CAUCHY – Jurnal Matematika Murni dan Aplikasi
Volume 11 (2) (2026), Pages 1754–1772
p-ISSN: 2086-0382; e-ISSN: 2477-3344

Hybrid ARIMAX–LSTM Modeling with Elastic Net for Forecasting the Education Index of Jambi Province

Mhd Teguh Arrozik*, Tri Candra Nur Muhaimin, and Melkisedek Sampari Koibur

Department of Statistics, Faculty of Science, Technology, and Mathematics (FSTeM), Brawijaya University, Indonesia

Abstract

Forecasting the Education Index is important for evidence-based educational planning and policy evaluation. This study compares hybrid ARIMAX–LSTM models with and without Elastic Net-based variable selection for forecasting the Education Index of Jambi Province, Indonesia. Annual data from 2010 to 2024 were chronologically divided into a training period of 2010–2022 and a testing period of 2023–2024. All candidate exogenous predictors were standardized using parameters estimated exclusively from the training data. Elastic Net variable selection was performed using three-fold expanding-window time-series cross-validation. The selected hyperparameters were $\lambda = 0.001151$ and $\alpha = 0.7$, and seven predictors were retained using an absolute standardized coefficient threshold of 0.01. The ARIMAX$(1,1,0)$ model was used to capture linear temporal dynamics and exogenous effects, while an LSTM network was applied to the ARIMAX residuals to model the remaining nonlinear patterns. In the holdout evaluation, the ARIMAX–LSTM model without Elastic Net obtained an RMSE of 0.5343, an MAE of 0.4668, and a MAPE of 0.7063%. The Elastic Net–ARIMAX–LSTM model obtained lower errors, with an RMSE of 0.3864, an MAE of 0.3271, and a MAPE of 0.4949%. Rolling-origin evaluation also produced lower errors for the Elastic Net model, with an RMSE of 0.2615, an MAE of 0.2326, and a MAPE of 0.3542%. These results indicate that Elastic Net-based variable selection improved forecasting performance under the evaluation settings used in this study. However, the findings should be interpreted cautiously because of the limited number of annual observations.

Keywords: Education Index; hybrid forecasting; Elastic Net; ARIMAX; LSTM; Jambi Province

1. Introduction

Education plays a fundamental role in regional development as it directly contributes to human capital formation, economic productivity, and social welfare. Educational performance is commonly evaluated using composite indicators such as the Education Index, which reflects access to education, participation rates, and educational attainment. These indicators are widely used by policymakers to monitor progress and design targeted interventions. Recent studies emphasize that improvements in educational outcomes are closely linked to long-term economic growth and social development, particularly in developing regions [1].

Despite its importance, forecasting the Education Index remains a challenging task due to the complex interactions among socio-economic and educational factors. Variables such as income

*Corresponding author. E-mail: mhdteguharrozik23@gmail.com

Received: January 06, 2026 Revised: May 19, 2026 Accepted: July 16, 2026
DOI: https://doi.org/10.18860/cauchy.v11i2.40192

per capita, poverty levels, teacher availability, school participation rates, and years of schooling often exhibit strong interdependencies and multicollinearity. These characteristics complicate the modeling process and may reduce forecasting accuracy when conventional approaches are applied without proper variable selection or dimensionality control [2, 3].

Classical time series models, particularly the Autoregressive Integrated Moving Average (ARIMA) framework, have been widely adopted due to their interpretability and solid statistical foundation. However, ARIMA relies solely on the historical dynamics of the target variable and does not incorporate exogenous influences. The ARIMAX extension addresses this limitation by including external regressors X_t to explain variations in the response y_t. Nevertheless, when the predictor matrix exhibits high correlation and the time dimension T is small, ARIMAX estimation becomes unstable. Multicollinearity inflates parameter variance, weakens identifiability, and increases the risk of overfitting, ultimately compromising forecasting accuracy and residual reliability [4].

In parallel, advances in machine learning have introduced deep learning approaches for time-series forecasting. Long Short-Term Memory (LSTM) networks are particularly effective in capturing nonlinear temporal dependencies and long-range interactions. However, deep learning architectures are data-intensive by design. When applied to short annual time series, LSTM models may overparameterize the data, leading to poor generalization and unstable convergence. Furthermore, LSTM does not explicitly separate linear and nonlinear structures, potentially resulting in inefficient representation of linear components that could be more parsimoniously modeled within a statistical framework [5].

To address these limitations, hybrid ARIMAX–LSTM models have been proposed to decompose the data-generating process into linear and nonlinear components. Under this framework, the observed series can be represented as $y_t = y_t^{(L)} + y_t^{(NL)}$, where $y_t^{(L)}$ captures linear dynamics and exogenous effects modeled by ARIMAX, and $y_t^{(NL)}$ represents nonlinear residual structures learned by LSTM. The effectiveness of this decomposition critically depends on the stability and statistical adequacy of the linear component. If ARIMAX estimation is ill-conditioned due to multicollinearity, the extracted residuals $\hat{\epsilon}_t$ may contain noise amplification rather than informative nonlinear signals, thereby weakening the second-stage learning process. Hence, hybridization should not be viewed merely as model stacking, but as a structured estimation pipeline in which the quality of each stage determines overall forecasting performance [6, 7].

Despite the growing popularity of hybrid forecasting models, limited attention has been devoted to integrating regularized estimation within the linear stage of hybrid architectures. Existing studies largely assume that ARIMAX parameters can be reliably estimated before residual modeling, thereby overlooking the bias–variance trade-off induced by correlated regressors in small-sample contexts. In educational forecasting settings, where socio-economic indicators are structurally interrelated, this omission may lead to unstable coefficient estimates and degraded residual informativeness.

This study addresses this methodological gap by embedding Elastic Net regularization into the ARIMAX estimation process. Elastic Net imposes a convex combination of ℓ_1 and ℓ_2 penalties, enabling simultaneous variable selection and coefficient shrinkage under multicollinearity. Formally, the regularized estimation improves conditioning of the exogenous parameter space, enhances estimator stability, and controls overfitting. By stabilizing the linear-exogenous component prior to residual extraction, the proposed framework generates residual series with improved signal-to-noise characteristics, thereby enabling LSTM to focus more effectively on nonlinear dynamics [8, 9]. Table 1 compares the models used in previous studies.

In addition, a clearer comparison with prior hybrid forecasting studies further highlights the unresolved methodological gaps. For example, Mustikasari and Pahrany (2023) demonstrated that Elastic Net provides stable regularization under correlated predictors, effectively addressing multicollinearity in standalone regression settings; however, their approach was not embedded within a hybrid time series decomposition framework [9]. Meanwhile, Sakina (2025) proposed a

Table 1: Comparison of Previous Studies on Models

Study	Model(s)	Key Metric(s)	Main Findings
Mustikasari and Pahrany (2023)	Elastic Net	Selection/prediction performance	Elastic Net provides stable regularization for correlated predictors and is relevant for exogenous-variable selection before ARIMAX estimation.
Sakina (2025)	Hybrid ARIMAX–LSTM	MAPE	The ARIMAX–LSTM combination outperforms other models; integrating exogenous variables into ARIMAX and using LSTM to model nonlinearities improves accuracy.
This Study	Hybrid Elastic Net ARIMAX–LSTM	MAE, RMSE, MAPE	Elastic Net improves forecasting accuracy through effective variable selection and residual learning.

hybrid ARIMAX–LSTM model that improved forecasting performance in terms of MAPE, yet the study did not explicitly address multicollinearity or parameter instability in the ARIMAX stage prior to residual modeling [7]. These findings suggest that although regularization and hybridization have been explored separately in the literature, their structural integration within a unified linear–nonlinear forecasting pipeline remains underdeveloped. In particular, limited attention has been given to how instability in the linear estimation stage may propagate to the residual series and subsequently affect nonlinear learning in LSTM. By embedding Elastic Net regularization directly into the ARIMAX estimation stage, the present study bridges this gap and advances hybrid forecasting beyond performance-oriented model combination toward a statistically grounded and stability-aware estimation framework. This study formally integrates penalized regression theory into a two-stage linear–nonlinear decomposition framework, providing a statistically coherent solution for small-sample multivariate forecasting under correlated predictors. Figure 1 presents a schematic illustration of the sequential integration of Elastic Net-based variable selection, ARIMAX modeling, residual extraction, and LSTM-based nonlinear learning. This framework is applied to the Education Index of Jambi Province along with selected socio-economic and educational indicators, including income per capita, poverty level, teacher–student ratios, school participation rates, expected years of schooling, and average years of schooling from 2010 to 2024, with the objective of examining how exogenous variable selection and residual learning influence forecasting performance.

2. Methods

This section describes the complete modeling framework adopted in this study, including data preparation, partitioning strategy, model construction, and evaluation procedures. The methodological design ensures strict chronological separation between training and testing data to prevent information leakage and to guarantee valid out-of-sample forecasting assessment.

2.1. Data Collection and Preparation

This study uses annual time-series data on the Education Index of Jambi Province, along with a set of socio-economic and educational indicators serving as exogenous variables. The data were obtained from the Central Statistics Agency of Indonesia (*Badan Pusat Statistik*, BPS), the official government institution responsible for collecting and disseminating national statistical data [10]. The dataset spans 2010–2024 and consists of one target variable (Education Index) and eleven socio-economic and educational indicators considered as potential exogenous predictors.

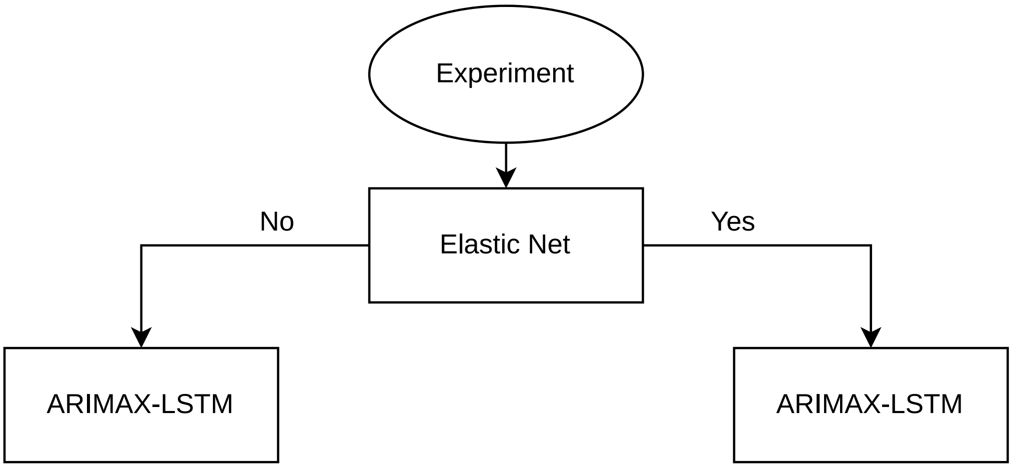


Fig. 1: Model configuration: hybrid architecture with and without an Elastic Net selection strategy.

A detailed description of the research variables is provided in Table 2.

Table 2: Description of Study Variables

No	Variable	Unit	Role
1	Education Index of Jambi Province	Index	Target variable
2	Income per Capita	Rupiah	Exogenous variable
3	Population Living in Poverty	Thousand persons	Exogenous variable
4	Teacher–Student Ratio (Elementary School)	Ratio	Exogenous variable
5	Teacher–Student Ratio (Junior High School)	Ratio	Exogenous variable
6	Teacher–Student Ratio (Senior High School)	Ratio	Exogenous variable
7	School Participation Rate (Elementary)	Percent (%)	Exogenous variable
8	School Participation Rate (Junior High)	Percent (%)	Exogenous variable
9	School Participation Rate (Senior High)	Percent (%)	Exogenous variable
10	Expected Years of Schooling	Years	Exogenous variable
11	Mean Years of Schooling	Years	Exogenous variable
12	Literacy Rate	Percent (%)	Exogenous variable

2.2. Data Partitioning and Evaluation Protocol

To ensure strict chronological out-of-sample evaluation, a time-based holdout strategy was initially implemented. The dataset covering the period 2010–2024 was partitioned as follows:

- **Training period:** 2010–2022
- **Testing period:** 2023–2024

All preprocessing procedures, including Elastic Net variable selection, hyperparameter tuning, and predictor standardization, were performed exclusively on the training data. The estimated transformations were subsequently applied to the testing period without re-estimation in order to prevent information leakage. Forecasting performance was evaluated using Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and Mean Absolute Percentage Error (MAPE), computed on the out-of-sample testing period (2023–2024).

To provide a more comprehensive assessment of model performance and to mitigate the limitation of a short holdout testing period, an expanding-window rolling-origin cross-validation procedure was additionally conducted. In this approach, the model was repeatedly re-estimated using all available observations up to year t and then used to forecast year $t + 1$. Four evaluation folds were considered: (2010–2020 $\rightarrow$ 2021), (2010–2021 $\rightarrow$ 2022), (2010–2022 $\rightarrow$ 2023), and (2010–

2023 → 2024). The rolling-origin procedure was applied separately to both model configurations using the same forecasting origins to ensure a consistent comparison.

For the Elastic Net–ARIMAX–LSTM model, the predictors selected from the training period were used consistently within the rolling-origin evaluation, while predictor standardization was re-estimated exclusively from each training fold. Forecasting performance across the folds was summarized using the average RMSE, MAE, and MAPE values. This additional validation procedure provides supplementary evidence regarding the stability of model performance across multiple temporal data partitions. Figure 2 illustrates the chronological separation between training and testing data used in this study.

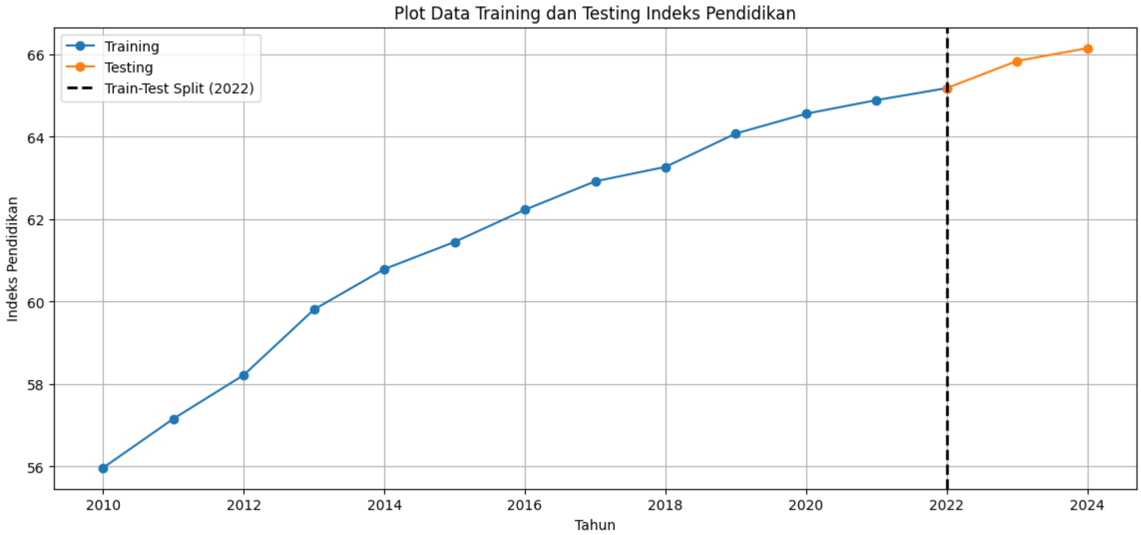


Fig. 2: Training and testing periods

2.3. Stationarity Testing

Stationarity is an important assumption in time series modeling because nonstationary series may lead to spurious parameter estimates and unreliable forecasts [11]. Therefore, stationarity testing was conducted on the original Education Index series using the training data prior to model estimation. In addition to visual inspection of the time-series plot, the Augmented Dickey–Fuller (ADF) test was performed under several specifications, namely: (1) intercept only, (2) intercept with deterministic trend, and (3) first-differenced data. The results are presented in Table 3.

Table 3: Results of the Augmented Dickey–Fuller (ADF) Stationarity Test

Data Series	**ADF Specification**	**ADF Statistic**	***p*-value**	**Conclusion (5%)**
Original target	Intercept only	-3.44741	0.00944	Stationary
Original target	Intercept + trend	-1.16257	0.91790	Non-stationary
First difference	Intercept only	-0.74999	0.83337	Non-stationary
First difference	Intercept + trend	-3.25300	0.04235	Stationary

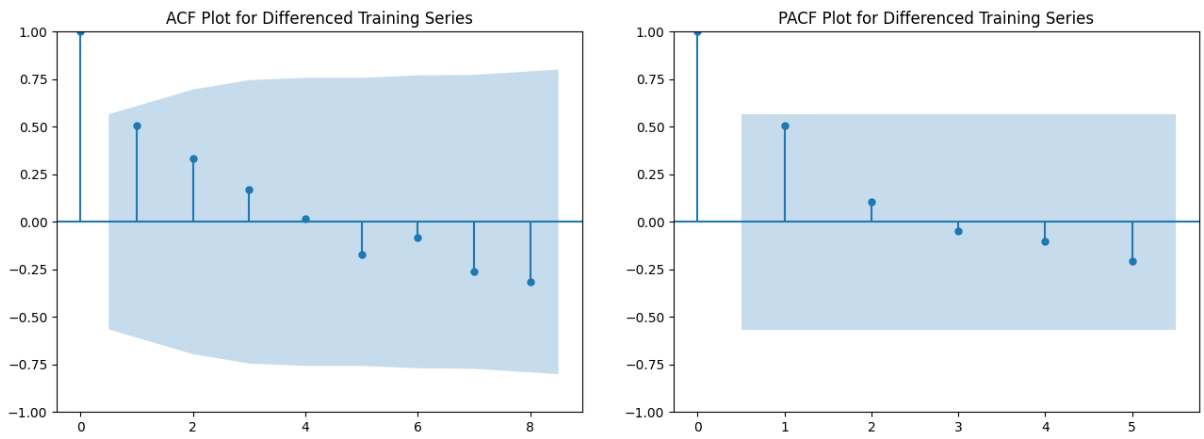
Table 3 shows that the original Education Index series appears stationary when the ADF test includes only an intercept term (ADF $= -3.44741$, p-value $= 0.00944$). However, when a deterministic trend is incorporated, the ADF statistic becomes -1.16257 with a p-value of 0.91790, indicating that the null hypothesis of a unit root cannot be rejected. Because the Education Index exhibits a visible upward trend over time, the specification including both an intercept and a deterministic trend was considered more appropriate for determining stationarity. Under this specification, the level series was classified as nonstationary.

The first-differenced series was subsequently examined. The ADF test with intercept and trend produced a statistic of -3.25300 and a p-value of 0.04235, indicating stationarity at the 5%

significance level. Therefore, first-order differencing was applied and the integration order was set to $d = 1$ for time series model estimation. Given the limited number of annual observations and the presence of a deterministic upward trend in the original series, the use of first-order differencing was adopted as a more conservative modeling strategy to ensure stationarity before parameter estimation.

2.4. Autocorrelation and Lag Identification

After first-order differencing was applied to achieve stationarity, the autocorrelation function (ACF) and partial autocorrelation function (PACF) plots of the differenced series were examined to identify candidate model orders, as presented in Figure 3.



(a) ACF Plot **(b)** PACF Plot

Fig. 3: Autocorrelation function (ACF) and partial autocorrelation function (PACF) plots for the differenced training series

Based on the ACF plot, the autocorrelation values gradually decrease (tailing off), indicating the presence of an autoregressive (AR) pattern in the series. Meanwhile, the PACF plot shows a significant spike at lag 1 and then cuts off after the first lag. This pattern suggests that the data are likely represented by an AR(1) process. Based on the ACF and PACF characteristics, several candidate models were considered, namely ARIMA$(1, 1, 0)$, ARIMA$(0, 1, 1)$, and ARIMA$(1, 1, 1)$. To determine the best model, the Akaike Information Criterion (AIC) was employed, where a lower AIC value indicates a better model fit. The AIC comparison results are presented in Table 4.

Table 4: AIC Values for Candidate Models

Model Order	**AIC**
(1,1,0)	-171.4036
(0,1,1)	-131.3300
(1,1,1)	-98.0460

Based on Table 4, among the candidate models that successfully converged, ARIMA$(1, 1, 0)$ produced the lowest AIC value (-171.4036) and was therefore selected as the most appropriate linear specification for the Education Index series. Candidate models that failed to converge were not used as the basis for final model selection.

Following the Box–Jenkins modeling framework, the identified ARIMA order was subsequently extended by incorporating exogenous variables into the model. As a result, the selected ARIMA$(1, 1, 0)$ specification became an ARIMAX$(1, 1, 0)$ model after the inclusion of external predictors. This approach is consistent with the theoretical formulation of ARIMAX, which extends the ARIMA framework by adding exogenous explanatory variables while retaining the autoregressive, differencing, and moving-average structure previously identified from the target series [12]. The exogenous variables are incorporated through a regression component that captures external influences on the response variable. Therefore, ARIMAX$(1, 1, 0)$ was adopted as the linear component for both the ARIMAX–LSTM and Elastic Net–ARIMAX–LSTM frameworks.

2.5. Experimental Setup

Two hybrid modeling strategies were implemented to evaluate the forecasting performance of the proposed ARIMAX–LSTM framework for the Education Index of Jambi Province. The first strategy represents a baseline hybrid configuration without prior variable selection. The ARIMAX model is estimated using the training period (2010–2022), incorporating all available socio-economic and educational indicators as exogenous variables. The in-sample residuals generated from this model over the training period are subsequently used to train the LSTM component. Table 5 shows the configuration details.

Table 5: Data Splitting Scenarios for Each Model

Model	Training Period	Testing Period	Forecasting Period
Hybrid ARIMAX–LSTM	2010–2022	2023–2024	2025–2027
Elastic Net Hybrid ARIMAX–LSTM	2010–2022	2023–2024	2025–2027

Residuals from the testing period (2023–2024) were reserved strictly for out-of-sample evaluation and are not included in LSTM training. The second strategy incorporates Elastic Net regularization prior to ARIMAX estimation. Elastic Net is applied exclusively to the training data (2010–2022) to identify relevant exogenous predictors. Variables were retained when the absolute value of their standardized Elastic Net coefficient was greater than or equal to 0.01. This threshold was used to exclude predictors with negligible coefficients after regularization. The residuals obtained from the ARIMAX model during the training period were then used to train the LSTM component following the same sequential structure as in the baseline configuration.

This approach enhances parsimony and mitigates multicollinearity while preserving the separation between linear and nonlinear modeling stages. For both strategies, the dataset is partitioned chronologically into a training period (2010–2022) and a testing period (2023–2024). Model performance is evaluated using MAE, RMSE, and MAPE, computed exclusively on the out-of-sample testing period. Given the annual frequency and limited sample size, a parsimonious LSTM architecture is adopted, using short input sequences consisting of the two most recent residual observations to capture nonlinear dependencies without overfitting.

After evaluating model performance over the testing period (2023–2024), the selected hybrid model is re-estimated using the full dataset (2010–2024) to generate forecasts for 2025–2027. Future values of the exogenous variables were unavailable for the forecasting horizon (2025–2027). Therefore, the values observed in 2024 were assumed to remain constant throughout the forecasting period and were used as inputs to the ARIMAX component. This assumption allows future forecasts to be generated while avoiding the introduction of additional uncertainty from separate exogenous-variable forecasting models.

2.6. Hybrid ARIMAX–LSTM Construction

ARIMAX Model Construction Let y_t denote the annual Education Index of Jambi Province. The linear dynamics are modeled using an ARIMAX specification:

$$y_t = c + \sum_{i=1}^{p} \phi_i y_{t-i} + \mathbf{X}_t \boldsymbol{\beta} + \epsilon_t, \tag{1}$$

where c is the intercept, ϕ_i represent the autoregressive parameters, and $\mathbf{X}_t$ is the vector of exogenous variables. The coefficient vector $\boldsymbol{\beta}$ captures the structural influence of socio-economic and educational indicators on the Education Index. ACF–PACF diagnostics and the Augmented Dickey–Fuller (ADF) test indicated stationarity after differencing, leading to an ARIMAX$(1, 1, 0)$ specification. Residual adequacy was assessed using the Ljung–Box test to ensure that the remaining serial correlation was not statistically significant before proceeding to nonlinear modeling.

LSTM Model Construction To capture nonlinear dynamics unexplained by the linear ARIMAX component, a Long Short-Term Memory (LSTM) network was applied to the residual series. Let e_t denote the ARIMAX residuals:

$$e_t = y_t - \hat{y}_t^{(L)}. \tag{2}$$

Given the annual frequency and limited sample size, a parsimonious architecture was adopted, consisting of a single LSTM hidden layer with 16 hidden units followed by a dense output layer with one output neuron. A sequence length of two was employed, meaning that the two most recent residual observations were used to predict the subsequent residual value. Because neural network estimation may be sensitive to random weight initialization, the training procedure was repeated five times. The resulting residual forecasts were subsequently averaged to obtain the final nonlinear correction term used in the hybrid forecasting framework.

Because the dataset consists of only 15 annual observations, of which 13 observations were available for model training, a highly parsimonious LSTM architecture was intentionally adopted to reduce the risk of overfitting. The network contains a single hidden layer with 16 hidden units and a short sequence length of two time steps, thereby limiting the number of trainable parameters relative to the available sample size.

Several additional measures were implemented to mitigate overfitting. First, early stopping based on training loss was applied with a patience value of 30 epochs, allowing training to terminate before unnecessary parameter updates occurred. Second, the residual series modeled by the LSTM component was substantially simpler than the original Education Index series because the linear dynamics had already been captured by the ARIMAX model. Third, the final residual forecast was obtained by averaging the results of five independent training runs, thereby reducing the influence of random initialization and improving prediction stability.

The LSTM model was trained using the Adam optimizer with a learning rate of 0.001 and mean squared error as the loss function. The maximum number of training epochs was set to 200 with a batch size of 1. Early stopping was applied by monitoring the training loss with a patience value of 30, and the best model weights were restored. A random seed of 42 was used to improve reproducibility. The training process was repeated five times using different seed initializations, and the average residual prediction was used as the final nonlinear correction. The nonlinear correction term is estimated as

$$\hat{e}_t = f(\mathbf{E}_{t-1}), \tag{3}$$

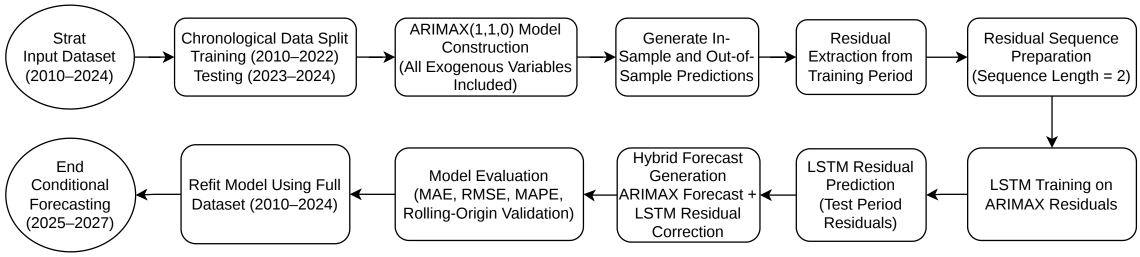
where $f(\cdot)$ represents the nonlinear mapping learned by the network and $\mathbf{E}_{t-1} = [e_{t-1}, e_{t-2}]$.

Hybrid ARIMAX–LSTM Framework The hybrid model integrates both components sequentially. The ARIMAX$(1, 1, 0)$ model first produces linear fitted values, $\hat{y}_t^{(L)}$. Residuals from the training period (2010–2022) are used to train the LSTM, while observations from 2023–2024 are reserved strictly for out-of-sample evaluation. The final forecast is obtained through the additive decomposition

$$\hat{y}_t = \hat{y}_t^{(L)} + \hat{e}_t. \tag{4}$$

This sequential structure assigns complementary roles to each component: ARIMAX captures linear temporal dependence and exogenous effects, while LSTM models residual nonlinear adjustments. The resulting framework improves predictive performance while preserving interpretability and parsimony, which are essential for annual educational time series with limited observations.

Fig. 4: Hybrid ARIMAX–LSTM framework



2.7. Hybrid Elastic Net–ARIMAX–LSTM Model

Elastic Net is a regularization technique that combines the L_1 and L_2 penalties to perform simultaneous variable selection and coefficient shrinkage. This method is particularly suitable for educational and socio-economic datasets, where strong correlations commonly exist among predictors [8, 9]. The Elastic Net estimator is obtained by minimizing the following objective function:

$$\min_{\boldsymbol{\beta}} \left\{ \frac{1}{2N} \sum_{i=1}^{N} (y_i - \mathbf{x}_i\boldsymbol{\beta})^2 + \lambda \left(\alpha \|\boldsymbol{\beta}\|_1 + \frac{1-\alpha}{2} \|\boldsymbol{\beta}\|_2^2 \right) \right\}. \tag{5}$$

The Elastic Net hyperparameters (α and λ) were determined using time-series cross-validation applied exclusively to the training dataset (2010–2022). The optimal parameter combination was selected based on the minimum cross-validated mean squared error. The test dataset (2023–2024) was not involved in the tuning process to ensure strict out-of-sample evaluation. Variable selection was performed using Elastic Net regression with a time-series split cross-validation scheme consisting of three expanding-window folds. The hyperparameter search included λ values generated using `np.logspace(-4,0,50)`, corresponding to 50 candidate values ranging from 10^{-4} to 1.0. The mixing parameter α was evaluated over the set $\{0.1, 0.3, 0.5, 0.7, 0.9, 1.0\}$.

Prior to Elastic Net estimation, all predictor variables and the response variable were standardized using `StandardScaler`. To prevent information leakage, the scaler was fitted exclusively on the training dataset (2010–2022) and subsequently applied to the testing dataset (2023–2024). The optimal hyperparameter combination selected through cross-validation was $\lambda = 0.001151$ and $\alpha = 0.7$, corresponding to the minimum cross-validated mean squared error. The analysis was implemented in Python using the `scikit-learn` package, and a random seed of 42 was specified to ensure reproducibility.

The penalty term is used solely during estimation and does not appear in the final model structure. Variables with coefficients shrunk to zero are excluded, resulting in a reduced set of relevant exogenous variables to be incorporated into the ARIMAX model. Using an absolute standardized coefficient threshold of 0.01, Elastic Net retained seven predictors: income per capita, population living in poverty, teacher–student ratio at junior high school, teacher–student ratio at senior high school, elementary school participation rate, expected years of schooling, and mean years of schooling.

The hybrid Elastic Net–ARIMAX–LSTM model extends the conventional ARIMAX–LSTM framework by incorporating Elastic Net regularization as an initial variable selection stage. Unlike the standard hybrid approach, Elastic Net is not employed for direct forecasting; instead, it is used solely to identify the most relevant exogenous predictors prior to ARIMAX estimation. Let y_t denote the annual Education Index at time t, and let:

$$\mathbf{X}_t = [X_{t,1}, X_{t,2}, \ldots, X_{t,M}], \tag{6}$$

represent the complete set of candidate exogenous variables. Elastic Net regularization is first applied to the training dataset (2010–2022) in order to shrink irrelevant coefficients toward zero and reduce multicollinearity among predictors. Variables associated with nonzero coefficients are retained to form the reduced exogenous vector:

$$\mathbf{X}_t^{\text{sel}} = [X_{t,1}^{\text{sel}}, X_{t,2}^{\text{sel}}, \dots, X_{t,K}^{\text{sel}}] \quad (K \leq M). \tag{7}$$

The selected predictors are subsequently incorporated into the ARIMAX$(1,1,0)$ model, formulated as:

$$y_t = c + \phi_1 y_{t-1} + \mathbf{X}_t^{\text{sel}} \boldsymbol{\beta}^{\text{sel}} + \epsilon_t. \tag{8}$$

This specification allows the model to mitigate multicollinearity while preserving the linear temporal dependence and structural relationships within the time series. The fitted values generated by the ARIMAX component are denoted by:

$$\hat{y}_t^{(L)} = c + \phi_1 y_{t-1} + \mathbf{X}_t^{\text{sel}} \boldsymbol{\beta}^{\text{sel}}. \tag{9}$$

Residuals from the ARIMAX model are then computed as:

$$e_t = y_t - \hat{y}_t^{(L)}. \tag{10}$$

These residuals capture nonlinear and unexplained dynamics that are not adequately represented by the linear ARIMAX structure. The LSTM component subsequently models the residual sequence as:

$$\hat{e}_t = g(e_{t-1}, e_{t-2}, \dots, e_{t-L}), \tag{11}$$

where L denotes the sequence length. Finally, the hybrid prediction is obtained by combining the linear ARIMAX forecast with the nonlinear residual correction generated by the LSTM model:

$$\hat{y}_t = \hat{y}_t^{(L)} + \hat{e}_t. \tag{12}$$

This hybrid framework enables Elastic Net to perform variable selection, ARIMAX to capture linear temporal and exogenous relationships, and LSTM to model nonlinear residual patterns, thereby improving forecasting robustness and predictive performance for educational time-series data.

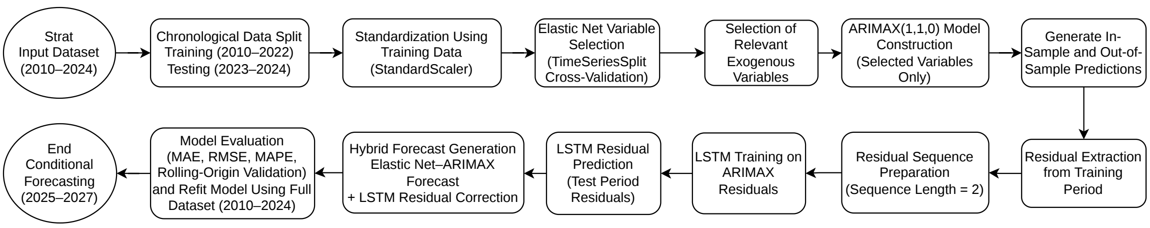


Fig. 5: Hybrid ARIMAX–LSTM framework with Elastic Net

All hybrid variants were evaluated using MAE, RMSE, and MAPE to ensure comparability with standalone models and to assess the impact of data partitioning on forecasting accuracy.

2.8. Evaluation and Comparison

To comprehensively evaluate the predictive performance of all forecasting models developed in this study, three standard accuracy metrics are employed: Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and Mean Absolute Percentage Error (MAPE) [13, 14]. These metrics are widely used in time-series forecasting and provide complementary perspectives on model accuracy and robustness [14].

MAE measures the average magnitude of forecast errors regardless of direction, offering a general assessment of prediction accuracy. The MAE is defined as:

$$\text{MAE} = \frac{1}{n} \sum_{t=1}^{n} |y_t - \hat{y}_t|. \tag{13}$$

RMSE penalizes larger errors more heavily by squaring the deviations before averaging, making it particularly sensitive to substantial forecast deviations. This metric is useful for evaluating the stability of forecasting models and is expressed as:

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{t=1}^{n}(y_t - \hat{y}_t)^2}. \tag{14}$$

MAPE measures the prediction error as a percentage of the actual value, enabling scale-independent and intuitive comparisons across models. The MAPE is computed as:

$$\text{MAPE} = \frac{1}{n}\sum_{t=1}^{n}\left|\frac{y_t - \hat{y}_t}{y_t}\right| \times 100\%. \tag{15}$$

The use of multiple evaluation metrics allows for a balanced and comprehensive assessment of forecasting performance. While MAE provides insight into average prediction accuracy, RMSE emphasizes the impact of larger errors, and MAPE offers a relative measure of accuracy that is easily interpretable by policymakers and practitioners [15, 16]. This combination is particularly appropriate for evaluating long-term educational indicators, which typically evolve gradually and require robust and interpretable forecasting measures [17].

The 95% prediction intervals for the hybrid forecasts were calculated using an approximate combination of uncertainty components. The standard error obtained from the ARIMAX forecasting distribution was combined with the standard deviation of the LSTM residual prediction errors. The resulting total standard error was calculated as the square root of the sum of the two variance components:

$$\text{SE}_{\text{total}} = \sqrt{\text{SE}^2_{\text{ARIMAX}} + \text{SD}^2_{\text{LSTM}}}. \tag{16}$$

Therefore, the reported intervals should be interpreted as approximate hybrid prediction intervals rather than exact probabilistic intervals.

3. Results and Discussion

This section presents and discusses the forecasting results obtained from the two model configurations evaluated in this study. The analysis focuses on comparing hybrid forecasting approaches with and without Elastic Net regularization, emphasizing the contribution of nonlinear residual learning and Elastic Net-based variable selection. Model performance is assessed using standard accuracy metrics to evaluate both predictive accuracy and model robustness in forecasting the Education Index of Jambi Province.

3.1. Evaluation Overview

This subsection summarizes the overall forecasting performance of all proposed models and provides a comparative assessment based on three widely used accuracy metrics: Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and Mean Absolute Percentage Error (MAPE). Each metric captures a different aspect of forecasting accuracy: MAE measures the average magnitude of forecast errors, RMSE penalizes larger deviations more heavily and is sensitive to extreme errors, while MAPE expresses prediction errors in percentage terms, allowing intuitive interpretation and comparability across time periods.

To ensure a realistic evaluation, all models were trained using historical data from 2010 to 2022 and evaluated on out-of-sample observations for 2023–2024. This out-of-sample evaluation framework reflects a genuine forecasting scenario in which future educational outcomes are predicted without access to subsequent information. By applying the same training–testing split and evaluation metrics across all model configurations, the comparison remains fair and consistent, enabling an objective assessment of the added value of hybrid modeling and Elastic

Net regularization. Because the annual dataset is relatively small and the holdout period contains only two observations (2023–2024), an expanding-window rolling-origin cross-validation procedure was additionally conducted for the period 2021–2024. This supplementary validation was intended to evaluate the stability of forecasting performance across multiple temporal partitions and to reduce the dependence of conclusions on a single train–test split.

3.2. ARIMAX–LSTM Model without Elastic Net

The ARIMAX–LSTM model without Elastic Net represents a hybrid forecasting approach that integrates the ARIMAX model to capture linear temporal dynamics and the influence of exogenous variables on the Education Index, while the LSTM network is employed to learn nonlinear patterns and long-term dependencies present in the ARIMAX residuals. In this configuration, all available exogenous variables are included in the ARIMAX component without any prior variable selection or regularization.

Fig. 6: ARIMAX–LSTM forecasts without Elastic Net

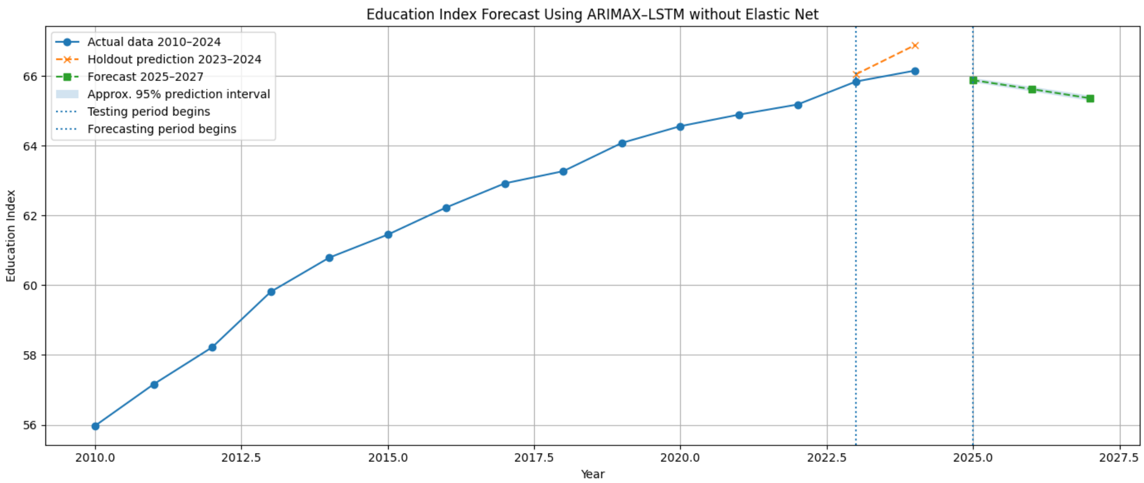


Figure 6 illustrates the comparison between the actual Education Index and the forecasts generated by the ARIMAX–LSTM model without Elastic Net. The figure shows that the model is generally capable of tracking the overall upward trend of the Education Index. However, noticeable deviations between predicted and actual values remain, particularly toward the end of the observation period. Table 6 presents the actual and predicted values for the test years 2023 and 2024.

Table 6: Actual and Predicted Values of the ARIMAX–LSTM Model without Elastic Net

Year	Actual	Predicted Value
2023	65.84	66.0468
2024	66.15	66.8767

As shown in Table 6, the model slightly overestimated the Education Index in both test years. In 2023, the actual value of 65.84 was predicted as 66.0468, while in 2024 the actual value of 66.15 was predicted as 66.8767. Although the prediction errors remain relatively small, the larger deviation observed in 2024 indicates a tendency toward instability in the forecasts. The absence of a regularization mechanism causes the ARIMAX component to remain sensitive to model complexity and potential multicollinearity among exogenous variables, which in turn affects the quality of residuals learned by the LSTM model. Table 7 presents the forecasted Education Index values generated by the ARIMAX–LSTM model without Elastic Net for the period 2025–2027.

The ARIMAX–LSTM model without Elastic Net produces relatively stable forecasts of the Education Index for the period 2025–2027. The forecast values decrease gradually from 65.8787

Table 7: Forecast Values of the ARIMAX–LSTM Model without Elastic Net

Year	Forecast	Lower 95% PI	Upper 95% PI
2025	65.8787	65.8164	65.9410
2026	65.6226	65.5548	65.6903
2027	65.3600	65.2844	65.4355

in 2025 to 65.6226 in 2026 and 65.3600 in 2027. This pattern indicates a gradual downward trajectory under the assumption that all future exogenous variables remain constant at their 2024 observed values; however, the minimal variation across years suggests limited sensitivity to subtle dynamic changes in the data. One factor contributing to the relatively flat forecasting trajectory is the assumption that all exogenous variables remain constant at their 2024 values during the forecasting horizon.

Under this conditional forecasting scenario, the ARIMAX model no longer receives additional information from future changes in educational and socio-economic indicators. As a result, the forecasts are determined mainly by the autoregressive structure of the model and tend to converge toward a stable long-run trajectory. The 95% prediction intervals remain relatively narrow, ranging from $[65.8164, 65.9410]$ in 2025 to $[65.2844, 65.4355]$ in 2027. These intervals provide an approximate measure of forecast uncertainty and indicate that the predicted values are expected to remain within a limited range under the assumed exogenous-variable scenario. However, given the small number of annual observations available for model estimation, the prediction intervals should be interpreted with caution and viewed as indicative rather than definitive measures of uncertainty.

3.3. ARIMAX–LSTM Model with Elastic Net

The ARIMAX–LSTM model with Elastic Net extends the previous hybrid framework by incorporating Elastic Net regularization at the ARIMAX stage. Elastic Net serves two primary purposes: selecting the most relevant exogenous variables and controlling model complexity through coefficient shrinkage. From the initial set of 11 candidate exogenous variables, Elastic Net retained seven predictors using an absolute standardized coefficient threshold of 0.01. Table 8 presents the coefficient estimation results.

Table 8: Elastic Net Coefficient Estimation and Variable Selection Results

No	Variables	Coefficient	Standardized Coeff.	Status
1	Mean Years of Schooling	0.4011	0.027417	Selected
2	Expected Years of Schooling	0.3878	-0.049959	Selected
3	Income per Capita	0.0913	-0.001684	Not Selected
4	Literacy Rate	0.0643	-0.060223	Selected
5	School Participation Rate (Senior High)	0.0310	0.048930	Selected
6	Teacher–Student Ratio (Elementary School)	0.0000	-0.010974	Selected
7	Population Living in Poverty	0.0000	-0.001971	Not Selected
8	School Participation Rate (Junior High)	0.0000	0.000000	Not Selected
9	School Participation Rate (Elementary)	-0.0000	0.564537	Selected
10	Teacher–Student Ratio (Senior High School)	0.0000	0.429661	Selected
11	Teacher–Student Ratio (Junior High School)	-0.0000	0.000000	Not Selected

Based on Table 8, Elastic Net retained seven predictors: income per capita, population living in poverty, teacher–student ratio at junior high school, teacher–student ratio at senior high school, elementary school participation rate, expected years of schooling, and mean years of schooling. Expected years of schooling and mean years of schooling obtained the largest absolute standardized coefficients, with values of 0.564537 and 0.429661, respectively. These results indicate that the two schooling-duration indicators provided the strongest relative contributions within the standardized Elastic Net model.

The remaining selected variables had smaller standardized coefficients but exceeded the predetermined threshold of 0.01. The selected coefficients should be interpreted as measures of relative predictive contribution after standardization and regularization, rather than as evidence of direct causal relationships. Consequently, the reported standardized coefficients reflect the relative contribution of each predictor after adjusting for differences in measurement scales. Variables whose coefficients were shrunk to zero were excluded from the final model because their contribution became negligible after regularization. Therefore, the selected variables should be interpreted as the most relevant predictors within the Elastic Net framework rather than as indicators of direct causal relationships.

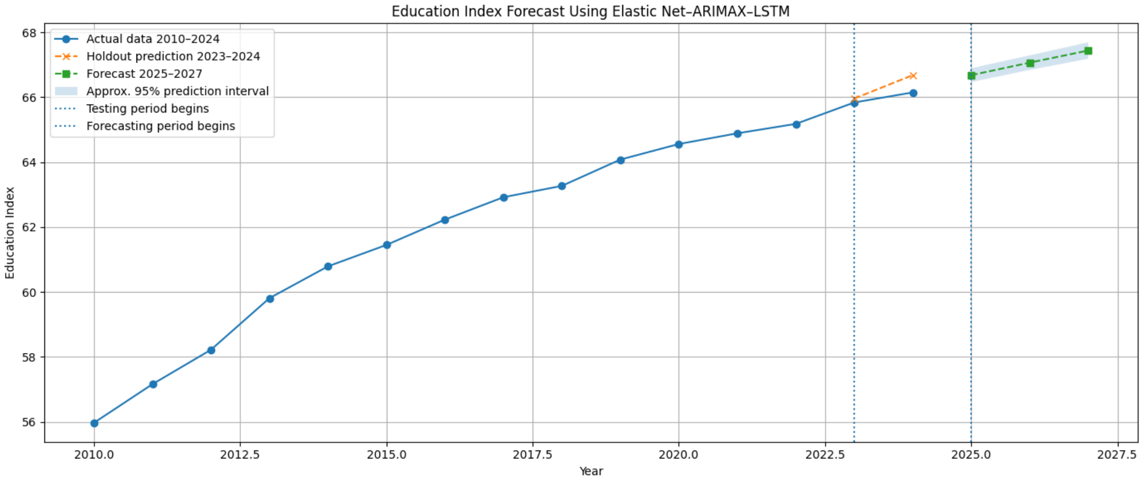


Fig. 7: ARIMAX–LSTM forecasts with Elastic Net

Fig. 7 shows the forecast results of the ARIMAX–LSTM model with Elastic Net. The predicted values were relatively close to the actual Education Index values during the testing period, although the model slightly overpredicted the target in both 2023 and 2024. Table 9 summarizes the actual and predicted values for the years 2023 and 2024.

Table 9: Actual and Predicted Values of the ARIMAX–LSTM Model with Elastic Net

Year	Actual	Prediction
2023	65.84	65.9613
2024	66.15	66.6828

The results in Table 9 indicate relatively small prediction errors in both test years. Compared to the model without Elastic Net, the forecasts are more stable and closer to the actual values. Specifically, the actual value of 65.84 in 2023 was predicted as 65.9613, while the actual value of 66.15 in 2024 was predicted as 66.6828. The lower forecasting errors indicate that Elastic Net-based variable selection improved predictive performance within the evaluation period used in this study. The improvement is associated with a more parsimonious exogenous-variable set; however, the results do not provide direct evidence that the residuals became cleaner or that multicollinearity was completely eliminated. Table 10 presents the forecasted values for the period 2025–2027.

In Table 10, the Elastic Net–ARIMAX–LSTM model forecasts a gradual increase in the Education Index from 66.6302 in 2025 to 67.0282 in 2026 and 67.4726 in 2027. Unlike the model without Elastic Net, this forecasting pattern is more consistent with the historical upward trend observed in the Education Index. The increasing trajectory is mainly driven by the ARIMAX component after Elastic Net regularization has retained the most relevant educational predictors while removing less informative variables. By reducing model complexity and multicollinearity,

Table 10: Forecast Values of the ARIMAX–LSTM Model with Elastic Net

Year	Forecast	Lower 95% PI	Upper 95% PI
2025	66.6302	66.4696	66.8952
2026	67.0282	66.8447	67.2916
2027	67.4726	67.1874	67.6931

Elastic Net produces a more stable linear structure that better preserves the long-term trend embedded in the historical data.

These are conditional forecasts because all future exogenous variables are assumed to remain constant at their 2024 values. Under this assumption, the model projects the future Education Index based on the historical relationships estimated from the available data. The 95% prediction intervals remain relatively narrow, ranging from $[66.4696, 66.8952]$ in 2025 to $[67.1874, 67.6931]$ in 2027. These intervals suggest a moderate degree of forecast certainty under the assumed scenario, although the limited sample size implies that the intervals should be interpreted as approximate measures of uncertainty rather than precise probabilistic bounds.

3.4. Residual Diagnostics Comparison

Residual diagnostics were performed to assess whether the residuals of the linear models satisfied the assumption of serial independence. In time-series forecasting, residuals that behave as white noise indicate that the model has adequately captured the underlying temporal structure of the data. Therefore, residual variance, residual standard deviation, residual autocorrelation function (ACF), and the Ljung–Box test were examined for both the ARIMAX and Elastic Net–ARIMAX models to evaluate the adequacy of the fitted models and to compare the impact of Elastic Net variable selection on residual behavior.

The ARIMAX model produced a residual variance of 279.3188 and a residual standard deviation of 16.7128, while the Elastic Net–ARIMAX model yielded a residual variance of 302.0760 and a residual standard deviation of 17.3803. The Ljung–Box test resulted in statistics of 0.4923 (p-value = 0.7818) for ARIMAX and 0.3155 (p-value = 0.8541) for Elastic Net–ARIMAX. Since both p-values exceed the 0.05 significance level, there is no evidence of significant residual autocorrelation in either model, indicating that the serial dependence structure of the data has been adequately captured. This conclusion is further supported by the residual ACF plots, in which most autocorrelation coefficients remain within the confidence bounds. Although the Elastic Net–ARIMAX model achieved a slightly higher Ljung–Box p-value, its residual variance and residual standard deviation were marginally larger than those of the ARIMAX model. Therefore, the primary contribution of Elastic Net in this study should not be interpreted as substantially improving residual diagnostics, but rather as enhancing variable selection and forecasting performance, as reflected by the lower forecasting errors obtained in the Elastic Net–ARIMAX–LSTM framework.

The relatively large residual variance values were strongly influenced by the first fitted residual, which reflects the initialization effect of the integrated ARIMAX model. Therefore, residual variance and residual standard deviation should be interpreted cautiously and were not used as the primary basis for comparing model quality. The Ljung–Box test and residual ACF were emphasized to evaluate remaining serial dependence.

3.5. Model Comparison and Evaluation

The comparison between the two models was conducted using both holdout evaluation (2023–2024) and rolling-origin cross-validation results. Visually, the ARIMAX–LSTM model with Elastic Net exhibited smaller prediction deviations and a curve that was more consistent with the observed Education Index values than the model without Elastic Net.

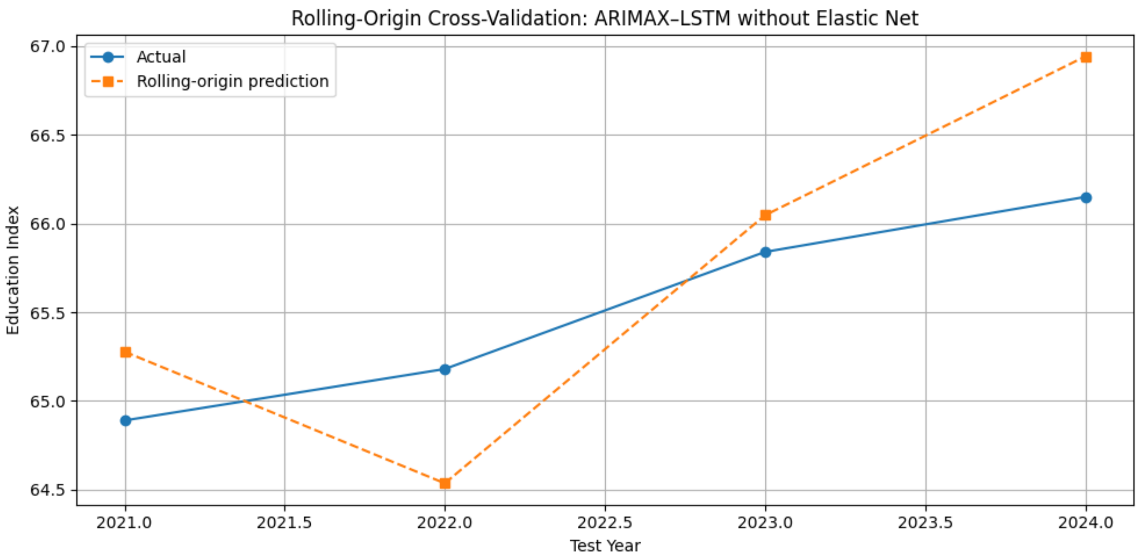
Based on Table 11A, the holdout evaluation for the 2023–2024 testing period shows that the

Table 11A: Model Comparison and Evaluation (Holdout Evaluation 2023–2024)

Model	RMSE	MAE	MAPE
ARIMAX–LSTM without Elastic Net	0.5343	0.4668	0.7063%
ARIMAX–LSTM with Elastic Net	0.3864	0.3271	0.4949%

Elastic Net–ARIMAX–LSTM model achieved lower forecasting errors than the ARIMAX–LSTM model without Elastic Net. The RMSE decreased from 0.5343 to 0.3864, the MAE decreased from 0.4668 to 0.3271, and the MAPE decreased from 0.7063% to 0.4949%. These results indicate that, within the fixed holdout period, the use of Elastic Net for exogenous-variable selection was associated with improved forecasting accuracy. Nevertheless, because the holdout period consists of only two annual observations, these results should be interpreted as preliminary out-of-sample evidence rather than definitive proof of model superiority.

Fig. 8: Rolling-origin cross-validation forecasts from ARIMAX–LSTM without Elastic Net



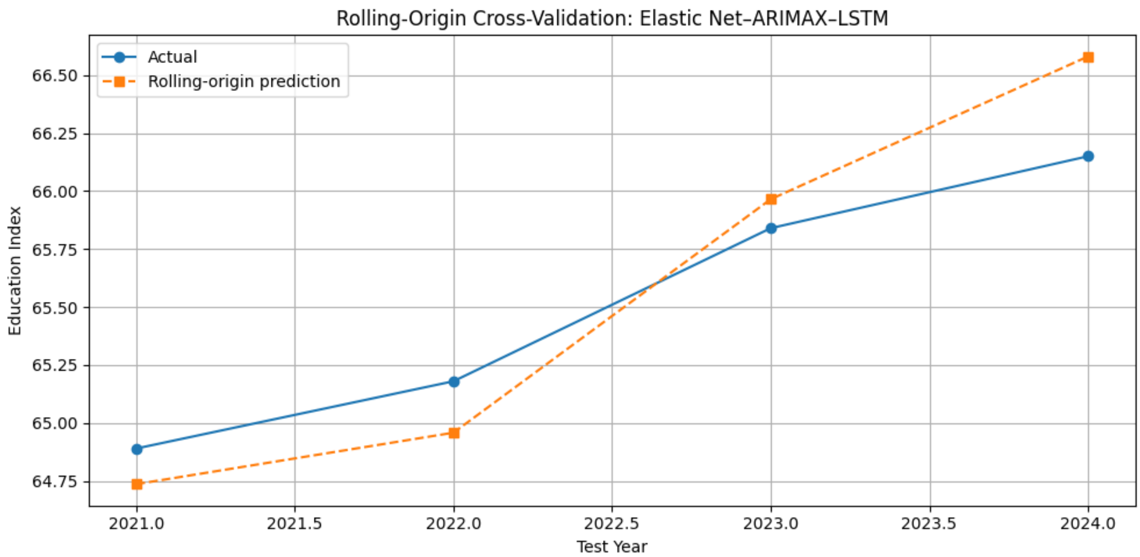


Fig. 9: Rolling-origin cross-validation forecasts from ARIMAX–LSTM with Elastic Net

Based on Table 11B, the rolling-origin evaluation provides additional evidence across several temporal partitions. The Elastic Net–ARIMAX–LSTM model obtained an RMSE of 0.2615, an

Table 11B: Model Comparison and Evaluation (Rolling-Origin Evaluation 2021–2024)

Model	RMSE	MAE	MAPE
ARIMAX–LSTM without Elastic Net	0.5557	0.5076	0.7742%
ARIMAX–LSTM with Elastic Net	0.2615	0.2326	0.3542%

MAE of 0.2326, and a MAPE of 0.3542%, while the model without Elastic Net yielded an RMSE of 0.5557, an MAE of 0.5076, and a MAPE of 0.7742%.

The lower values across all three error metrics indicate that the Elastic Net-based model produced more accurate forecasts over the evaluated rolling origins. These findings suggest that variable selection contributed to a more parsimonious forecasting structure and was associated with more consistent predictive performance across the temporal splits considered. However, the results should still be interpreted cautiously because the rolling-origin assessment involved only four annual forecasting origins and a relatively small overall dataset.

4. Conclusion

This study compared ARIMAX–LSTM models with and without Elastic Net-based variable selection for forecasting the Education Index of Jambi Province. Elastic Net retained seven of the eleven candidate exogenous predictors using an absolute standardized coefficient threshold of 0.01. Within the holdout evaluation for 2023–2024, the Elastic Net–ARIMAX–LSTM model obtained an RMSE of 0.3864, an MAE of 0.3271, and a MAPE of 0.4949%, while the model without Elastic Net obtained an RMSE of 0.5343, an MAE of 0.4668, and a MAPE of 0.7063%. The rolling-origin evaluation also showed lower forecasting errors for the Elastic Net model, with an RMSE of 0.2615, an MAE of 0.2326, and a MAPE of 0.3542%.

These results indicate that Elastic Net-based variable selection was associated with improved forecasting performance within the evaluation settings used in this study. The selected model produced conditional forecasts of 66.6824, 67.0682, and 67.4402 for 2025, 2026, and 2027, respectively, under the assumption that future exogenous variables remain constant at their 2024 observed values.

The findings should be interpreted cautiously because the dataset contains only 15 annual observations, the fixed holdout period consists of only two observations, and the future exogenous variables were not independently forecast. Future research should consider longer or higher-frequency datasets, additional rolling-origin folds, alternative exogenous-variable scenarios, and more formal uncertainty estimation procedures.

CRediT Authorship Contribution Statement

Mhd. Teguh Arrozik: Conceptualization, Methodology, Data Curation, Formal Analysis, Writing – Original Draft.
Tri Candra Nur Muhaimin: Visualization, Formal Analysis, Writing – Review & Editing.
Melkisedek Sampari Koibur: Validation, Visualization, Writing – Review & Editing.

Declaration of Generative AI and AI-assisted technologies

The authors acknowledge the use of AI-assisted tools, specifically **ChatGPT (OpenAI, GPT-4, June 2025 version)**, to support the academic writing process. The tool was used to help structure content, improve phrasing, and ensure academic clarity. All suggestions generated by the AI were reviewed, edited, and approved by the authors to ensure accuracy, originality, and alignment with the intended meaning. The final manuscript reflects the authors' own analysis, interpretation, and scholarly judgment.

Declaration of Competing Interest

The authors declare that they have no known competing financial or personal interests that could have influenced the work reported in this paper.

Funding and Acknowledgments

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors. The authors acknowledge the availability of open statistical data provided by official Indonesian institutions, which enabled the completion of this study.

Data and Code Availability

The data used in this study consist of annual Education Index indicators and related socio-economic variables for Jambi Province, obtained from publicly available sources, including Badan Pusat Statistik (BPS) and official regional publications. The datasets are openly accessible through these institutions. The computational codes supporting the findings of this study are not publicly available but may be provided by the corresponding author upon reasonable request for academic purposes.

References

[1] E. A. Hanushek and L. Woessmann. "Education, knowledge capital, and economic growth". In: *Economics of Education Review* (). DOI: 10.1016/B978-0-12-815391-8.00014-8.

[2] G. Bontempi, S. Ben Taieb, and Y.-A. Le Borgne. *Machine learning strategies for time series forecasting.* 3rd. doi: RePEc:ulb:ulbeco:2013/167761. European Business Review.

[3] L. J. Slater et al. "Hybrid forecasting: using statistics and machine learning to integrate predictions from dynamical models". In: *Eur. Geosci. Union* 1.September (2022), pp. 1–35. DOI: 10.5194/hess-2022-334.

[4] R. H. Shumway and D. S. Stoffer. *Time Series Analysis and Its Applications.* 5th. Springer Texts in Statistics. Springer. DOI: 10.1007/978-3-031-70584-7.

[5] B. Lim, S. Ö. Arık, N. Loeff, and T. Pfister. "Temporal Fusion Transformers for interpretable multi-horizon time series forecasting". In: *Int. J. Forecast.* 37.4 (2021), pp. 1748–1764. DOI: 10.1016/j.ijforecast.2021.03.012.

[6] K. Kashif and R. Lepaczuk. "LSTM-ARIMA As A Hybrid Approach In Algorithmic Investment Strategies". In: *Work. Pap.* 7.443 (2024), pp. 44–50. DOI: 10.33138/2957-0506.2024.7.443.

[7] D. F. Sakina and A. Muhaimin. "Application of ARIMAX-LSTM Model in Forecasting the Price of Broiler Chicken in Central Java". In: *J Stat.* 18.1 (2025), pp. 909–919. DOI: 10.36456/jstat.vol18.no1.a10555.

[8] P. Breheny and J. Huang. *High-Dimensional Regression Modeling.* 1st. Chapman and Hall/CRC. https://myweb.uiowa.edu/pbreheny/7240/s25/book.pdf.

[9] A. Mustikasari and A. D. Pahrany. "Lasso Regression and Elastic Net in Analysing Factors Affecting the Open Unemployment Rate". In: *J. Varian* 8.2 (2025), pp. 151–164. DOI: 10.30812/xey50x64.

[10] Badan Pusat Statistik Indonesia. *Statistik Pendidikan 2024.* Jakarta, Indonesia, 2024. https://www.bps.go.id/id/publication/2024/11/22/c20eb87371b77ee79ea1fa86/statistik-pendidikan-2024.html.

[11] Y. Liu and H. Wu. "Non-stationary Transformers: Exploring the Stationarity in Time Series Forecasting". In: *Advances in Neural Information Processing Systems (NeurIPS)*. Vol. 35. 2022, pp. 9881–9893. DOI: 10.5555/3600270.3600988.

[12] O. Basheer and A. Ojutomori. "Fitting Autoregressive Integrated Moving Average with Exogenous Variables Model with Lognormal Error Term". In: *J. Sci. Res. Reports* 30.10 (2024), pp. 158–168. DOI: 10.9734/jsrr/2024/v30i102442.

[13] T. Fissler, R. Frongillo, J. Hlavinov, and B. Rudloff. "Forecast evaluation of quantiles, prediction intervals, and other". In: *Electron. J. Stat.* 15.1657598 (2021), pp. 1034–1084. DOI: 10.1214/21-EJS1808.

[14] R. J. Hyndman and G. Athanasopoulos. *Forecasting: Principles and Practice*. 3rd. Melbourne: OTexts, 2021. https://otexts.com/fpp3/.

[15] R. D. Gidaro and J. L. Harvill. "Evaluating Approximations of Count Distributions and Forecasts for Poisson-Lindley Integer Autoregressive Processes". In: *arXiv preprint stat.ME* 1 (2024), pp. 1–40. DOI: 10.48550/arXiv.2405.20342.

[16] G. Buturac. "Measurement of Economic Forecast Accuracy: A Systematic Overview of the Empirical Literature". In: *Risk Financ. Manag.* 15.1 (2022), pp. 1–28. DOI: 10.3390/jrfm15010001.

[17] F. Petropoulos et al. "Forecasting: theory and practice". In: *Int. J. Forecast.* 38.3 (2022), pp. 705–871. DOI: 10.1016/j.ijforecast.2021.11.001.