Matrik: Jurnal Manajemen, Teknik Informatika, dan Rekayasa Komputer
Vol. 25, No. 3, July 2026, pp. 523∼538
ISSN: 2476-9843, accredited by Kemenristekdikti, Decree No: 10/C/C3/DT.05.00/2025
DOI: 10.30812/matrik.v25i3.6026

Assessing the Effectiveness of Statistical and Temporal Imputation Methods for Bi-LSTM-Based Forecasting on Environmental and Climate Time Series Data

Adelia Desyana Eka Putri, Aji Prasetya Wibawa, Adelia Khansa Ristiaputri, Adhelia Wida Khaidir, Dhia Rafifah Thifal, Agung Bella Putra Utama
Universitas Negeri Malang, Malang, Indonesia

Article Info

Article history:

Received December 15, 2025
Revised January 27, 2026
Accepted March 27, 2026

Keywords:

Bi-LSTM;
Imputation;
Missing value;
Particle Swarm Optimization;
Time series forecasting.

ABSTRACT

Time series data in climatology and environmental research are highly susceptible to missing values that can disrupt temporal structures and degrade forecasting performance. This study evaluates the effectiveness of several imputation methods in improving the predictive performance of a Bidirectional Long Short-Term Memory model across three missing-data mechanisms: Missing Completely at Random, Missing at Random, and Missing Not at Random. The compared methods include mean, median, mode, k-nearest neighbors, multiple imputation by chained equations, and last observation carried forward, with data deletion serving as the baseline. All datasets were normalized using the min–max technique, and model hyperparameters were optimized through Particle Swarm Optimization. Performance was assessed using mean absolute percentage error, root mean square error, and the coefficient of determination. The findings indicate that proper imputation significantly enhances forecasting accuracy compared to deleting incomplete observations. In Dataset 1, the last observation carried forward achieved the best performance with a coefficient of determination of 0.923 and a root mean square error of 3.373. Similarly, Dataset 2 showed optimal results with the same method, producing a coefficient of determination of 0.950 and a root mean square error of 14.458. The most substantial improvement was observed in Dataset 3, where mean imputation reduced the mean absolute percentage error from 3.219 to 0.329 while increasing the coefficient of determination to 0.986. These results highlight the critical role of selecting an imputation strategy in deep learning-based time series forecasting and provide practical guidance for handling incomplete environmental datasets.

Corresponding Author:

Aji Prasetya Wibawa,
Department of Electrical and Informatics Engineering,
Universitas Negeri Malang, Malang, Indonesia,
Email: **aji.prasetya.ft@um.ac.id**

How to Cite:
A. D. Eka Putri, A. P. Wibawa, A. K. . Ristiaputri, A. W. Khaidir, D. R. Thifal, and A. B. Putra Utama, "Assessing the Effectiveness of Statistical and Temporal Imputation Methods for Bi-LSTM-Based Forecasting on Environmental and Climate Time Series Data", *MATRIK: Jurnal Manajemen, Teknik Informatika, dan Rekayasa Komputer*, Vol. 25, No. 3, pp. 523-538, July, 2026.

1. INTRODUCTION

Time series data in climatology, environmental monitoring, and solar activity research are crucial for understanding atmospheric dynamics, climate variability, and large-scale environmental change [1]. The reliability of these data strongly influences scientific analysis, weather forecasting, climate modeling, and monitoring of astrophysical phenomena such as sunspots [2]. However, time-series observational data are highly susceptible to missing values due to sensor malfunctions, extreme environmental conditions, instrument degradation, and irregular acquisition processes [3]. Missing values disrupt temporal continuity and damage the sequential patterns required by deep learning forecasting models, ultimately reducing prediction accuracy and increasing analytical uncertainty [4]. Therefore, addressing missing data effectively is essential to ensure the integrity and reliability of time-series forecasting systems.

Numerous studies have proposed methods to address missing data problems. Statistical techniques such as Mean, Median, and Mode imputation are widely used due to their simplicity and computational efficiency [5]. Distance-based approaches, such as K-Nearest Neighbors (KNN), estimate missing values using the similarity of neighboring observations, while model-based approaches, such as Multiple Imputation by Chained Equations (MICE), provide robust estimation for multivariate datasets [6]. Temporal methods, including Last Observation Carried Forward (LOCF) and interpolation, are frequently applied to climatological datasets to preserve chronological consistency [7]. On the predictive modeling side, Long Short-Term Memory (LSTM) and Bidirectional Long Short-Term Memory (Bi-LSTM) networks have demonstrated strong capabilities for capturing long-term dependencies in sequential data [8]. Despite these advances, the effectiveness of each imputation method varies depending on the characteristics of the missing data and the forecasting model employed.

Previous research has separately explored the use of imputation techniques and deep learning models; however, many studies assume complete datasets and do not thoroughly evaluate how different imputation strategies affect forecasting performance [9]. Several works have compared imputation methods, yet only a limited number consider the mechanisms of missing data, such as Missing Completely At Random (MCAR), Missing At Random (MAR), and Missing Not At Random (MNAR), despite their distinct statistical properties [10, 11]. Other studies focus solely on a single dataset or imputation approach, limiting the generalizability of their findings. Furthermore, the relationship between imputation quality and Bi-LSTM forecasting performance remains insufficiently investigated, particularly in environmental and climatological contexts that heavily depend on temporal structure [12, 13]. Existing studies also rarely ensure optimal forecasting conditions by optimizing hyperparameters when evaluating imputation effectiveness. These limitations highlight the need for a more comprehensive and mechanism-aware evaluation framework that integrates imputation and forecasting processes.

The difference between this research and previous studies lies in its systematic evaluation of multiple statistical, proximity-based, model-based, and temporal imputation methods across three missing data mechanisms (MCAR, MAR, and MNAR), while simultaneously measuring their impact on forecasting performance using an optimized Bi-LSTM model [14–16]. Unlike prior research, which typically examines imputation and forecasting independently, this study integrates both within a unified evaluation framework. The novelty of this research lies in the development of a comprehensive framework that connects mechanisms for missing data, diverse imputation strategies, and optimization-based Bidirectional Long Short-Term Memory forecasting. By incorporating Particle Swarm Optimization to obtain optimal model configurations, this study provides a more reliable assessment of how imputation quality influences predictive accuracy under incomplete data conditions [17, 18]. This integrated approach enables a deeper understanding of how data preprocessing decisions directly affect downstream predictive performance.

The purpose of this study is to evaluate the effectiveness of various imputation methods, including Mean, Median, Mode, KNN, MICE, and LOCF, on environmental and climate time-series datasets under MCAR, MAR, and MNAR scenarios, and to analyze their influence on forecasting performance using an optimized Bi-LSTM model. Model performance is assessed using Mean Absolute Percentage Error (MAPE), Root Mean Square Error (RMSE), and the Coefficient of Determination (R^2) [19, 20]. The main findings indicate that the choice of imputation method significantly affects forecasting accuracy, underscoring the importance of selecting appropriate strategies based on the characteristics of missing data. Therefore, the proposed framework offers practical guidance for improving prediction reliability in time-series analysis when dealing with incomplete observational data. Overall, this study advances methodological rigor in time-series forecasting under incomplete data conditions.

2. RESEARCH METHOD

This study adopts a structured research methodology to systematically evaluate the effectiveness of various imputation methods on forecasting performance using the Bi-LSTM model. The proposed research framework, illustrated in Figure 1, outlines the study's sequential stages: data collection and preprocessing, followed by missing-data simulation under MCAR, MAR, and MNAR mechanisms. The incomplete datasets are then treated using multiple imputation techniques, including Mean, Median, Mode, KNN, MICE, and LOCF. Subsequently, an optimized Bi-LSTM model using PSO is developed for time-series forecasting. Finally, model performance is evaluated using MAPE, RMSE, and R^2, enabling a fair and systematic comparison across all experimental scenarios.

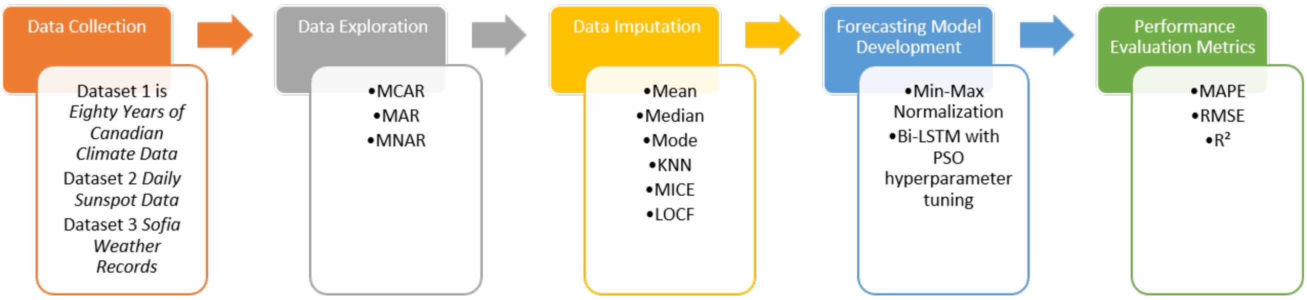

Figure 1. Research Framework

2.1. Data Collection

Data collection for this study was conducted using three climatological datasets from Kaggle. Dataset 1: Eighty Years of Canadian Climate Data; dataset 2: Daily Sunspot Data; dataset 3; Sofia Weather Records. The entire dataset represents long-term environmental and astronomical phenomena, spans long time series, and exhibits complex temporal dynamics. The statistics for each dataset are shown in Table 1.

Table 1. Descriptive Statistics Dataset

Statistics	Dataset 1	Dataset 2	Dataset 3
Number of instances	27,53	73.718	87.599
Average	-0.42°C	79.25	1.647
Standard deviation	12.87	77.47	9.695
Minimum	-48.1°C	-1 (anomaly/invalid value)	-36.9
Quartile 1 (25%)	-8.4°C	15	-3.6
Median (50%)	1.9°C	58	1.7
Quartet 3 (75%)	10.0°C	125	8.3
Maximum	23.9°C	528	29.1
Missing values	1,691	3204	50

From Table 1, Dataset 1 contains 27,530 instances that reflect temperature variations from extreme winter to summer. The target attribute used is MEAN_TEMPERATURE_WHITEHORSE, which represents the daily average temperature and exhibits fairly high seasonal variability. This dataset also contains 1,691 missing values, especially in periods of very low temperatures. The data is real-world and obtained without modification, so it reflects the actual weather recording conditions.

Dataset 2 records the number of daily sunspots from January 11818 to October 31, 2019, for a total of 73,718 instances. The target attribute is the Number of Sunspots because it provides a measure of solar activity that directly affects various scientific and technological phenomena on Earth. This dataset has 3,204 missing data values. The data are historical and collected from various international observatories, making it suitable for time-series-based scientific forecasting studies.

Dataset 3 consists of 87,599 daily air temperature records of the city of Sofia, Bulgaria. Its target attribute is air_temperature, which indicates daily temperature variations with clear seasonal patterns. This dataset contains 50 missing data values. The data comes from official weather records and reflects atmospheric conditions that vary throughout the year. This dataset is used for its rich data structure and is relevant to environmental time-series modeling.

2.2. Data Exploration

The three datasets used have missing values, so it is necessary to identify the type of loss before imputation. This step is important because each type of missing value has distinct characteristics and analytical effects on the time series structure and the forecasting model's performance [18]. The identification helps determine the imputation method that best fits the data loss pattern [19]. A summary of the definitions, causes, and handling for each type of missing value is shown in Table 2.

Table 2. Types of Missing Values

Type of Missing Value	Definition	Main Causes	Impact	General Handling
MCAR	Lost randomly, not affected by any variables	Human error, random recording failure.	Does not cause bias; The analysis remains valid.	Mean, Median, Mode, Drop Missing Values.
MAR	Other variables, not the target, influence the loss	Environmental variables affect the record.	It may be biased if it uses simple imputation.	KNN, MICE, multivariat.
MNAR	The target value itself influences loss.	The sensor fails at the extreme value.	High bias, loss of critical value.	LOCF, model-based, multivariate method.

Based on Table 2, Dataset 1 exhibits an MNAR missing-value pattern, with missing values mainly occurring during extreme temperature periods, so an imputation method is needed that maintains extreme values and temporal continuity. Dataset 2 is classified as MCAR because missing values occur randomly, independent of the target value or other variables. Meanwhile, Dataset 3 includes MAR because the missing values in air_temperature are influenced by other variables and are related to environmental factors or sensor conditions. These three characteristics are the basis for selecting the appropriate imputation method for each dataset.

2.3. Data Imputation

Imputation is the process of filling in missing values so that the dataset is complete and ready for modeling. In time series, this step is crucial because missing values can disrupt the temporal structure and reduce prediction accuracy [21]. Imputation helps preserve important information, reduce bias, and ensure optimal learning models [20]. This study applied several imputation methods: Mean, Median, Mode, KNN, MICE, and LOCF with Drop Missing Values as a comparison. All of these methods represent global, local, statistical, value-proximity, and multivariate approaches, as shown in Table 4.

Table 3. Characteristics of the Imputation Method

Imputation Method	Characteristics	Excess	Deficiency
Mean	Fill in the missing value with the average of all observations.	Simple, fast, suitable for MCAR data.	Damage variance and temporal patterns are not effective for extreme data/MNAR.
Median	Use the distribution's middle value for imputation.	More stable against outliers; good for asymmetrical distribution.	Ignoring temporal patterns is inaccurate for seasonal fluctuations.
Mode	Fill in with the values that appear most often.	Very effective for categorical data or repetitive values.	Not suitable for continuous data; can distort the original distribution.
KNN	Fill in based on the proximity of the value between observations.	Keeping local patterns, suitable for MAR and time series.	High compute time; sensitive to k parameters.
MICE	Use an iterative multivariate model to predict lost values.	Maintain inter-feature correlation; accurate for MAR.	Complex, requires many iterations, risks overfitting.
LOCF	Populate with the previous value in the time series.	Maintain temporal continuity; good for time-series data.	Unable to capture trend changes; bias in the event of a spike in value.
Drop Missing Values	Deletes all rows that contain missing values.	Simple, it does not give rise to new assumptions.	Loss of information; not suitable for datasets with>5% missingness, as it risks deleting important patterns.

Table 4 shows that each imputation method has its own advantages and limitations. Mean works well on MCAR data but can undermine variance and temporal patterns. The median is more robust to outliers, but it still ignores temporal dynamics. Modes are suitable for categorical data, but are less suitable for continuous data because they can distort the distribution. KNN is effective for MAR because it preserves local context, even though the computation is intensive and sensitive to the choice of k. LOCF maintains the continuity of the time series even though it cannot capture trend changes. Meanwhile, Drop Missing Values is the simplest but can lead to the loss of important information when the data is quite large.

2.4. Data Imputation

Imputation is the process of filling in missing values so that the dataset is complete and ready for modeling. In time series, this step is crucial because missing values can disrupt the temporal structure and reduce prediction accuracy [21]. Imputation helps preserve important information, reduce bias, and ensure optimal learning models [20]. This study applied several imputation methods: Mean, Median, Mode, KNN, MICE, and LOCF with Drop Missing Values as a comparison. All of these methods represent global, local, statistical, value-proximity, and multivariate approaches, as shown in Table 4.

Table 4. Characteristics of the Imputation Method

Imputation Method	Characteristics	Excess	Deficiency
Mean	Fill in the missing value with the average of all observations.	Simple, fast, suitable for MCAR data.	Damage variance and temporal patterns are not effective for extreme data/MNAR.
Median	Use the distribution's middle value for imputation.	More stable against outliers; good for asymmetrical distribution.	Ignoring temporal patterns is inaccurate for seasonal fluctuations.
Mode	Fill in with the values that appear most often.	Very effective for categorical data or repetitive values.	Not suitable for continuous data; can distort the original distribution.
KNN	Fill in based on the proximity of the value between observations.	Keeping local patterns, suitable for MAR and time series.	High compute time; sensitive to k parameters.
MICE	Use an iterative multivariate model to predict lost values.	Maintain inter-feature correlation; accurate for MAR.	Complex, requires many iterations, risks overfitting.
LOCF	Populate with the previous value in the time series.	Maintain temporal continuity; good for time-series data.	Unable to capture trend changes; bias in the event of a spike in value.
Drop Missing Values	Deletes all rows that contain missing values.	Simple, it does not give rise to new assumptions.	Loss of information; not suitable for datasets with>5% missingness, as it risks deleting important patterns.

Table 4 shows that each imputation method has its own advantages and limitations. Mean works well on MCAR data but can undermine variance and temporal patterns. The median is more robust to outliers, but it still ignores temporal dynamics. Modes are suitable for categorical data, but are less suitable for continuous data because they can distort the distribution. KNN is effective for MAR because it preserves local context, even though the computation is intensive and sensitive to the choice of k. LOCF maintains the continuity of the time series even though it cannot capture trend changes. Meanwhile, Drop Missing Values is the simplest but can lead to the loss of important information when the data is quite large.

2.5. Forecasting Model Development

The forecasting process in this study used Bi-LSTM, with data normalization as the initial step after imputation. Normalization is necessary to ensure all values are on a uniform scale, making model training more stable and temporal patterns easier to learn [22]. Without normalization, differences in scale between values can inhibit convergence and lower prediction accuracy [23]. This study uses Min–Max Normalization because it preserves the data distribution while scaling the data to 0–1. The normalization formula is shown in Equation 1.

$$X_{new} = \frac{X_{old} - X_{min}}{X_{max} - X_{min}} \tag{1}$$

Equation 1 represents the Min–Max Normalization, which rescales the original value x_old into a new value x_new within the range 0–1. The transformation is performed by subtracting the minimum value of the feature and dividing it by the range $(X_{max} - X_{min})$. This ensures that all values are mapped proportionally while preserving the original data distribution.

After normalization, forecasting is performed using a Bi-LSTM. This model reads the sequence of data in both directions, enabling it to capture both short- and long-term temporal patterns more effectively than one-way LSTMs [24]. This advantage makes Bi-LSTM more suitable for time series with seasonal patterns and complex fluctuations [25]. The Bi-LSTM architecture is shown in Figure 1.

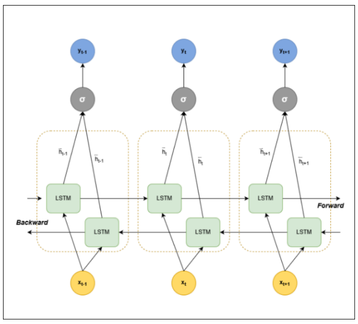

Figure 2. Bi-LSTM Architecture

Based on Figure 2, the Bi-LSTM architecture integrates bidirectional processing to generate richer temporal representations, enabling the model to capture better periodic patterns and high variability in environmental and astronomical data. This capability contributes to more stable and accurate predictions. To ensure optimal model performance, hyperparameter optimization is conducted using PSO to identify the best Bi-LSTM configuration for each dataset. To provide methodological justification, this study considers several commonly used hyperparameter optimization approaches, including Grid Search, Random Search, Bayesian Optimization, and Particle Swarm Optimization. A conceptual comparison of these methods is presented in Table 5.

Table 5. Conceptual Comparison of Hyperparameter Optimization

Method	**Search Strategy**	**Exploration Capability**	**Computational Cost**	**Strengths**	**Limitations**
Grid Search	Exhaustive search over a predefined grid	Limited to predefined combinations	Very High	Systematic and reproducible	Computationally expensive for large search spaces
Random Search	Random sampling of hyperparameters	Moderate	Moderate	Faster than grid search	Performance may vary across runs
Bayesian Optimization	Sequential model-based optimization	High (guided exploration)	High	Efficient convergence in complex spaces	Computational overhead and model dependency
Particle Swarm Optimization	Population-based global search	High (collective adaptive exploration)	Moderate	Balances global exploration and local exploitation	Sensitive to swarm configuration parameters

As shown in Table 5, Grid Search provides systematic coverage but becomes computationally impractical for high-dimensional hyperparameter spaces. Random Search reduces computational burden but lacks adaptive guidance toward promising regions. Bayesian Optimization offers guided search through surrogate modeling; however, its sequential updates increase computational complexity when tuning multiple interdependent parameters.

PSO was selected in this study because it provides a balanced trade-off between exploration capability and computational efficiency. Its population-based global search mechanism enables adaptive exploration of complex hyperparameter spaces while maintaining moderate computational requirements. Therefore, the selection of PSO is based on methodological suitability rather than an untested claim of absolute superiority. The hyperparameter search space is presented in Table 6.

Table 6. PSO Hyperparameter Tuning Search

Hyperparameter	**Value Range/Option**	**Note**
Hidden Layers	02-Oct	Generally < 5 at optimal configuration
Neurons per Layer	1-100	The best compromise between complexity and accuracy
Activation Function	tanh, ReLU, sigmoid	Stable for time-series forecasting
Optimizer	Adam, RMSprop, SGD	Selected PSOs based on convergence
Loss Function	MSE, MAE	Reduce prediction errors
Batch Size	32-64	Stable for large datasets
Epochs	50-100	Avoiding overfitting/underfitting
Dropout	0.0-0.5	Reduces overfitting
PSO Parameters	swarm=10, generations=5, velocity=[0,1], pbest=1.5, gbest=2.0	Controlling the movement of particles

After the PSO explores the hyperparameter space in Table 6, this algorithm generates a different optimal configuration for each dataset. These variations reflect the influence of missing-value characteristics and time-series patterns on the optimal parameter combination. The tuning results are summarized in Tables 7, 8, and 9, which present the final parameters, including the number of neurons, activation functions, dropout rates, and optimizers selected by the PSO. This difference in optimal configuration suggests that the temporal structure and stability of the data distribution greatly influence the PSO search process. Thus, the tuning results table provides evidence that PSOs can customize the Bi-LSTM architecture to the characteristics of each dataset.

The tuning results table shows that each dataset produces a different optimal hyperparameter combination. This difference shows that the missing-value characteristics and temporal patterns of each dataset directly affect the Bi-LSTM architecture's requirements. The variation in the number of layers, neurons, activation functions, and dropout rates reflects the PSO's adaptation to balance the model's complexity and training stability. This confirms that PSOs can customize the configuration to ensure the model works most effectively on the imputed data.

Table 7. Dataset 1 Tuning Results

Parameter	Drop Missing Value	KNN	LOCF	MEAN	MEDIAN	MICE	MODE
Hidden Layer	4	6	2	3	7	6	2
Neurons	43	247	125	39	40	43	219
Activation Function	linear	relu	relu	relu	relu	relu	relu
Optimizer	adam	adam	adam	adam	adam	adam	adam
Loss Function	mse	mse	mse	mse	mse	mse	mse
Epochs	18	37	65	67	40	83	96
Batch Size	16	21	20	16	60	50	39

Table 8. Dataset 2 Tuning Results

Parameter	Drop Missing Value	KNN	LOCF	MEAN	MEDIAN	MICE	MODE
Hidden Layer	7	4	2	6	5	8	3
Neurons	79	246	33	30	252	174	84
Activation Function	linear	relu	tanh	relu	relu	relu	relu
Optimizer	adam	adam	rmsprop	adam	rmsprop	rmsprop	adam
Loss Function	mse	mse	mse	mse	mse	mse	mse
Epochs	45	97	74	33	48	49	13
Batch Size	51	42	51	56	56	22	40

Table 9. Dataset 3 Tuning Results

Parameter	Drop Missing Value	KNN	LOCF	MEAN	MEDIAN	MICE	MODE
Hidden Layer	6	4	7	3	4	8	2
Neurons	161	123	154	253	39	16	186
Activation Function	linear	tanh	linear	relu	linear	relu	relu
Optimizer	rmsprop	adam	rmsprop	adam	rmsprop	adam	adam
Loss Function	mse	mse	mse	mse	mse	mse	mse
Epochs	72	12	72	59	25	22	7
Batch Size	39	53	48	34	27	38	38

2.6. Performance Evaluation Metrics

Model performance was evaluated using three main metrics: accuracy, stability, and the model's ability to explain data variation. The first metric used is MAPE, which provides a percentage measure of prediction error. Equation (2) describes the MAPE calculation as the average of the relative error between the actual value and the prediction.

$$MAPE = \frac{100\%}{n} \sum_{i=0}^{n} \left| \frac{y_i - \hat{y}_i}{y_i} \right| \tag{2}$$

Equation (3) defines RMSE as the root mean square error of the difference between the predicted value and the actual value. Because errors are squared, RMSE becomes sensitive to outliers and major errors. A small RMSE value indicates that the model produces stable predictions across the entire data range.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2} \tag{3}$$

Equation (4) describes R^2 as the proportion of actual data variance that the model can explain. A value close to 1 indicates that the prediction closely follows a temporal pattern. Conversely, a low R^2 value indicates that the model failed to learn the structure of the data variance.

$$R^2 = 1 - \frac{\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}{\sum_{i=1}^{n} (y_i - \bar{y}^2)} \tag{4}$$

3. RESULT AND ANALYSIS

This chapter presents the results of the performance evaluation of the Bi-LSTM model across three datasets with different missing-value characteristics: MNAR, MCAR, and MAR. The discussion compared the Drop Missing Values approach as a baseline with various imputation methods to assess the effect of data reconstruction on forecasting accuracy. The analysis considers not only differences in metric values but also the temporal context, seasonal patterns, and dynamics of change that can affect the temporal structure of each dataset. Thus, the evaluation presented emphasizes the role of the missing-value trait in determining the model's sensitivity to a particular imputation method and the quality of the resulting predictions.

Dataset 1 illustrates an MNAR mechanism in which missing observations are associated with extreme target values. In this scenario, the removal of observations may disproportionately affect critical segments of the time series, potentially altering the temporal structure. The evaluation results for Dataset 1 are presented in Table 10. It should be noted that the analysis was conducted under a fixed proportion of missing data; therefore, the conclusions reflect performance under this specific sparsity level.

Table 10. Dataset Evaluation Results 1

Imputation Method	MAPE	RMSE	R2
Drop Missing Values	263.453	404.659	0.88820
Mean	136.013	419.693	0.87804
Mode	0.78580	480.834	0.84381
Median	0.96151	445.807	0.86239
MICE	129.009	411.897	0.88253
LOCF	0.83314	337.314	0.92318
KNN	124.549	422.770	0.87625

Based on Table 10, the data-deletion approach yielded a MAPE of 2.63453, an RMSE of 4.04659, and an R^2 of 0.8882. Compared to this baseline, several imputation methods improved forecasting performance. The last observation carried forward method achieved the lowest RMSE (3.37314) and the highest RR^2 (0.92318), indicating better preservation of temporal variance structure under the MNAR setting. It also substantially reduced the MAPE to 0.83314. In contrast, distribution-based methods such as mean, median, and mode imputation produced mixed results. While these methods reduced MAPE compared to data deletion, their RMSE and R^2 values remained lower than those obtained using last observation carried forward. The MICE and KNN approaches also showed moderate improvements over data deletion, but did not surpass the performance of last observation carried forward across all metrics. These results suggest that, for the evaluated MNAR proportion, methods that preserve temporal continuity tend to perform better than those that rely solely on distributional averages. However, since the analysis was conducted at a single missing-data proportion, a sensitivity analysis across different sparsity levels would be necessary to determine whether this superiority remains consistent at higher or lower missing rates.

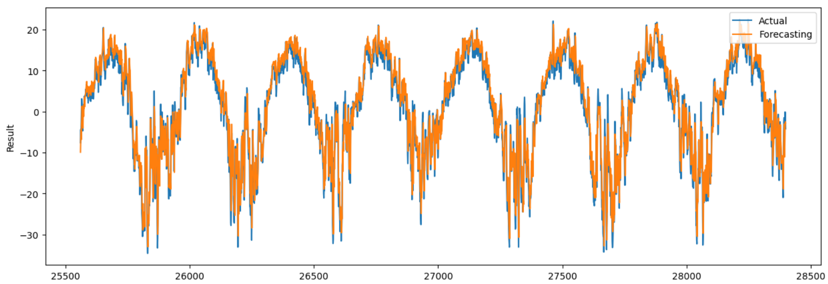

Figure 3. Graph of Bi-LSTM Forecasting Results Using the LOCF Method in Dataset 1

Figure 3 illustrates the forecasting results of the Bi-LSTM model using last-observation-carried-forward imputation for Dataset 1 (MNAR). The visual alignment between the predicted and actual series is supported by quantitative performance metrics, where this configuration achieved an MAPE of 0.83314, an RMSE of 3.37314, and an R^2 of 0.92318 (see Table 10). The relatively low RMSE indicates that the average magnitude of residual deviations remains limited across the time horizon. Furthermore, the high R^2 (0.92318) confirms that more than 92% of the variance in the actual series is explained by the model, supporting the visual observation that the predicted curve follows both periodic cycles and extreme fluctuations. Although minor discrepancies are observable around sharp peaks and troughs, these deviations are consistent with the residual magnitude reflected in the error metrics. Compared with other imputation methods, this configuration yielded the lowest RMSE and the highest R^2 for Dataset 1, quantitatively confirming the visual fit shown in the figure.

Dataset 2 follows the MCAR mechanism, in which missing observations are independent of the target variable and other predictors. Under this condition, data deletion does not, in principle, introduce systematic bias into the distribution of the time series. However, removing observations may still affect temporal continuity depending on the proportion of missing data. The results presented in Table 11 reflect model performance under a fixed proportion of missing data; therefore, interpretations are limited to this evaluated sparsity level.

Table 11. Dataset Evaluation Results 2

Imputation Method	MAPE	RMSE	R2
Drop Missing Values	263.453	404.659	0.88820
Mean	136.013	419.693	0.87804
Mode	0.78580	480.834	0.84381
Median	0.96151	445.807	0.86239
MICE	129.009	411.897	0.88253
LOCF	0.83314	337.314	0.92318
KNN	124.549	422.770	0.87625

Based on Table 11, the data-deletion approach yielded a MAPE of 3.677, an RMSE of 15.11929, and an R^2 of 0.93957. These values indicate that, under MCAR conditions, deleting incomplete observations does not substantially degrade predictive performance. Several imputation methods showed mixed behavior across metrics. Mean and mode imputation resulted in lower MAPE (1.16036 and 1.07173, respectively); however, both methods produced higher RMSE (21.88851 and 24.87812, respectively) and lower R^2 (0.8865 and 0.85338, respectively) than data deletion. This indicates that improvements in relative error do not necessarily correspond to better overall predictive accuracy when evaluated using variance-based metrics. Median and MICE also yielded higher RMSE values than data deletion, suggesting that these approaches did not enhance model performance under the evaluated MCAR setting. In contrast, the last observation carried forward method achieved the RMSE (14.45783) and the highest R^2 (0.95048), while maintaining a MAPE of 3.13373. These results indicate that preserving temporal continuity can yield modest performance gains even when missing values are random. Overall, the findings for Dataset 2 suggest that, at the tested MCAR proportion, the impact of imputation is less pronounced than under structured missing mechanisms. While certain methods improve specific metrics, only the last observation carried forward approach consistently enhances variance explanation and reduces absolute error. Nevertheless, because the analysis was conducted at a single sparsity level, further sensitivity analysis across varying missing-data proportions would be required to determine whether these trends remain consistent at higher or lower levels of data loss.

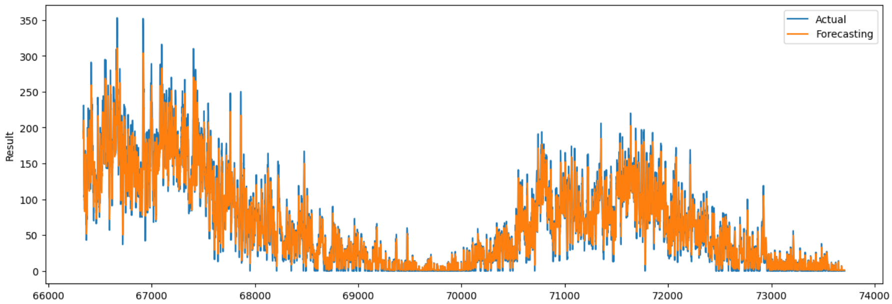

Figure 4. Graph of Bi-LSTM Forecasting Results Using the LOCF Method in Dataset 2

Figure 4 presents the forecasting results of the Bi-LSTM model with last-observation-carried-forward imputation for Dataset 2 (MCAR). The visual alignment between the predicted and actual series is supported by the quantitative metrics reported in Table 11, where this configuration achieved an RMSE of 14.45783 and an R^2 of 0.95048, the highest among all evaluated imputation methods for this dataset. The relatively high R^2 indicates that approximately 95% of the variance in the observed series is explained by the model, confirming that the major oscillatory patterns are effectively captured. Although deviations are evident during high-amplitude peaks, the magnitude of these residuals is reflected in the reported RMSE, which remains lower than that of the data-deletion approach (15.11929). Compared with other imputation strategies, last observation carried forward provides the best balance between absolute error and variance explained in Dataset 2. Therefore, the quantitative metrics corroborate the visual fit shown in Figure 4, indicating that preserving temporal continuity can enhance predictive performance even when missing values are randomly distributed.

Dataset 3 follows the MAR mechanism, in which missing observations depend on other variables but are not directly related to the target value. Under this condition, the distribution of the target variable remains relatively stable despite incomplete observations. The results presented in Table 12 reflect performance under a fixed missing data proportion; therefore, the interpretation is limited to this evaluated sparsity level.

Table 12. Dataset Evaluation Results 3

Imputation Method	**MAPE**	**RMSE**	**R^2**
Drop Missing Values	321.864	133.139	0.98002
Mean	0.32898	111.433	0.98599
Mode	0.35386	120.996	0.98349
Median	0.41440	171.982	0.96663
MICE	107.673	269.188	0.91849
LOCF	0.60395	175.908	0.96511
KNN	0.57502	182.262	0.96289

Based on Table 12, the data-deletion approach yielded a MAPE of 3.21864, an RMSE of 1.33139, and an R^2 of 0.98002. Compared to this baseline, several imputation methods improved forecasting performance across multiple metrics. Mean imputation achieved the lowest MAPE (0.32898) and the lowest RMSE (1.11433), while also producing the highest R^2 (0.98599). Mode imputation yielded comparable performance, with a MAPE of 0.35386 and an R^2 of 0.98349. These results indicate that distribution-based imputation methods can effectively preserve predictive performance under the evaluated MAR condition. Median imputation showed moderate performance, with a higher RMSE (1.71982) and a lower R^2 (0.96663) compared to the mean and mode. Methods such as MICE, last observation carried forward, and KNN produced higher errors and lower coefficients of determination than mean-based imputation, suggesting that more complex or continuity-based approaches did not provide additional benefits in the tested MAR setting. Overall, the findings for Dataset 3 suggest that, at the evaluated proportion of missing data, simple distribution-based imputation methods yield the most favorable balance between error- and variance-based metrics. However, because the analysis was conducted at a single sparsity level, a sensitivity analysis across varying proportions of missing data would be necessary to determine whether these performance patterns remain consistent with varying degrees of data incompleteness.

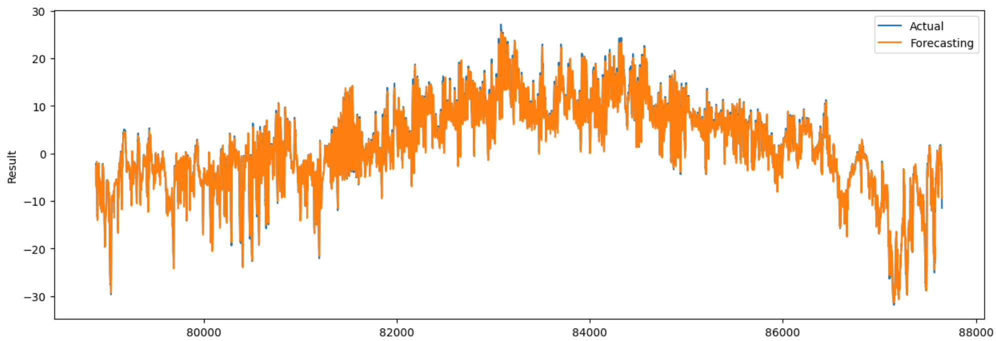

Figure 5. Graph of Bi-LSTM Forecasting Results Using the Mean Method in Dataset 3

Figure 5 illustrates the forecasting results of the Bi-LSTM model using mean imputation for Dataset 3 (MAR). The close visual alignment between the predicted and actual series is supported by the quantitative results reported in Table 11, which show that this configuration achieved the lowest MAPE (0.32898), the lowest RMSE (1.11433), and the highest R^2 (0.98599) among all evaluated imputation methods for this dataset. The high R^2 indicates that approximately 98.6% of the variance in the observed data is explained by the model, confirming that the overall temporal structure is well preserved. The relatively low RMSE reflects limited residual magnitude across the prediction horizon, consistent with the minimal visible deviation between the two curves. Compared with other imputation approaches, mean imputation yields the most consistent improvement across error- and variance-based metrics in Dataset 3. Therefore, the quantitative evaluation substantiates the visual correspondence shown in Figure 4, indicating that under the evaluated MAR condition, distribution-based imputation can effectively support forecasting performance.

Based on the overall evaluation results across the three datasets, the imputation method provides a significant performance improvement over the Drop Missing Values approach for the Bi-LSTM model, especially in Datasets 1 (MNAR) and 3 (MAR). The decrease in MAPE and the consistent increase in R^2 indicate that reconstructing missing values restores the time series' structure, providing the model with a more complete temporal context for analysis. In the MCAR dataset, Drop Missing Values still produces stable performance due to the truly random nature of data loss, but continuity-based methods such as LOCF provide additional improvements. This difference in performance shows that the characteristics of missing values are highly influential on the effectiveness of imputation, so the method selection cannot be generalized across datasets. Thus, imputation is not just a preprocessing step but a crucial element that determines the final quality of a time-series-based deep learning model.

To strengthen the quantitative interpretation of the comparative results, an independent t-test was conducted to assess the statistical significance of imputation versus data deletion. The results indicate that for Dataset 1 and Dataset 3, the obtained p-value was 0.0001, which is below the significance threshold of 0.05. This demonstrates that applying imputation methods significantly affects evaluation metrics and yields statistically meaningful improvements over deleting missing values. However, for Dataset 2, the p-value obtained was 0.6361, which is greater than 0.05. This indicates that there is no statistically significant difference in forecasting performance between the imputed and deleted data. In other words, although certain imputation methods produced numerically better metric values, the improvement cannot be considered statistically significant. This suggests that Dataset 2 is relatively robust to missing-value handling and that the impact of imputation on forecasting accuracy is less pronounced than in the other datasets.

To address the limitation of using a single deep learning architecture, this study extends the analysis by comparing the forecasting performance of Bi-LSTM with other recurrent-based deep learning models, namely Long Short-Term Memory, Gated Recurrent Unit, and Recurrent Neural Network. The comparative results for each dataset are presented in Table 13.

Table 13. Comparison Evaluation Results

Dataset	Method	MAPE	RMSE	R^2
Dataset 1	Bi-LSTM	0.83314	337.314	0.92318
	GRU	0.92128	363.496	0.91079
	RNN	207.995	723.555	0.64654
	LSTM	0.90747	370.117	0.90648
Dataset 2	Bi-LSTM	313.373	1.445.783	0.95048
	GRU	327.918	1.807.289	0.84806
	RNN	700.709	1.529.486	0.88033
	LSTM	661.441	1.599.668	0.91919
Dataset 3	Bi-LSTM	0.32898	111.433	0.98599
	GRU	0.49631	136.801	0.97889
	RNN	285.135	943.251	0.14101
	LSTM	0.85895	130.905	0.98084

The primary objective of this study is to develop a systematic mapping between missing-value mechanisms (MNAR, MCAR, MAR) and their corresponding optimal imputation strategies for deep learning-based time-series forecasting. The results consistently demonstrate that Bi-LSTM achieves superior performance across all datasets. In Dataset 1, Bi-LSTM achieved the lowest MAPE (0.83314) and RMSE (3.37314), and the highest R^2 (0.92318), outperforming LSTM, GRU, and RNN. A similar pattern is observed in Dataset 2, where Bi-LSTM achieved an R^2 of 0.95048, higher than that of all the other architectures compared. The superiority becomes even more pronounced in Dataset 3, where Bi-LSTM produced the lowest MAPE (0.32898) and the R^2 (0.98599).

The empirical evidence indicates that the effectiveness of imputation strategies remains stable across recurrent-based deep learning architectures, with Bi-LSTM exhibiting comparatively stronger temporal representation capability than unidirectional models. This consistency suggests that the observed improvements are not architecture-specific but instead stem from the interaction between missing-value mechanisms and the reconstruction strategy applied prior to model training. Accordingly, the results directly address the identified research gap by showing that forecasting performance is shaped not merely by model complexity but by the alignment between missingness characteristics and imputation approach. In other words, predictive stability depends on mechanism-aware data reconstruction rather than on architectural sophistication alone.

The findings of this study also show that each type of missing value has the most optimal imputation method based on its statistical patterns and temporal dynamics. LOCF proved to be most effective in the MNAR dataset because it was able to maintain the continuity of the lost extreme values; MCAR is not very sensitive to method choice because data loss has no pattern; while Mean is the best method on the MAR dataset because the data distribution remains stable even though some values are lost. This success confirms that imputation must account for the mechanism of data loss to avoid bias or distortion of time-series patterns. Overall, this study strengthens the understanding that the success of forecasting is determined not solely by the Bi-LSTM model but also by the quality of the imputed data, which serves as the main foundation.

This study has several limitations that should be considered when interpreting the results. First, all experiments were conducted using univariate time series models. Consequently, multivariate imputation methods such as MICE cannot fully exploit inter-variable correlations, which are their primary strength. In the absence of additional explanatory variables, MICE operated under constrained univariate conditions; therefore, its performance in this study does not reflect its full multivariate capability. Second, although several recurrent-based architectures (LSTM, GRU, and RNN) were included for comparison, the study did not evaluate more advanced sequence models, such as Temporal Convolutional Networks or Transformer-based forecasters. As a result, architectural generalization remains limited to recurrent neural network families. Third, the proportion of missing values was evaluated at a fixed sparsity level, and a systematic sensitivity analysis across multiple missing-data ratios was not conducted. Therefore, the robustness of each imputation method under varying levels of data incompleteness remains an open research question. Fourth, the study did not simulate more complex MNAR mechanisms, such as threshold-based measurement failures, sensor saturation effects, or physically constrained observational gaps, which may alter the statistical distribution and temporal continuity of the series. Finally, although three publicly available datasets were used to represent MNAR, MCAR, and MAR mechanisms, all datasets originated in the environmental and climatological domains. This domain consistency was intentionally maintained to control for structural differences in temporal dynamics and in the physical processes that generate data. However, it limits the direct generalizability of the findings to other time-series contexts such as financial markets, industrial monitoring, or medical signal analysis, which may exhibit distinct volatility patterns, regime shifts, or irregular sampling structures. Accordingly, the conclusions of this study should be interpreted within structured environmental time-series settings.

Despite these limitations, the findings make meaningful contributions to climatology and environmental data science, particularly in handling missing values in environmental monitoring systems. The results demonstrate that the effectiveness of an imputation strategy depends strongly on the underlying mechanism generating missing values. Continuity-based approaches, such as LOCF, show advantages under MNAR conditions; distribution-based methods, such as mean imputation, perform effectively under MAR settings; and MCAR scenarios exhibit relatively lower sensitivity to method selection. These findings offer practical guidance for designing more reliable pipelines for preprocessing climatological data. Moreover, the consistent improvement in Bi-LSTM forecasting performance after appropriate imputation underscores that forecasting accuracy is not determined solely by model architecture but also by the quality of data reconstruction before training. This reinforces the importance of mechanism-aware preprocessing strategies over generic imputation techniques. Overall, this study provides an evaluation-based framework for selecting imputation methods aligned with missing-value mechanisms in environmental time-series forecasting. Future research extending this framework to multivariate datasets and heterogeneous domains would further validate and broaden the applicability of the proposed approach.

4. CONCLUSION

The experimental results demonstrate that forecasting performance is strongly influenced by the interaction between imputation strategy and temporal modeling rather than by model architecture alone. Across all datasets, Bi-LSTM consistently achieved superior accuracy, indicating its stronger ability to capture bidirectional temporal dependencies. In Dataset 1, Bi-LSTM produced a MAPE of 0.83314, an RMSE of 3.37314, and an R^2 of 0.92318. Similar robustness was observed in Dataset 2, where Bi-LSTM achieved an RMSE of 14.45783 and an R^2 of 0.95048. In contrast, Dataset 3 exhibited the highest predictive stability, with a MAPE of 0.32898, an RMSE of 1.11433, and an R^2 of 0.98599. These numerical outcomes confirm that the forecasting models effectively learned temporal patterns when supported by appropriate data reconstruction.

Beyond model performance, the results reveal a clear mapping between missing value mechanisms and optimal imputation strategies. Continuity-preserving methods, such as LOCF, demonstrated superior effectiveness under MNAR conditions by maintaining extreme-value trajectories. In contrast, mean imputation performed optimally under MAR scenarios, preserving distributional stability. In contrast, MCAR scenarios showed relatively low sensitivity to the choice of imputation method. This mechanism-specific behavior explains the observed improvements in Bi-LSTM performance across datasets. It highlights that predictive stability arises from alignment between the characteristics of missingness and the imputation strategy, rather than from increasing architectural complexity.

These findings reinforce the importance of mechanism-aware preprocessing for environmental time-series forecasting and provide a structured framework for selecting imputation methods based on data-loss characteristics. Although the current analysis is limited to univariate environmental datasets, the consistent performance gains suggest strong potential for extension to multivariate settings and other time-series domains with different statistical properties. Future research may expand this framework by incorporating heterogeneous variables, evaluating multiple missing-data ratios, and exploring advanced sequence models to further generalize the proposed approach's applicability across a broader range of forecasting contexts.

5. ACKNOWLEDGEMENTS

The authors sincerely thank Anna Turner, Abhinand, and NikiTricky for publicly sharing the Eighty Years of Canadian Climate Data, Daily Sunspot Data, and Sofia Weather Records datasets through Kaggle. Their efforts in curating and disseminating these datasets have significantly supported the empirical validation of this study.

6. DECLARATIONS

AI USAGE STATEMENT

The authors state that Artificial Intelligence tools, such as ChatGPT developed by OpenAI, were employed to assist with language polishing, grammatical corrections, and paraphrasing during the manuscript preparation. The authors affirm that all concepts, data analyses, and conclusions presented in this work are entirely their own and were not produced by the AI tool.

AUTHOR CONTIBUTION

Adelia Desyana Eka Putri: original draft, formal analysis, investigation, writing - review and editing, Aji Prasetya Wibawa: writing - review and editing, supervision, project administration, Adelia Khansa Ristiaputri: original draft, visualization, investigation, writing - review and editing, Adhelia Wida Khaidir: original draft, software, investigation, writing - review and editing, Dhia Rafifah Thifal: original draft, data collection, investigation, writing - review and editing, Agung Bella Putra Utama: formal analysis, writing - review and editing.

FUNDING STATEMENT

This research was carried out without financial support.

COMPETING INTEREST

The authors declare no conflict of interest.

REFERENCES

[1] C. Betancourt, C. W. Li, F. Kleinert, and M. G. Schultz, "Graph Machine Learning for Improved Imputation of Missing Tropospheric Ozone Data," *Environmental Science and Technology*, vol. 57, no. 46, pp. 18 246–18 258, 2023, https://doi.org/10.1021/acs.est.3c05104.

[2] C. Xie, C. Huang, D. Zhang, and W. He, "Bilstm-i: A deep learning-based long interval gap-filling method for meteorological observation data," *International Journal of Environmental Research and Public Health*, vol. 18, no. 19, pp. 1–12, 2021, https://doi.org/10.3390/ijerph181910321.

[3] V. Hua, T. Nguyen, M.-S. Dao, H. D. Nguyen, and B. T. Nguyen, "The impact of data imputation on air quality prediction problem," *PLOS ONE*, vol. 19, no. 9, pp. 1–39, 09 2024. [Online]. Available: https://doi.org/10.1371/journal.pone.0306303

[4] T. Kim, J. Kim, W. Yang, H. Lee, and J. Choo, "Missing value imputation of time-series air-quality data via deep neural networks," *International Journal of Environmental Research and Public Health*, vol. 18, no. 22, pp. 1–8, 2021, https://doi.org/10.3390/ijerph182212213.

[5] C. Wongoutong, "Performance Comparison of Various Imputation Methods for Missing Data Mechanisms (MAR , MCAR , and MNAR) in a Nonstationary Time Series," *International Journal of Mathematics and Mathematical Sciences*, vol. 25, no. 1, pp. 1–16, October, 2025, https://doi.org/10.1155/ijmm/3031708.

[6] W. Alahamade, I. Lake, C. E. Reeves, and B. De La Iglesia, "Evaluation of multivariate time series clustering for imputation of air pollution data," *Geoscientific Instrumentation, Methods and Data Systems*, vol. 10, no. 2, pp. 265–285, 2021, https://doi.org/10.5194/gi-10-265-2021.

[7] A. T. Prihatno, H. Nurcahyanto, F. Ahmed, and H. Rahman, "Forecasting PM2 . 5 Concentration Using a Single-Dense Layer BiLSTM Method," *Electronics*, vol. 10, no. 15, p. 1808, 2021, https://doi.org/10.3390/electronics10151808.

[8] A. S. Firmansyah and A. T. Putra, "Implementation of Bidirectional Long-Short Term Memory (Bi- LSTM) and Attention to Detect Political Fake News Using IndoBERT and GloVe Embedding," *Recursive Journal of Informatics*, vol. 3, no. 2, pp. 68–76, 2025, https://doi.org/10.15294/rji.v3i2.159.

[9] J. F. Ruma, M. S. G. Adnan, A. Dewan, and R. M. Rahman, "Particle swarm optimization based LSTM networks for water level forecasting: A case study on Bangladesh river network," *Results in Engineering*, vol. 17, no. July 2022, p. 100951, 2023, https://doi.org/10.1016/j.rineng.2023.100951.

[10] K. J. Lee, J. B. Carlin, J. A. Simpson, and M. Moreno-betancur, "Assumptions and analysis planning in studies with missing data in multiple variables: moving beyond the MCAR/MAR/MNAR classification," *International Journal of Epidemiology*, vol. 52, no. 4, pp. 1268–1275, 2023, https://doi.org/10.1093/ije/dyad008.

[11] A. A. R. Khattab, N. M. Elshennawy, and M. Fahmy, "GMA: Gap Imputing Algorithm for time series missing values," *Journal of Electrical Systems and Information Technology*, vol. 10, no. 1, pp. 1–20, 2023, https://doi.org/10.1186/s43067-023-00094-1.

[12] J. Wang, Z. Yan, T. Pan, Z. Zhu, X. Song, and D. Yang, "applied sciences Drilling Parameters Multi-Objective Optimization Method Based on PSO-Bi-LSTM," *Applied Sciences*, vol. 13, no. 21, p. 11666, 2023, https://doi.org/10.3390/app132111666.

[13] A. Mechanisms, "BiLSTM-MLAM : A Multi-Scale Time Series Prediction Model for Sensor Data Based on Bi-LSTM and Local Attention Mechanisms," *Sensors*, vol. 24, no. 12, p. 3962, 2024, https://doi.org/10.3390/s24123962.

[14] M. A. Helaly, S. Rady, M. Mabrouk, M. M. Aref, and S. Villarroya, "Advancements in water quality prediction : a practical review of machine learning and deep learning approaches," *Cluster Computing*, vol. 28, no. 9, pp. 1–17, 2025, https://doi.org/10.1007/s10586-025-05221-3.

[15] N. Ahmad and V. Kumar, "Enhancing PM 2 . 5 Air Pollution Forecasting with Novel Random Imputation Based on Hybrid RNN - Bidirectional GRU (nRI RNN - BiGRU) Model," *SN Computer Science*, vol. 6, no. 2, p. 637, 2025, https://doi.org/10.1007/s42979-025-04167-y.

[16] V. V. Guggilam and G. Sundaram, "Spatio-Temporal Bi-LSTM Based Variational Auto-Encoder for Multivariate IoT Data Imputation," *International Journal of Intelligent Engineering and Systems*, vol. 17, no. 3, pp. 352–367, 2024, https://doi.org/10.22266/ijies2024.0630.28.

[17] S. Surono, K. W. Goh, C. W. Onn, A. Nurraihan, N. S. Siregar, A. B. Saeid, and T. T. Wijaya, "Optimization of Markov Weighted Fuzzy Time Series Forecasting Using Genetic Algorithm (GA) and Particle Swarm Optimization (PSO)," *Emerging Science Journal*, vol. 6, no. 6, pp. 1375–1393, 2022, https://doi.org/10.28991/ESJ-2022-06-06-010.

[18] J. H. Alkhateeb, "Fine Tuning Hyperparameters of Deep Learning Models Using Metaheuristic Accelerated Particle Swarm Optimization Algorithm," *IEEE Access*, vol. 13, no. July, pp. 134 506–134 518, 2025, https://doi.org/10.1109/ACCESS.2025.3591403.

[19] D. Chicco, M. J. Warrens, and G. Jurman, "The coefficient of determination R-squared is more informative than SMAPE , MAE , MAPE , MSE and RMSE in regression analysis evaluation," *PeerJ Computer Science*, vol. 7, no. e623, pp. 1–24, 2021, https://doi.org/10.7717/peerj-cs.623.

[20] M. Chen, H. Zhu, Y. Chen, and Y. Wang, "A Novel Missing Data Imputation Approach for Time Series Air Quality Data Based on Logistic Regression," *Atmosphere*, vol. 13, no. 7, p. 1044, jun 2022, https://doi.org/10.3390/atmos13071044.

[21] A. Galdelli, G. Narang, S. Tomassini, L. D. Agostino, A. N. Tassetti, and A. Mancini, "Data imputation in large and small - scale spatiotemporal time series gaps using BackForward Bi - LSTM," *Journal of Big Data*, vol. 12, p. 115, May, 2025, https://doi.org/10.1186/s40537-025-01163-0.

[22] J. G. Kim, S. Y. Lee, and I. B. Lee, "The Development of an LSTM Model to Predict Time Series Missing Data of Air Temperature inside Fattening Pig Houses," *Agriculture (Switzerland)*, vol. 13, no. 4, pp. 1–18, 2023, https://doi.org/10.3390/agriculture13040795.

[23] H. Karnati, A. Soma, and A. Alam, “Comprehensive analysis of various imputation and forecasting models for predicting PM2 . 5 pollutant in Delhi,” *Neural Computing and Applications*, vol. 37, no. 17, pp. 11 441–11 458, 2025, https://doi.org/10.1007/s00521-025-11047-2.

[24] K. Tzoumpas, “A Data Filling Methodology for Time Series Based on CNN and (Bi) LSTM Neural Networks,” *IEEE Access*, vol. 12, pp. 31 443–31 460, February, 2024, https://doi.org/10.1109/ACCESS.2024.3369891.

[25] G. B. Y. A. Kara, E. Pekel, E. Ozcetin, “Genetic algorithm optimized a deep learning method with attention mechanism for soil moisture prediction,” *Neural Computing and Applications*, vol. 7, no. 36, pp. 1761–1772, 2024, https://doi.org/10.1007/s00521-023-09168-7.

[This page intentionally left blank.]