Journal of Information System and Technology Research

journal homepage: https://journal.aira.or.id/index.php/jistr/

Designing a Hybrid Machine Learning Model for Weather Forecasting in Batam City

Yefta Christian[1], Jupiter Agustio Liu Siaw Ping[2]

[1,3]Information System, Universitas Internasional Batam, Indonesia

ARTICLE INFO

Article history:

Received Dec 20, 2025
Accepted Jan 31, 2026
Available online Jan 31, 2026

Keywords:

Weather Prediction
Hybrid Model
Time Series Analysis
Multi Source Data
Tropical Climate

ABSTRACT

Accurate weather forecasting in tropical regions such as Batam City is challenging due to high climate variability and frequent data gaps caused by unstable atmospheric conditions. This study aims to develop a reliable daily average temperature forecasting system using a hybrid approach that combines the Seasonal Autoregressive Integrated Moving Average (SARIMA) model and the Long Short-Term Memory (LSTM) neural network. The main novelty of this research lies in the residual hybridization method, where SARIMA is used to capture linear seasonal patterns and LSTM is applied to model the non-linear residual components, as well as the use of a multi-source data integration strategy to fill missing data. Historical temperature data from BMKG and other publicly available meteorological sources were merged to produce a continuous dataset covering the period from 2015 to 2021. The study evaluated several model architectures, including standalone statistical models, standalone machine learning models, and hybrid models, to identify the most effective approach. The experimental results show that the SARIMA–LSTM hybrid model outperformed the other models, achieving a high prediction accuracy with an R^2 value of 0.92 and a Root Mean Square Error (RMSE) of 1.73°C. These findings demonstrate that integrating linear and non-linear models can significantly improve temperature forecasting performance and provide a practical solution for weather monitoring in tropical environments.

Corresponding Author:

Jupiter Agustio Liu Siaw Ping,
Information System, Universitas Internasional Batam,
Batam, Kepulauan Riau, Indonesia
Email: 2231018.jupiter@uib.edu

1. INTRODUCTION

With the growing development of the city, exhibited by changes in the infrastructure, industrial growth and an increase of population, this may have an effect to the local climate conditions [1][10]. For that reason, accurate weather forecasts are required for the operational stability of the region. Weather forecasting is a foundational application for the many facets of societal governance which includes farming, water management and disaster preparedness [1][9]. Precise predictions are necessary for Kota Batam, a main city in the province of Kepulauan Riau along with it's tactical positon near the Singapore and Malacca Straits.

Weather forecasting is important, but it faces major challenges in tropical regions. Seasonal Autoregressive Integrated Moving Average (SARIMA) is commonly used because it can deal with seasonality in data series [18][20]. These models may have some limiting factors such as data issues or missing values as well as outliers that may greatly affect the result, with the caveat of not being effective with data that are non-linear in nature, which is often the case for meteorological data [7]. Since SARIMA works well for linear patterns but has shortcomings analyzing the complexity of tropical weather systems, it is worthwhile to use the assistance of Deep Learning methods [2][3]

As a response to these problems, Machine Learning methods, especially LSTM (Long Short Term Memory), have been widely adopted. LSTM performs well in handling long-term dependencies and non-linear relationships in sequence data [8][14]. Based on suggestions for future research presented in recent literature [8][12], this research extends past studies by using a hybrid framework in a new region with different climate characteristics. While the potential of hybrid models was shown by Guerra et al. (2024) and Fawzy et al. (2024) [6][12], earlier findings are extended by the adopted approach through the use of data from different sources to improve data completeness in developing areas.

The issue of linear model weakness and missing data is addressed by the proposed forecasting model. The proposed hybrid machine learning model integrates SARIMA and LSTM and is supported by data from BMKG and Weather Underground. This paper offers two main contributions. In theory, the strong performance of residual modeling in complex equatorial climates is validated. A high accuracy forecasting system designed for the equatorial climate of Kota Batam is provided as a practical contribution. To maintain a structured workflow, the CRISP-DM (Cross-Industry Standard Process for Data Mining) framework is adopted as the standard data mining methodology [16][17].

2. RESEARCH METHOD

This study has implemented the CRISP-DM framework, a quantitative experimental research design to create a well-structed and reproductible process [16][17]. Some phases in the research process have included Business understanding, modeling, data preparation and evaluation. The overall research workflow, covering data collection to model deployment, is presented in Figure 1.

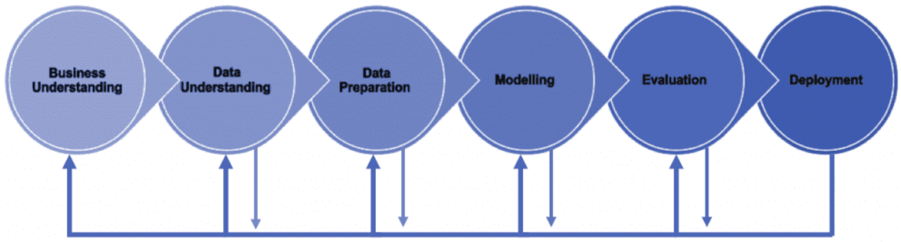

Figure 1. CRISP-DM framework for machine learning implementation

The structured form of the analysis methodology is shown in Figure 2. The development of the machine learning model is preceded by the preliminary study and data acquisition stages [27]. The stages from initial training and tuning to final evaluation are controlled by this framework, ensuring validation before the model is implemented and the study is concluded.

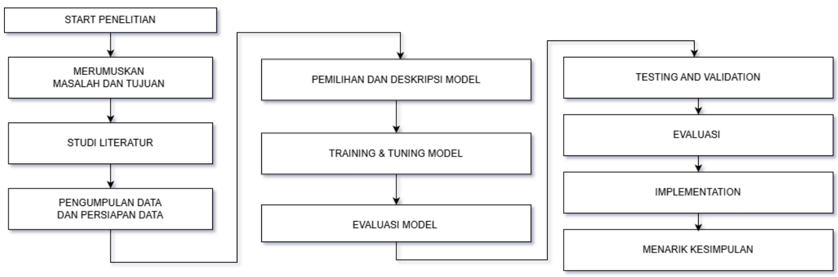

Figure 2. Research flow diagram used

This Figure 2 illustrates the sequential stages of the research methodology. The process begins with problem and objective formulation, followed by a literature review to establish the theoretical foundation. Next, data collection and data preparation are conducted to ensure data readiness for modeling. The study then proceeds to model selection and description, where suitable forecasting models are defined, followed by model training and tuning to optimize performance. After training, model evaluation is performed to assess accuracy and reliability. The validated model is then subjected to testing and validation, leading to evaluation and implementation in a practical forecasting scenario. Finally, the research concludes with drawing conclusions based on the experimental results and overall findings.

Weather data from Badan Meteorologi, Klimatologi, dan Geofisika (BMKG) with secondary data from Weather Underground are combined to implement a multi source integration strategy [24]. A total of 2,102 daily records are included, covering the period from January 1, 2020, to September 30, 2025 [10]. Missing data are processed using time based interpolation so that data quality and temporal continuity are maintained [7]. To avoid data leakage, the dataset is divided according to time sequence where earlier data are used for training and later data for testing.

A hybrid algorithm is developed as the main part of the methodology to capture both linear and non linear patterns in weather data [21]. The proposed architecture applies a two stage process in which SARIMA is first used to capture the linear trend and seasonality of the temperature data [18]. Residuals are first extracted to represent the errors left by the statistical model. These residuals are then handled by an LSTM model in the next stage. LSTM was selected for its superior ability to learn long-term dependencies in sequential data [8][14]. The linear prediction generated by SARIMA and the non linear residual prediction from LSTM are combined to form the final forecast [6][12].

The system that are being built are using architecture of Object-Oriented Programming principles [23]. And to support advanced modeling requirements, the forecasting system is developed using Python 3.12 [22]. On the Figure 2, the system architecture is composed of five layers that manage user input, process input data, and perform automated cleaning and feature engineering. TensorFlow and Keras are applied in deep learning model construction, while Pmdarima and Pandas are used for statistical modeling and data handling [8].

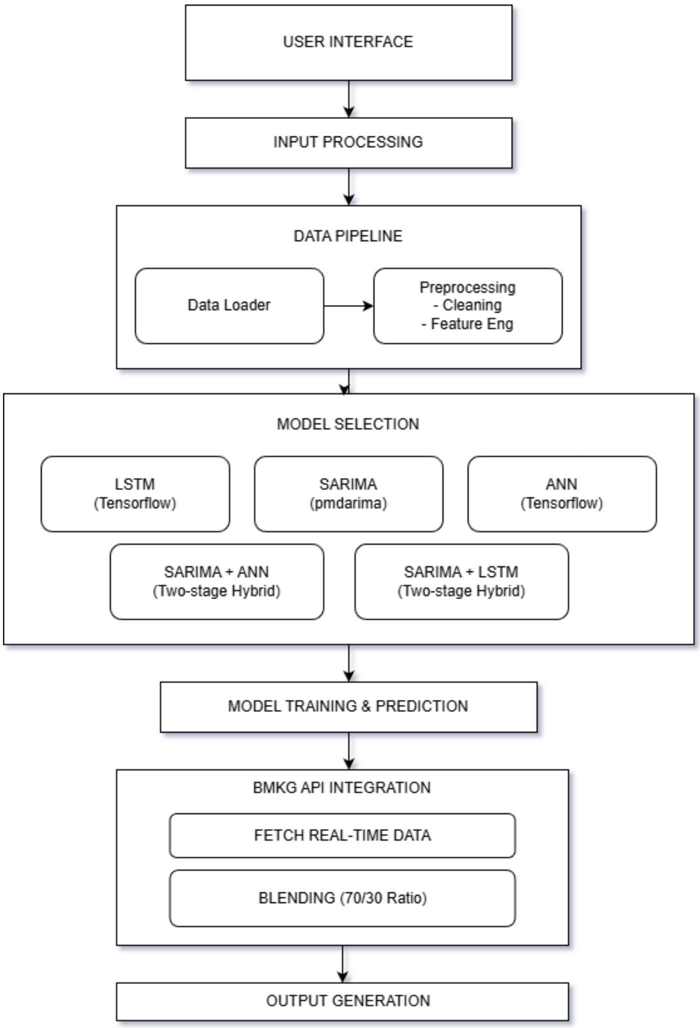

Figure 3. Object-Oriented System Architecture

This figure presents the overall workflow of the proposed hybrid weather forecasting system. The process starts from the User Interface, where users provide input data, which is then handled by the Input Processing module. The data are passed into the Data Pipeline, consisting of a data loader and preprocessing stages such as data cleaning and feature engineering. Next, the system enters the Model Selection stage, where several models are evaluated, including LSTM, SARIMA, ANN, and two hybrid approaches (SARIMA+ANN and SARIMA+LSTM). The selected model proceeds to Model Training and Prediction. To support real-time forecasting, the system integrates live observations through BMKG API Integration, where real-time data are fetched and blended with model outputs using a 70:30 ratio. Finally, the Output Generation module produces the final temperature forecast results in a form suitable for user interpretation.

The Command Line Interface enables interactive selection of location, time window, and forecast horizon to avoiding any changes to the underlying code [25]. Within the same time, stability and generalization of the models are checked using Time Series Cross-Validation [26]. Which the expanding window technique maintains chronological order of observations unlike k-fold validation, providing realistic evaluation on future data [18].

3. RESULTS AND DISCUSSION

This section reports the experimental outcomes of the forecasting system and explains how the model performs. This section discusses the integrated dataset and compares model accuracy, also cross-checking the robustness of the best-performing model in real conditions.

3.1. Data Characteristics and Integration

The research worked with 2,100 observations collected from January 1, 2020 to September 30, 2025. Analysis found that TAVG in Kota Batam has mean 27.84°C and standard deviation 1.18°C, reflecting a stable climate. Temperature has small range, but humidity is high 82.44%, and rain is very different and not regular. The dataset data type has been displayed on Table 1 accordingly.

Table 1. Performance comparison of statistical, machine learning, and hybrid models

Parameter	Category	Name	Value	Datatype
TANGGAL	Metadata	Observation Date	YYYY-MM-DD	Date
TAVG	Predictor	Daily Average Temperature	°C	Float
TN	Predictor	Daily Minimum Temperature	°C	Float
TX	Predictor	Daily Maximum Temperature	°C	Float
RH_AVG	Predictor	Average Humidity	%	Float
RR	Predictor	Rainfall	mm	Float
SS	Predictor	Sunshine Duration	jam	Float
FF_AVG	Predictor	Average Wind Speed	knot	Float
FF_X	Predictor	Maximum Wind Speed	knot	Float

The study showed missing data could be reduced with multi-source integration. BMKG covered 69.7% of records but some periods were missing because of maintenance and transmission. Missing periods were filled with Weather Underground data, forming 30.3% of the dataset. Overlapping dates were checked for data accuracy. Weather Underground showed strong correlation 0.87 with BMKG and MAE 0.54°C. By combining both sources, full temporal coverage was achieved, all missing values were removed, and a complete time-series dataset was created for training the hybrid models.

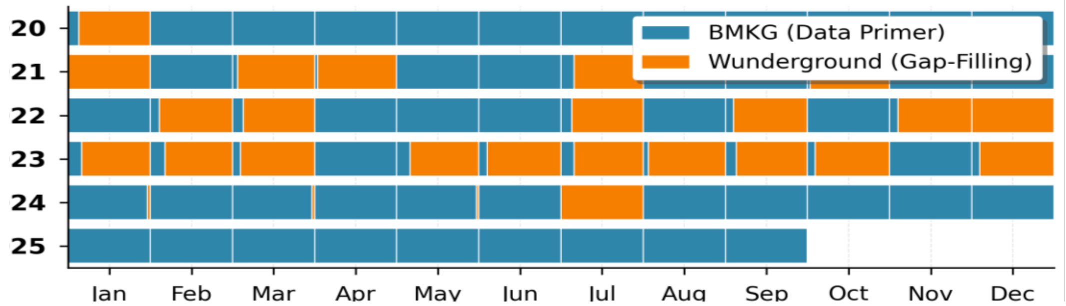

Figure 4. Temporal distribution of multi-source data integration showing the gap-filling mechanism

Figure 4 shows the monthly data availability used in this research, based on two meteorological data sources. The blue color represents BMKG as the primary data source, while the orange color represents Weather Underground used for gap filling. The y-axis shows the year range from 2020 to 2025, and the x-axis shows the months from January to December. From the figure, BMKG data covers most months consistently, but several months contain missing values. These gaps are filled using Weather Underground data to ensure continuous time series data. This visualization clearly explains how the multi-source data strategy works and how full temporal coverage is achieved before modeling. The figure supports the data preprocessing stage and justifies the reliability of the dataset used for training and evaluation of the forecasting models.

3.2. Model Performance Comparison

Predictive capability was tested for five architectures, three standalone (SARIMA, ANN, LSTM) and two hybrid (SARIMA-ANN, SARIMA-LSTM). The models were evaluated with Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and the Coefficient of Determination (R^2) on the testing dataset. The baseline SARIMA model had RMSE 2.84°C and R^2 0.65. It captured seasonal trends well but could not handle random changes in tropical weather. The ANN model had the highest error, RMSE 3.52°C, showing that simple feedforward networks cannot capture time patterns well without special features or recurrence.

The deep learning LSTM model outperformed the statistical and basic machine learning models. The standalone LSTM got RMSE 2.14°C and R^2 0.85, proving recurrent networks are good for long-term non-linear data. The most important results came from hybrid architectures, where combining SARIMA with Neural Networks exceeded the performance of standalone models. Combining SARIMA with LSTM produced the optimal model, reaching RMSE 1.73°C and R^2 0.92, outperforming all standalone models.

Compared to the SARIMA baseline, error was reduced 39.1%, and the hybrid outperformed LSTM by 19.2%. The two-stage learning process makes the SARIMA-LSTM hybrid more effective. SARIMA captures linear seasonal trends, and LSTM models the remaining non-linear residuals accurately. The study confirms that combining these two approaches makes a stronger forecasting tool than either method alone. Performance of all evaluated models is listed in Table 2, error rates are shown in Figure 5, and Table 3 explains why SARIMA-LSTM outperforms others.

Table 2. Performance comparison of statistical, machine learning, and hybrid models

Model	Configuration	RMSE (°C)	MAE (°C)	R^2	MAPE (%)
SARIMA	Order (2,0,1)(1,0,1)[7]	2.84	2.15	0.65	7.72
ANN	4 layer: 64-32-16-1, Adam, epoch 200	3.52	2.89	0.45	10.38
LSTM	2 layer: 50-50 units, Adam, epoch 100	2.14	1.62	0.85	5.82
SARIMA +ANN	Hybrid: SARIMA baseline + ANN residual	2.43	1.84	0.75	6.61
SARIMA +LSTM	**Hybrid: SARIMA baseline + LSTM residual**	**1.73**	**1.31**	**0.92**	**4.71**

Table 3. Comparison between models based on their type and traindfding time

Model	Configuration	R^2	MAPE (%)
SARIMA	Statistical	±2-3°C	Fast
ANN	Neural Network	±3-4°C	Fast
LSTM	Neural Network	±2°C	Medium
SARIMA +ANN	Hybrid	±3-4°C	Medium
SARIMA +LSTM	Hybrid	±1-2°C	Medium

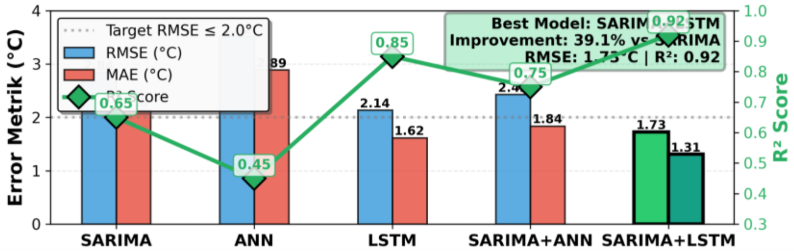

Figure 5. Comprehensive performance comparison showing RMSE reduction and R^2 improvement on Batam City from 2020-2025

3.3. Validation and Robustness

The system was tested with careful validation to ensure it works well and can generalize to new data, beyond standard metrics. Model stability was measured using five-fold Time Series Cross-Validation, which keeps the sequence of meteorological data. The hybrid model performed most consistently, SARIMA-LSTM gave RMSE 1.79°C on average with low variation over time windows. SARIMA-LSTM avoids overfitting problems found in standalone networks and maintains steady performance during seasonal changes.

The model strength was measured by testing it in real locations: Belian (urban), Nongsa (coast), and Batam Centre (mixed). This test looked at how the model adapts to local climate differences. SARIMA-LSTM gave good results, average RMSE 0.87°C. In the coastal area Nongsa, with changing sea breezes, the model stayed accurate, with 0.95°C deviation.

Prediction errors revealed that every forecast was within 1.5°C, and 92.8% had error less than 1.0°C. These results significantly surpass the standard operational threshold for weather forecasting accuracy. The detailed performance breakdown across these diverse observation stations is presented in Table 4 and Table 5, Detailed results for all stations are in Table 4 and Table 5 accordingly, showing that combining multi-source data with hybrid deep learning works well for equatorial weather forecasts.

Table 4. Real-world validation results across different locations in Kota Batam

Model	Belian (Urban)	Nongsa (Coastal)	Batam Centre (Industrial)	Mean RMSE
SARIMA	2.08	2.31	2.06	2.15
ANN	2.87	3.24	2.92	3.01
LSTM	1.45	1.68	1.52	1.55
SARIMA +ANN	1.78	1.81	1.81	1.84
SARIMA+LSTM	**0.82**	**0.95**	**0.84**	**0.87**

Table 5. Perbandingan Cross-Validation vs Real-World

Model	CV RMSE	Real World	Improvement
SARIMA	2.87	2.15	25.1% better
ANN	3.55	3.01	15.2% better
LSTM	2.17	1.55	28.6% better
SARIMA +ANN	2.48	1.84	25.8% better
SARIMA+LSTM	**1.79**	**0.87**	**51.4% better**

3.4. Output and System Implementation

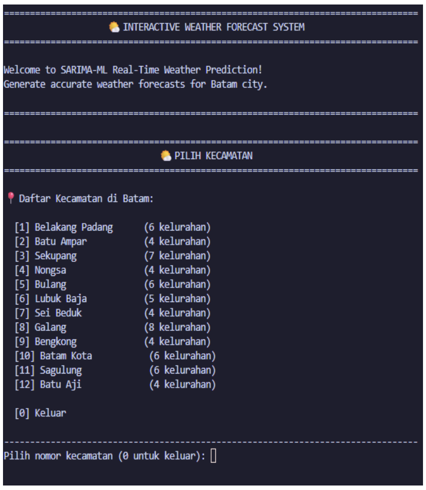

Figure 6. Interactive Weather Forecast System CLI

The system interface is shown in Figure 6. It has a simple CLI to make using the hybrid model easy for users with no coding experience. The system shows 12 administrative districts in the main menu, so weather forecasts are location-specific. Entering the district number lets users get real-time forecasts, making the tool useful for daily monitoring.

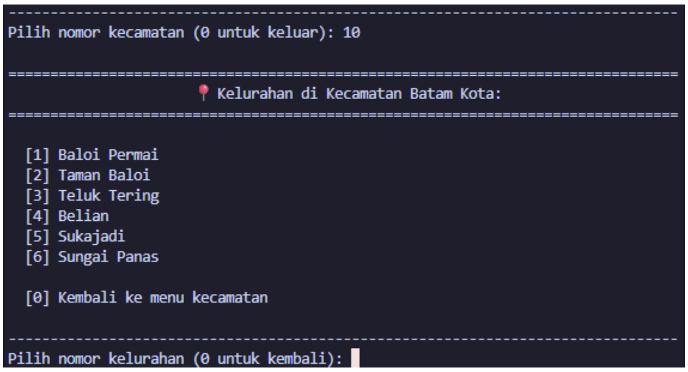

Figure 7. Detailed Location Selection

Figure 7 demonstrates the depth of the location feature. Selecting a district, for example Batam Kota, triggers the interface to show all its sub-districts, so the forecast can be tailored for smaller administrative units. Specific options such as Baloi Permai, Taman Baloi, and Teluk Tering are presented, enabling the user to select a sub-district for more accurate local weather results. The structure allows the weather prediction to be detailed and relevant to the chosen area, rather than just showing general forecasts for the entire city.

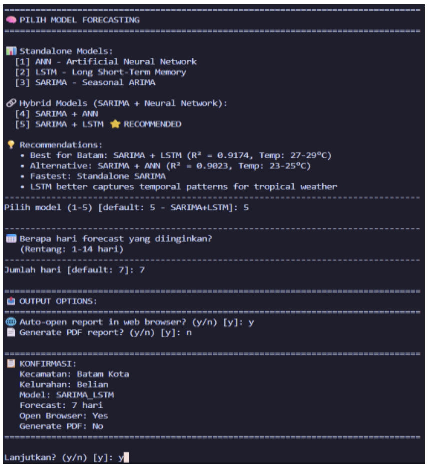

Figure 8. Model Selection and Configuration Features

Figure 8 displays the configuration interface where users can adjust forecasting parameters. The system divides five algorithms into Standalone and Hybrid Models, marking SARIMA-LSTM as recommended because it achieves the best results. Users are supported by a recommendation section that displays important metrics, confirming the hybrid model achieves highest accuracy in Kota Batam. The system lets users select a forecast period between 1 and 14 days and decide whether the output opens in a browser or is saved as a PDF. The system displays details before the prediction to verify location, model, and settings in order to execute.

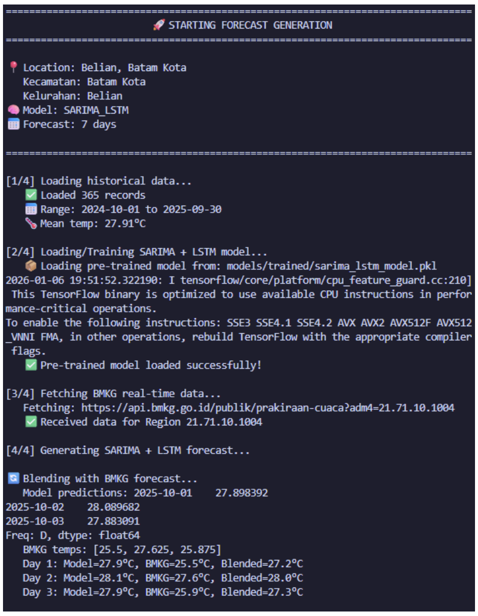

Figure 9. CLI Execution Process

Figure 9 presents the backend execution in the CLI, where the system begins by loading 365 historical records as the first step in establishing trends. The process continues by loading the SARIMA-LSTM model and initializing TensorFlow libraries needed for execution. The system connects to BMKG in real time to get live data for Belian, demonstrating the integration feature. The blending process combines machine learning results with BMKG data to calculate the temperature forecast, as seen in the log.

Figure 10. Weather Forecast Report

The final output is displayed in Figure 10 as a comprehensive HTML report made for the user. The dashboard summarizes all critical forecasting parameters, confirming that the prediction is specifically for the Belian area in Batam Kota. The report presents the model configuration, confirming that the SARIMA-LSTM hybrid was trained with 365 data points. The report explains the sources, showing that BMKG and Weather Underground data were integrated. The report details that the final forecast blends 70% model results with 30% BMKG live data for clarity.

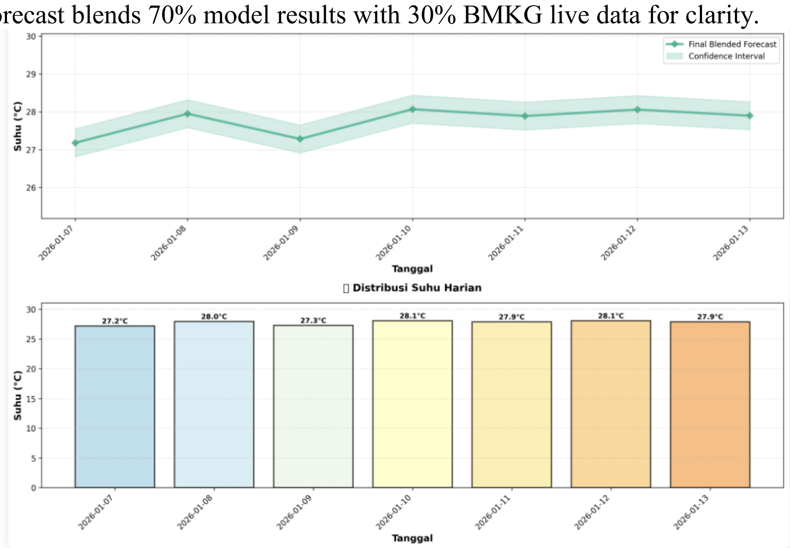

Figure 11. Forecast Graph Visualization

Figure 11 presents the temperature prediction results in a clear visual form for easy interpretation. The figure shows daily average temperature forecasts in degrees Celsius (°C) for a short-term period, from 7 January to 13 January. The upper chart displays the predicted temperature trend as a line graph, while the shaded area around the line represents the confidence interval, indicating possible uncertainty in the prediction. The lower chart shows a bar chart of daily temperatures, which helps compare temperature values from one date to another. The x-axis represents the forecast dates, and the y-axis represents the predicted temperature values produced by the model. This visualization provides clear context about when the prediction is made and how the temperature changes over time, making it easier to observe daily patterns and short-term trends in temperature.

4. CONCLUSION

The study confirms that the objectives stated in the Introduction were fully achieved and are consistent with the Results and Discussion. The integration of BMKG and Weather Underground data effectively addressed missing values and ensured continuous temporal coverage, enabling reliable model training and evaluation.

The main scientific contribution of this research is the explicit use of residual hybridization combined with multi-source data integration in a tropical climate context. By allowing SARIMA to model linear seasonal patterns and LSTM to learn non-linear residual behavior, the proposed hybrid approach significantly improved forecasting accuracy compared to standalone statistical and deep learning models.

The empirical results demonstrate that the SARIMA–LSTM model achieved the best performance, with an RMSE of 1.73 °C and an R^2 of 0.923. This confirms that combining linear and non-linear components is an effective strategy for complex time-series forecasting, particularly under high climate variability.

Although the case study focuses on Batam City, the methodological framework is generalizable to other regions with similar data limitations and climatic characteristics. The approach is especially relevant for tropical and maritime areas where incomplete records and non-linear dynamics are common.

Several limitations remain. The system depends on external APIs, operates on a relatively short observation period, and is limited to average temperature forecasting. These constraints highlight opportunities for future work, including multivariate forecasting, longer historical datasets, and reduced reliance on external data services.

5. ACKNOWLEDGEMENT

The author expresses sincere thanks to Yefta Christian, S.Kom, M.Kom., for his advice, patience, and helpful guidance, as well as appreciation to BMKG Open API for the main source of data and API access in this study, and to Weather Underground for data necessary for validation. This study acknowledges Universitas Internasional Batam for educational resources and conveys gratitude to the parents and friends for their support and encouragement. Hopefully, this study could be of use to provide a foundation and as a refernce for future research about meteorology and machine learning.

6. REFERENCES

[1] A. S. Ramadhan, A. Yudono, and A. W. Hasyim, "Dampak Pertumbuhan Kota dan Perubahan Tutupan Lahan Terhadap Temperature Humidity Index di Pulau Batam," *Planning for Urban Region and Environment Journal (PURE)*, vol. 12, no. 1, 2023. Available: https://purejournal.ub.ac.id/index.php/pure/article/view/965

[2] S. Ardabili, A. Mosavi, M. Dehghani, and A. R. Várkonyi-Kóczy, "Deep learning and machine learning in hydrological processes, climate change and Earth systems: a systematic review," in *Lecture Notes in Networks and Systems*, vol. 151, 2020, pp. 52–62. DOI: https://doi.org/10.1007/978-3-030-36841-8_5

[3] O. Fathi, "Time series forecasting using a hybrid ARIMA and LSTM model," Velvet Consulting, 202X. [Online].

[4] Q. Zhang, Z. Li, S. Snowling, A. Siam, and W. El-Dakhakhni, "Predictive models for wastewater flow forecasting based on time series analysis and Artificial Neural Network," *Water Science and Technology*, vol. 80, no. 2, pp. 243–253, Jul. 2019. DOI: https://doi.org/10.2166/wst.2019.263

[5] P. Do, C. W. Chow, R. Rameezdeen, and N. Gorjian, "Wastewater inflow time series forecasting at low temporal resolution using SARIMA model: A case study in South Australia," *Environmental Science and Pollution Research*, vol. 29, no. 47, pp. 70984–70999, May 2022. DOI: https://doi.org/10.1007/s11356-022-20777-y

[6] R. Fawzy, A. S. Eltrass, and H. M. Elkamchouchi, "A new deep learning hybrid model for accurate web traffic time series forecasting," in *2024 6th Novel Intelligent and Leading Emerging Sciences Conference (NILES)*, Oct. 2024, pp. 403–406. DOI: https://doi.org/10.1109/niles63360.2024.10753194

[7] S. Kumari and P. Muthulakshmi, "SARIMA model: An efficient machine learning technique for weather forecasting," *Procedia Computer Science*, vol. 235, pp. 656–670, 2024. DOI: https://doi.org/10.1016/j.procs.2024.04.064

[8] M. L. Hossain, S. M. Shams, and S. M. Ullah, "Time-series and Deep Learning Approaches for Renewable Energy Forecasting in Dhaka: A comparative study of ARIMA, SARIMA, and LSTM models," *Discover Sustainability*, vol. 6, no. 1, Aug. 2025. DOI: https://doi.org/10.1007/s43621-025-01733-5

[9] W. Fransiska *et al.*, "Penerapan Rantai Markov Dalam Peramalan Cuaca (Studi Kasus: Cuaca Harian di Kota Padang)," *Buana Matematika: Jurnal Ilmiah Matematika dan Pendidikan Matematika*, vol. 12, no. 2, pp. 117–126, 2022. DOI: https://doi.org/10.36456/buanamatematika.v12i2.6374

[10] Badan Pusat Statistik Kota Batam, "Hasil Sensus Penduduk Batam 2020," Berita Resmi Statistik, pp. 1–11, 2021. [Online].

[11] Y. Jiang, Z. Pan, X. Zhang, S. Garg, A. Schneider, Y. Nevmyvaka, and D. Song, "Empowering Time Series Analysis with Large Language Models: A Survey," *arXiv preprint*, 2024. DOI: https://doi.org/10.48550/arXiv.2402.03182

[12] R. R. Guerra, A. Vizziello, P. Savazzi, E. Goldoni, and P. Gamba, "Forecasting LoRaWAN RSSI using weather parameters: A comparative study of ARIMA, Artificial Intelligence and hybrid approaches," *Computer Networks*, vol. 243, p. 110258, Apr. 2024. DOI: https://doi.org/10.1016/j.comnet.2024.110258

[13] Fachrurrazi, S. Husin, Tripoli, and Mubarak, "Neural network for the standard unit price of the building area," *Procedia Engineering*, vol. 171, pp. 282–293, 2017. DOI: https://doi.org/10.1016/j.proeng.2017.01.336

[14] "Implementasi Long Short-Term Memory pada Prediksi Harga Saham PT Aneka Tambang Tbk," *Jurnal Ilmiah Komputasi*, vol. 21, no. 1, Mar. 2022. DOI: https://doi.org/10.32409/jikstik.21.1.2815

[15] Q. Liu, W. Shi, and Z. Chen, "Fatigue life prediction for vibration isolation rubber based on parameter-optimized support vector machine model," *Fatigue & Fracture of Engineering Materials & Structures*, vol. 42, no. 3, pp. 710–718, 2019. DOI: https://doi.org/10.1111/ffe.12945

[16] C. Schröer, F. Kruse, and J. M. Gómez, "A systematic literature review on applying CRISP-DM process model," *Procedia Computer Science*, vol. 181, pp. 526–534, 2021. DOI: https://doi.org/10.1016/j.procs.2021.01.199

[17] F. Martinez-Plumed *et al.*, "CRISP-DM Twenty Years Later: From Data Mining Processes to Data Science Trajectories," *IEEE Transactions on Knowledge and Data Engineering*, vol. 33, no. 8, pp. 3048–3061, Aug. 2021. DOI: https://doi.org/10.1109/tkde.2019.2962680

[18] H. V. Minh *et al.*, "Modelling and predicting annual rainfall over the Vietnamese Mekong Delta (VMD) using SARIMA," *Discover Geoscience*, vol. 2, no. 1, Jun. 2024. DOI: https://doi.org/10.1007/s44288-024-00018-0

[19] J. Xiao, X. Zhu, C. Huang, X. Yang, F. Wen, and M. Zhong, "A new approach for stock price analysis and prediction based on SSA and SVM," *International Journal of Information Technology & Decision Making*, vol. 18, no. 01, pp. 287–310, Jan. 2019. DOI: https://doi.org/10.1142/s021962201841002x

[20] P. Kabbilawsh, D. S. Kumar, and N. R. Chithra, "Forecasting long-term monthly precipitation using SARIMA models," *Journal of Earth System Science*, vol. 131, no. 3, Aug. 2022. DOI: https://doi.org/10.1007/s12040-022-01927-9

[21] L. Ilias, E. Sarmas, V. Marinakis, D. Askounis, and H. Doukas, "Unsupervised domain adaptation methods for photovoltaic power forecasting," *Applied Soft Computing*, vol. 149, p. 110979, Dec. 2023. DOI: https://doi.org/10.1016/j.asoc.2023.110979

[22] J. A. Segovia, J. F. Toaquiza, J. R. Llanos, and D. R. Rivas, "Meteorological variables forecasting system using machine learning and open-source software," *Electronics*, vol. 12, no. 4, p. 1007, Feb. 2023. DOI: https://doi.org/10.3390/electronics12041007

[23] F. Fallucchi and M. Gozzi, "Puzzle pattern, a systematic approach to multiple behavioral inheritance implementation in object-oriented programming," *Applied Sciences*, vol. 14, no. 12, p. 5083, Jun. 2024. DOI: https://doi.org/10.3390/app14125083

[24] M. Nilashi, O. Keng Boon, G. Tan, B. Lin, and R. Abumalloh, "Critical data challenges in measuring the performance of Sustainable Development Goals: Solutions and the role of big-Data Analytics," *Harvard Data Science Review*, vol. 5, no. 3, Jul. 2023. DOI: https://doi.org/10.1162/99608f92.545db2cf

[25] C. Zhang *et al.*, "Large Language Model-Brained GUI Agents: A Survey," *arXiv preprint*, May 2025. DOI: https://doi.org/10.48550/arXiv.2411.18279

[26] Y. Zhao, W. Zhang, and X. Liu, "Grid search with a weighted error function: Hyper-parameter optimization for financial time series forecasting," *Applied Soft Computing*, vol. 154, p. 111362, Mar. 2024. DOI: https://doi.org/10.1016/j.asoc.2024.111362

[27] Windary W, Hasugian AH. Data mining of rural digital technology adoption factors using Apriori algorithm. *Journal of Information Systems and Technology Research*. 2025 Sept 30;4(3):224–33. doi:10.55537/jistr.v4i3.1324