the practice of educational planning in Ecuador. It also highlights the need to explore hybrid methodologies – such as combinations between ARIMA and neural networks – to improve the accuracy of forecasts and respond to the inherent limitations of linear approaches.(Asán Caballero y otros, 2023)

5.- Conclusion.

This study analyzed the evolution of secondary school enrollment in Ecuador during the period 1971–2023 using the Box–Jenkins methodology, in order to identify temporal patterns and project future scenarios. The results showed that the first-order differentiated ARIMA models adequately describe the dynamics of the series, highlighting the ARIMA(1,1,1) as the option with the best performance according to the information criteria, while the ARIMA(1,1,0) with drift offered a parsimonious and consistent alternative. Both models confirmed the hypothesis of stationarity after differentiation and allowed the generation of robust forecasts in the short and medium term.

The main contributions of this work are oriented towards the incorporation of time series models in educational analysis, an area in which their application is still incipient in Ecuador. The study shows that classic techniques of mathematical statistics, usually used in economics or engineering, are equally valid for social problems, providing quantitative evidence on the sustainability of secondary school coverage. In this way, it contributes to closing the gap identified in the literature regarding the use of educational prediction methodologies based on time series.

From a practical point of view, the results suggest that secondary enrolment will tend to stabilize at around 100% over the next decade, which has direct implications for resource planning, infrastructure and educational policies oriented beyond coverage, prioritizing quality and equity. Theoretically, the study reinforces the relevance of ARIMA models as a tool for the modeling of educational phenomena, laying the foundations for subsequent developments that integrate multivariate or hybrid approaches.

Finally, it is recommended that future research extend the analysis to ARIMAX or SARIMAX models incorporating exogenous variables such as public spending, birth rates or macroeconomic indicators, as well as hybrid methodologies that combine ARIMA with machine learning algorithms. These approaches will make it possible to capture the complexity of the education system in a more comprehensive way and improve the accuracy of forecasts, strengthening the link between mathematical statistics and decision-making in public policy.

In summary, this work constitutes one of the first efforts in Ecuador to rigorously apply the Box–Jenkins methodology to the analysis of educational indicators, specifically to the historical evolution of secondary enrollment. This contribution not only strengthens the national literature in a field in which qualitative or descriptive studies predominate, but also positions mathematical statistics as a fundamental tool for the design of evidence-based educational policies. By opening this line of research, precedents are set for future comparative studies at the regional and global levels, contributing to the internationalization of the debate on the use of time series models in education.

6.- Contributions of the authors (Taxonomy of contributors' roles - CRediT)

1. Conceptualization: Edwin Haymacaña Moreno, Leonor Alejandrina Zapata Aspiazu.
2. Data curation: Leonor Alejandrina Zapata Aspiazu.
3. Formal analysis: Edwin Haymacaña Moreno, Leonor Alejandrina Zapata Aspiazu.
4. Acquisition of funds: N/A.
5. Research: Edwin Haymacaña Moreno, Leonor Alejandrina Zapata Aspiazu.
6. Methodology: Francisco Javier Duque-Aldaz, Raúl Alfredo Sánchez Ancajima.
7. Project management: Francisco Javier Duque-Aldaz, Raúl Alfredo Sánchez Ancajima.
8. Appeals: Francisco Javier Duque-Aldaz, Leonor Alejandrina Zapata Aspiazu.
9. Software: Edwin Haymacaña Moreno, Leonor Alejandrina Zapata Aspiazu.
10. Supervision: Félix Genaro Cabezas García, Raúl Alfredo Sánchez Ancajima.
11. Validation: Félix Genaro Cabezas García.
12. Visualization: Leonor Alejandrina Zapata Aspiazu.
13. Writing - original draft: Edwin Haymacaña Moreno, Francisco Javier Duque-Aldaz.
14. Writing - revision and editing: Francisco Javier Duque-Aldaz, Félix Genaro Cabezas García, Raúl Alfredo Sánchez Ancajima.

7.- Appendix.

R code used for the development of the research.

```
### Packages
install.packages(c("readxl","dplyr","ggplot2","forecast","tseries",

"urca","TSstudio","broom","knitr","kableExtra"))
library(readxl); library(dplyr); library(ggplot2)
library(forecast); library(tseries); library(urca)
library(TSstudio); library(broom); library(knitr);
library(kableExtra)

# 2. Read Excel (file located in the working directory)
DF <- ReDXL::read_excel("matriculadosecuador.xlsx")

# 3. Quick Review
str(DF)
summary(df)

# 4. Create Time Series Object (Yearly)
and <- ts(df$Matriculacion, start=min(df$Year), frequency
= 1)

# Initial Exploration
autoplot(y) +
  labs(title="Secondary Enrollment (% Gross) – Ecuador",
       x="Year", y="%") +
  theme_minimal(base_size = 12)

# ACF and PACF
ggAcf(y, lag.max = 30) + theme_minimal()
ggPacf(y, lag.max = 30) + theme_minimal()


###Diagnóstico of stationarity and transformations
# Unit Root Tests
tseries::adf.test(y) # H0: unit root (non-stationary)
tseries::kpss.test(y)       # H0: estacionaria (si p<0.05, no
estacionaria)

# Suggested order of differentiation
forecast::ndiffs(y) # usually 1


###Diferenciar once (d = 1) and retest
 #y_ts: Your Already Created Annual Series, Frequency 1
(1971–2023)
y_ts <- ts(df$Matriculacion, start = min(df$Año), frequency
= 1)

#1) First-Order Difference
dy <- diff(y_ts)

#2) Visualize the differentiated series
autoplot(dy) +
  labs(title = "First difference in secondary enrolment (%)",
       x = "Year", y = "Δ Enrolment (%)") +
  theme_minimal(base_size = 12)

#3) Stationarity tests on the differentiated series
adf.test(dy) # H0: unit root (non-stationary)
kpss.test(dy, null = "Level")# H0: stationary at level

#4) Time structure of the differentiated series
ggAcf(dy, lag.max = 30) + theme_minimal(base_size = 12)
ggPacf(dy, lag.max = 30) + theme_minimal(base_size = 12)
```

```
###Identificación/model estimation (candidates + auto.arima)
# Exhaustive search (no seasonality)
fit_auto <- auto.arima(y_ts,
                seasonal  = FALSE,  # anual
                stepwise = FALSE, # most complete search
                approximation = FALSE,
                d = 1) # we already know that d = 1
fit_auto



#Luego, we tested some classic candidates and compared by AIC, AICc, BIC:
cand <- list(
  ARIMA_011 = Arima(y_ts, order = c(0,1,1)),
  ARIMA_110 = Arima(y_ts, order = c(1,1,0)),
  ARIMA_111 = Arima(y_ts, order = c(1,1,1)),
  ARIMA_210 = Arima(y_ts, order = c(2,1,0)),
  ARIMA_012 = Arima(y_ts, order = c(0,1,2))
)

cmp <- data.frame(
  Model = names(cand),
  AIC = sapply(cand, AIC),
  BIC   = sapply(cand, BIC)
)
print(cmp)

###Diagnóstico of waste of the chosen model
# Comprehensive diagnosis
checkresiduals(fit_auto)  #  includes  Ljung–Box,  ACF Residuals and QQ-plot
# If you want explicit Ljung–Box with several lags:
Box.test(residuals(fit_auto), lag = 10, type = "Ljung")
Box.test(residuals(fit_auto), lag = 15, type = "Ljung")


###Validación out of sample (train/test) and metrics
# Temporary partition
y_tr <- window(y_ts, end = 2016)
y_te <- window(y_ts, start = 2017)

fit_tr <- auto.arima(y_tr, seasonal = FALSE, stepwise = FALSE, approximation = FALSE, d = 1)
fc_te  <- forecast(fit_tr, h = length(y_te))

# Validation Metrics
accuracy(fc_te, y_te)
autoplot(fc_te) + autolayer(y_te, series = "Observed") +
  labs(title="Out-of-sample forecast (2017–2023)",
       y = "% enrollment") + theme_minimal(base_size = 12)


###Pronóstico end (h = 5–10 years)
# Retrain with the whole series and predict
fit_all <- auto.arima(y_ts, seasonal = FALSE, stepwise = FALSE, approximation = FALSE, d = 1)
fc_10   <- forecast(fit_all, h = 10)

autoplot(fc_10) +
  labs(title = "Secondary Enrollment Forecast (%) – Ecuador",
      x = "Year", y = "%") +
  theme_minimal(base_size = 12)
```

7.- References.

[1] C. T. Simonino, L. F. dos Santos, and A. F. de Freitas, "Implementation of the National School Feeding Program in a pandemic context in the municipality of Viçosa/MG: implications for public management and family farming," *Administração Pública e Gestão Social*, vol. 17, no. 1, 2025. **URL:** https://www.redalyc.org/journal/3515/351580876003/

[2] M. Zanatta Idemori, A. R. Nogueira de Macedo, P. M. Galvão de Souza, and O. S. Gaion, "THE RELATIONSHIP BETWEEN GAMIFICATION AND EXPERIENCE IN DELIVERY APPS," *Revista Pensamento Contemporâneo em Administração*, vol. 19, no. 1, pp. 140–157, 2025. **DOI:** https://doi.org/10.12712/rpca.v.192.65109 UFF+1 Journals

[3] D. Escolar, "Conflict and Ethno-Racial Identifications in the Colonial Prelude to Caudillismo and Montonero Rebellions," *Bulletin of the Institute of Argentine and American History Dr. Emilio Ravignani*, no. 61, pp. 85–91, 2024. **DOI:** https://doi.org/10.34096/bol.rav.n61.14555 SciELO Argentina

[4] F. O. Corrêa Werle and J. A. Lago Fonseca, "School management: foundations and discussions about democratic management," *Essay: Evaluation and Public Policies in Education*, e0254968, 2025. **DOI:** https://doi.org/10.1590/S0104-40362025003304968 Educa FCC

[5] G. E. Castro Rosales, A. D. Torres Alvarado, L. S. Zalamea Cedeño, F. J. Duque-Aldaz, and F. R. Rodríguez-Flores, "Comprehensive Ergonomic Proposal for the Reduction of Musculoskeletal Risks in Soap Production: An Approach Based on Statistical Analysis and Postural Evaluation," *INQUIDE – Ingeniería Química y Desarrollo*, vol. 7, no. 2, 2025. **DOI:** https://doi.org/10.53591/iqd.v7i02.2416 Revistas UG

[6] M. J. Cabrera Valladolid, "Analysis of Time Series of Confirmed and Deceased Cases by Covid-19 in Ecuador," Escuela Superior Politécnica de Chimborazo, 2021. **URL:** https://dspace.espoch.edu.ec/handle/123456789/14812

[7] M. G. Zalduaromero, "Estimation of an import demand function in the plastic industry in Ecuador, period 1980–2016," UESS, 2017. **URL:** http://repositorio.uees.edu.ec/123456789/2008

[8] J. I. Cañarte Murillo, "Restricted growth of the balance of payments: Ecuadorian case of the Thirlwall Law 2000–2016," UESS, 2017. **URL:** http://repositorio.uees.edu.ec/123456789/1980
[9] M. Alós and M. Serio, "Participation in the labor market of older people in Argentina," *Pilquen Journal – Social Sciences Section*, vol. 27, no. 2, pp. 71–96, 2024. **URL:** https://www.redalyc.org/journal/3475/347580385004/
[10] F. J. Duque-Aldaz and E. G. Pazan Gómez, "Factors affecting entrepreneurial intention of Senior University Students," *Espacio*, vol. 39, no. 09, p. 32, 2017. **URL:** https://www.revistaespacios.com/a18v39n09/18390932.html
[11] A. Meneses Freire, L. Zúñiga Lema, J. Muñoz Cargua, J. Lara Sinaluisa, and W. Acurio Martínez, "Nonparametric functional predictive model in functional time series. Application in meteorological variables," Escuela Superior Politécnica de Chimborazo, 2022. **URL:** https://dspace.espoch.edu.ec/handle/123456789/19024
[12] L. Y. Ortega Villegas, "Analysis in the shipping industry in Ecuador as a result of the financial crisis in 2008," UESS, 2018. **URL:** http://repositorio.uees.edu.ec/123456789/2503
[13] D. M. Ichau Tabango, P. J. Pando Sigua, P. L. Cárdenas García, and S. V. Avilés Sacoto, "Demand and inventory management for the creation of an automated information management system: Case study applied to an Ecuadorian supermarket," USFQ, 2021. **URL:** http://repositorio.usfq.edu.ec/handle/23000/10650
[14] F. J. Duque-Aldaz, F. R. Rodríguez-Flores, and J. Carmona Tapia, "Identification of parameters in ordinary differential equation systems using artificial neural networks," *San Gregorio*, vol. 1, no. 2, 2025. **URL:** https://revista.sangregorio.edu.ec/index.php/REVISTASANGREGORIO/article/view/2826
[15] F. F. Sandoya Sanchez and A. G. Abad Robalino, "Comparison of the accuracy of the prediction of data from a time series between ARIMA models and Neural Networks," ESPOL, 2017. **URL:** http://www.dspace.espol.edu.ec/xmlui/handle/123456789/40345
[16] A. F. Mayorga Trujillo, "Research on the increase in productivity in the bolt factory in the company Galo G. Orbea O. Cía. Ltda. through the analysis of availability in the stages of its production process," Escuela Superior Politécnica de Chimborazo, 2017. **URL:** https://dspace.espoch.edu.ec/handle/123456789/6515
[17] D. A. Villarreal Godoy, L. F. Cóndor Simbaña, and C. A. Súarez Nuñez, "Inventory Control in an Ecuadorian Textile Company: Improvement in Production Planning and Control Processes," USFQ, 2022. **URL:** http://repositorio.usfq.edu.ec/handle/23000/12544
[18] N. Vela and M. C. Camacho Cordovez, "Oil Industry: Modeling of Demand Forecasting and Consumption Planning and Projection," USFQ, 2020. **URL:** http://repositorio.usfq.edu.ec/handle/23000/9206
[19] J. S. Fu-López, J. P. Fierro Aguilar, F. R. Rodríguez-Flores, and F. J. Duque-Aldaz, "Application of non-automated Lean strategies for quality improvement in manual assembly processes: A case study in the white goods industry," *INQUIDE – Ingeniería Química y Desarrollo*, vol. 7, no. 1, 2025. **DOI:** https://doi.org/10.53591/iqd.v7i02.2417
[20] J. M. Ausay Carrillo, "Proposal of an econometric model for the sales forecast of the EXIBAL-Riobamba feed factory," Escuela Superior Politécnica de Chimborazo, 2022. **URL:** https://dspace.espoch.edu.ec/handle/123456789/19903
[21] P. A. Eguiguren Calisto and S. V. Avilés Sacoto, "Demand projection models for high volatility and low volume products in sales within a food company," USFQ, 2019. **URL:** http://repositorio.usfq.edu.ec/handle/23000/8320
[22] D. S. Navarro Llivisaca, "Influence of VAT and income tax on tax collections, 2007–2016," UESS, 2017. **URL:** http://repositorio.uees.edu.ec/123456789/1773
[23] I. J. Figueroa Tigrero, "Ecuador's Tax Pressure, Period 2000–2018: An Analysis from the Laffer Curve Theory Approach," UESS, 2019. **URL:** http://repositorio.uees.edu.ec/123456789/3090
[24] A. M. Freire Engracia, A. J. Santos Castello, and I. D. Rivadeneyra, "Analysis of the relationship between foreign direct investment and tax collection: Effects on the trade sector," ESPOL, 2025. **URL:** http://www.dspace.espol.edu.ec/handle/123456789/65933
[25] K. N. Lema Remache, "Comparison between ARIMA and Tylor Kriging (TK) methods for wind speed time series prediction: Case study Chimborazo Meteorological Station," Escuela Superior Politécnica de Chimborazo, 2024. **URL:** https://dspace.espoch.edu.ec/handle/123456789/22839
[26] G. J. Morocho Choca, L. Á. Bucheli Carpio, and F. J. Duque-Aldaz, "Fuel oil fuel dispatch optimization through multivariate regression using local storage indicators," *INQUIDE*, vol. 6, no. 2, 2024. **DOI:** https://doi.org/10.53591/iqd.v6i02.477
[27] J. E. Pincay Moran, A. F. López Vargas, F. J. Duque-Aldaz, W. Villamagua Castillo, and R. Sánchez Casanova, "Evaluation and Proposal for an Environmental Management System in a Mango Plantation," *INQUIDE*, vol. 7, no. 1, 2025. **DOI:** https://doi.org/10.53591/iqd.v7i01.1991
[28] R. J. Guerrero Quinde and C. C. Pérez Siguenza, "Design of a dairy product supply model in a retail company," ESPOL, 2025. **URL:** http://www.dspace.espol.edu.ec/handle/123456789/66133
[29] G. Silva and L. C. Di Serio, "Innovation in 'Forgotten Business,'" *Innovation & Management Review*, vol. 18, no. 4, pp. 350–364, 2021. **URL:** https://www.redalyc.org/journal/5375/537569183001/
[30] F. Medeiros, L. C. Di Serio, and A. Moreira, "Avon Brasil: Optimization of Logistics Processes in a Direct Sale Company," *Revista de Administração*

Vol. 08 / Nº 01
e – ISSN: 3028-8533
ISSN – L: 3028-8533

Contemporânea, vol. 25, no. 4, e-190212, 2021. **DOI:** https://doi.org/10.1590/1982-7849rac2021190212.en

[31] L. M. García-Feria, L. Aguilar-Faisal, R. Pastor-Nieto, and J. C. Serio-Silva, "Changes in vegetation at small landscape scales and captivity alter the gut microbiota of black howler monkeys (Alouatta pigra: Atelidae)," *Acta Biológica Colombiana*, vol. 28, no. 1, pp. 154–164, 2023. **DOI:** https://doi.org/10.15446/abc.v28n1.93450

[32] C. A. García Vázquez, A. I. González Santos, and V. Pérez Garrido, "Methodology and validation algorithm to identify models of an air handling unit," *Electronic, Automatic and Communications Engineering*, vol. 42, no. 1, 2021. **URL:** http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S1815-59282021000100105&lang=es

[33] J. E. Mendoza Cota, "COVID-19 and Employment in Mexico: Initial Impact and Short-Term Forecasts," *Accounting and Administration*, vol. 65, no. 5, 2020. **URL:** https://dialnet.unirioja.es/servlet/articulo?codigo=7731039

[34] J. W. Tudela-Mamani and G. A.-M. Grisellx, "Impact of COVID-19 on Peru's International Tourism Demand: An Application of the Box-Jenkins Methodology," *Revista de Investigaciones Altoandinas*, vol. 24, no. 1, 2022. **DOI:** https://doi.org/10.18271/ria.2022.317

[35] L. Asán Caballero, J. Rojas Delgado, and G. E. Jiménez Moya, "Time Series Prediction Algorithm for Air Traffic Forecasting Based on Artificial Neural Networks," *Cuban Journal of Computer Sciences*, vol. 16, no. 4, 2023. **URL:** http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S2227-18992022000400084&lang=es