Imputation of Missing Weather Data in Automatic Weather Station Using the GRU Algorithm

Kustita Rhamadania [1]*, Makhsun[2]*, Choirul Basir [3], Naufal Ananda [4]*****
* Master of Informatics Engineering, Universitas Pamulang
** Department Mathematics, Faculty of Mathematics and Natural Sciences, Universitas Pamulang
*** Division of Instrumentation and Calibration, Region II of Meteorology Climatology and Geophysics Agency
ku5tit4@gmail.com [1], dosen00345@unpam.ac.id [2], dosen02278@unpam.ac.id [3], naufal.ananda17@gmail.com [4]

Article Info

Article history:

Received 2026-01-19
Revised 2026-02-18
Accepted 2026-02-27

Keyword:

AWS,
Imputation,
Machine Learning,
GRU,
Weather.

ABSTRACT

Missing data in Automatic Weather Stations (AWS) due to sensor failure or communication errors poses a significant challenge to the accuracy of weather and climate modeling. Traditional imputation methods often struggle with the complex temporal patterns of meteorological data. This study evaluates the performance of the Gate Recurrent Unit (GRU) algorithm for imputing missing air temperature and humidity data. The research utilized meteorological data from three AWS locations in Banten Province (AWS Staklim Banten, Golf Modern, and BSD Serpong) collected between January 2022 and January 2024. The model was tested using simulated missing data scenarios ranging from 5% to 50% and evaluated using R^2, RMSE, and MAPE metrics. The results demonstrate that the GRU model achieves high accuracy, maintaining a strong correlation ($R^2 > 0.9$) with actual sensor data even at high missing rates. The average MAPE was recorded at 1.79% for air temperature and 3.92% for humidity. Furthermore, the RMSE values met the precision criteria and measurement uncertainty limits set by the World Meteorological Organization (WMO). The study concludes that the GRU algorithm is a computationally efficient and highly accurate method for handling missing time-series weather data.

I. Introduction

Indonesia is one of the regions facing global challenges in terms of weather variability. This greatly affects various sectors of daily life. [1]. Weather parameters such as air temperature and humidity are important in weather and climate modelling to support cross-sectoral climate adaptation policies and strategies [2]. Meteorology, Climatology and Geophysics Agency (BMKG) is responsible for providing accurate, real-time weather data obtained from Automatic Weather Stations (AWS). [3]. AWS are automatic weather observation stations spread across almost the entire territory of Indonesia. Real-time weather data monitored by AWS is sent from the observation site to the BMKG database centre.[4].

Although AWS is effective in providing real-time weather data, there is often *missing data* on several weather parameters due to sensor damage, communication disruptions, or calibration and maintenance processes [5]. This incomplete data can reduce the accuracy and quality of weather and climate modelling, thereby impacting the accuracy of analyses and predictions required in various sectors. Therefore, imputation of *missing data* on weather parameters is an important step in improving the quality of data collected by AWS.

Imputation aims to replace missing values with estimates that can improve the accuracy of weather analysis and prediction [6]. However, traditional imputation methods such as mean imputation and linear interpolation have limitations in handling temporal and non-linear patterns in time series data [7]. In the context of time series data, a more adaptive and dynamic approach is needed to produce more accurate estimates. The application of machine learning models such as Gate Recurrent Unit (GRU) has been evaluated and proven effective in handling time series data imputation problems.[8].

Unlike previous studies that broadly compare various deep learning architectures, this research specifically focuses on validating the performance GRU algorithm against the rigorous World Meteorological Organization (WMO) data quality standards. GRU is selected as the primary focus due to its proven balance between accuracy and computational efficiency, specifically its lower memory usage compared to LSTM which is critical for future implementation on low-power Automatic Weather Station (AWS) loggers. This study aims to demonstrate that a standalone GRU model is sufficient to meet operational measurement uncertainty limits without the computational overhead of more complex ensembles. GRU, developed by Cho et al. (2014), offers a simpler architecture while still providing high performance [9]. Several studies have shown that GRU is effective in imputing missing data in various domains, including meteorology [10] [11]. This study aims to evaluate the GRU model in handling missing weather data on AWS.

II. METHOD

A. Data

Data used in this study is meteorological data obtained from three AWS locations in Banten Province. Figure 1 shows the distribution of AWS used in the study.

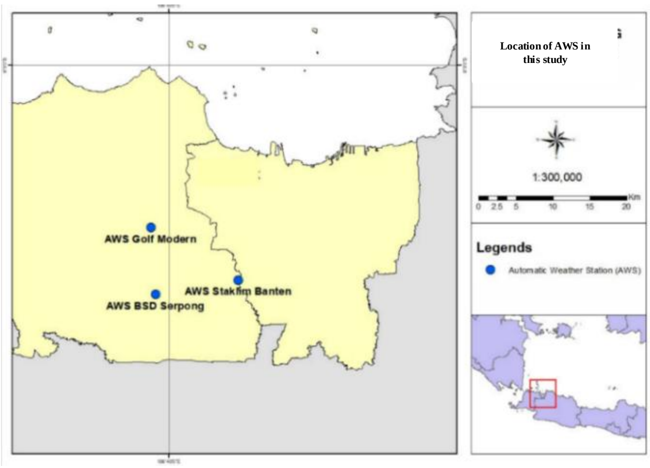

Figure 1.Location of AWS in This Study

Figure 1 shows the geographical distribution of the three AWS locations used as data sources. All three are located in a single, adjacent regional cluster. The three AWS are located in lowland areas with elevations between 18 and 24 metres above sea level. The coordinates and elevation of each AWS are shown in Table I.

In addition to elevation, the distance between AWSs is also a consideration in spatiotemporal analysis. Table 1 shows the distance matrix between AWS locations, which range from 9 to 13 kilometres. Variations in elevation and distance between AWSs provide additional context for evaluating the spatial performance of the GRU model [12].

TABEL I
COORDINATE AND ELEVATION AWS

AWS	Latitude	Longitude	Elevation
AWS Staklim Banten	-6,2614	106,7509	24
AWS Golf Modern	-6,1978	106,6448	18
AWS BSD Serpong	-6,2792	106,6503	18

AWS data is processed and tested spatially using data from temperature and humidity sensors located at AWS, namely HMP155 sensors and CR100 data loggers as the centre for data processing and analysis. The processed data can be obtained through the official BMKG portal (awscenter.bmkg.go.id) with a recording interval of every 10 minutes during the period from 1 January 2022 to 31 January 2024.

B. Pre-Processing Data

Pre-processing aims to prepare the dataset that will be used for modelling. The steps taken are to remove outliers in the raw data on air temperature, air pressure, humidity and solar radiation at the AWS. Outliers are unusual observations in a data set that deviate significantly from other observations. Outliers are identified through data quality control (QC) based on range checks and step checks [13].

Range checks are data checks based on the historical range of parameters. If the data is not within the normal range, it becomes suspect based on the range check.The normal range for air temperature data is -30 to 50 °C [14]. Step check is a check based on the temporal relationship between current data and previous data. If the difference between the current data and the previous data exceeds a certain threshold, the data becomes suspect based on the step check. The threshold value for the step check of solar radiation intensity data is that there should be no increase of 6 °C or a decrease of more than 9 °C over a period of 5 minutes .[15].

To ensure the model treats all input features equally and converges faster during training, we applied Standardization using the StandardScaler function. This technique transforms the data distribution so that the mean of observed values is 0 and the standard deviation is 1. The normalization formula is expressed as:

$$z = \frac{x - \mu}{\sigma}$$

Where z is the normalized value fed into the network, x is the original meteorological observation, μ is the mean of the training data, and σ is the standard deviation.

Considering the sequential nature of meteorological time-series data, random shuffling was avoided to prevent look-ahead bias. Instead, we employed a chronological split approach: the first 80% of the dataset (January 2022 – August 2023) was utilized for training, while the remaining 20% (September 2023 – January 2024) was reserved for testing and validation.

C. Flowchart

GRU model design is based on a variety of missing data scenarios. The *missing data* scenarios were applied to air temperature and humidity data at the Banten Climatology AWS, BSD Serpong AWS, and Golf Modern AWS alternately. The dataset was collected during 24 hours of AWS operation with a data interval of every 10 minutes. The dataset was divided into two parts, the first part being data from January 2022 to August 2023 as the imputation model and the dataset from September 2023 to January 2024 as the test data.

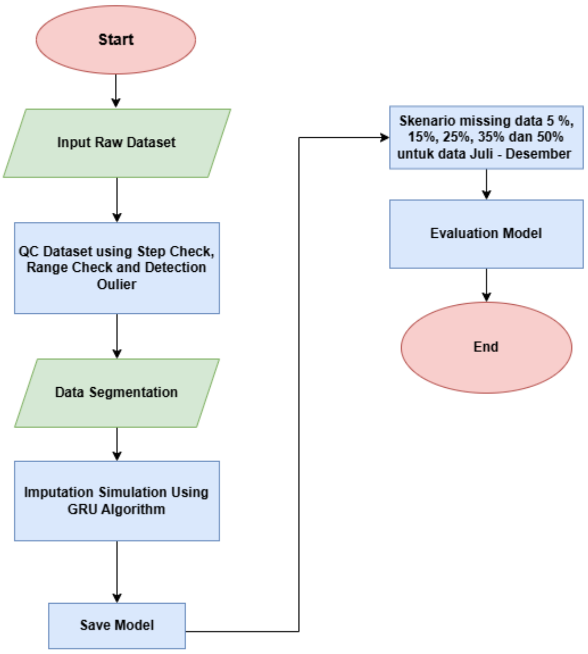

Figure 2. Research Flowchart

Figure 2 shows the experimental flow starting from the formation of the dataset, the missing data scenario usig GRU modelling, to model performance evaluation. To simulate sensor failure events, we applied a continuous block missingness strategy. Data was removed sequentially from the end of the test set, creating gaps corresponding to 5%, 15%, 25%, 35%, and 50% of the total dataset length.. This deliberately omitted data is then stored as comparative data when testing the imputation model [16]. GRU algorithms will work to impute missing air temperature and humidity data on AWS. Each estimator will be tested using evaluation metrics.

D. GRU Model Architecture

The configuration phase initiates with adjusting the input data format to ensure compatibility with the Gated Recurrent Unit (GRU) neural network architecture. Prior to being processed by the model layers, the data undergoes normalization using the Z-Score Standardization (*StandardScaler*) technique to transform the feature distribution to have a mean of 0 and a standard deviation of 1. This step aims to standardize the scale across different parameters, thereby stabilizing the learning process and accelerating gradient convergence during training. Subsequently, the data is reshaped into a 3-dimensional format (Samples, 2, 1). This dimensional structure is designed to allow the recurrent operations within the GRU to process the input sequences from neighboring stations as sequential data. In detail, the parameters employed in constructing the GRU model, ranging from pre-processing and network architecture to training configurations, are summarized in Table II below:

TABEL II
CORRELATION OF AIR HUMIDITY BETWEEN AWS LOCATIONS

Component	Parameter	Value
Pre-Processing	Normalization	StandardScaler
	Segmentation	80% Training ; 20 Testing
	Input	(Samples,2,1_
Model Architecture	Recurrent Layer	GRU (40 Units) , ReLU
	Hidden Layer	Dense (40 Units) , ReLU
	Output Layer	Dense (1 Units) , Linear
Training	Optimizer	Adam
	Loss Function	Mean Squared Error (MSE)
	Batch Size	32
	Epoch	50

E. Data Analysis

Air temperature and humidity imputation estimator model using the GRU algorithm needs to be tested to determine the accuracy performance level of each estimator. The analysis was carried out on the model for each estimator for each *missing data* scenario. The evaluation parameters used were R^2, RMSE, and MAPE. R^2 indicates the relationship between the imputed data variables and the actual data on a scale of 0-1 [17]. RMSE expresses the *error* value in units of solar radiation intensity. MAPE expresses the *error* value as a percentage. Actual data is expressed as y and imputed data is expressed as x. The equations for these three parameters are expressed as follows [18]:

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

$$\text{RMSE} = \sqrt{\frac{1}{m}\sum_{i=1}^{m}(y_i - \hat{y}_i)^2} \quad (2)$$

$$\text{MAPE} = \frac{100\%}{m}\sum_{i=1}^{m}\left|\frac{\hat{y}_i - y_i}{y_i}\right| \quad (3)$$

III. HASIL DAN PEMBAHASAN

A. *Result of Pre-Processing Data*

Pre-processing of AWS BSD Serpong, AWS Golf Modern, and AWS Staklim Banten data was reviewed by examining data availability and detecting outliers. The ideal amount of AWS data for the period 1 January 2022 - 31 January 2024 is 109,584 data points (6 data points per hour x 24 hours per day x 761 days). Table III shows the amount of data availability for the three AWS systems.

TABEL III
AVAILABILITY DATASET AWS

AWS	Air Temperature	Air Humidity
AWS Staklim Banten	106.553	106.553
AWS Golf Modern	105.907	105.907
AWS BSD Serpong	105.698	105.698

Table III shows that the availability rate of the AWS dataset 3 is quite good, with an availability rate of >90% for the period 1 January 2022 - 31 January 2024. Next, outlier detection and QC detection in the form of range checks and step checks were performed. The QC process successfully detected outliers in both parameters. The number of outliers did not significantly affect the availability of the AWS dataset. This is evidenced by the availability rate of all parameters remaining close to 90%. To complete the AWS dataset, missing data treatment was performed using the GRU algorithm.

Missing data occurs randomly at each location at different time segments. This pattern allows other AWS location variables to be used as estimators for the missing variable values. However, the determination of predictors still takes into account the correlation between variables at adjacent AWS locations [19]. Table IV shows the correlation between air temperatures across AWS locations.

TABEL IV
CORRELATION OF AIR TEMPERATURE BETWEEN AWS LOCATIONS

Site	AWS Staklim Banten	AWS Golf Modern	AWS BSD Serpong
AWS Staklim Banten	-	0.95	0.95
AWS Golf Modern	0.95	-	0.94
AWS BSD Serpong	0.95	0.94	-

Correlation between air temperatures across AWS locations shows very strong values. The correlation between AWS locations is > 0.9. This indicates that air temperature conditions in the South Tangerang area are fairly homogeneous [20]. Air temperature parameters in adjacent locations with identical characteristics have great potential for use as input for imputation models.

TABEL V
CORRELATION OF AIR HUMIDITY BETWEEN AWS LOCATIONS

Site	AWS Staklim Banten	AWS Golf Modern	AWS BSD Serpong
AWS Staklim Banten	-	0.90	0.91
AWS Golf Modern	0.90	-	0.92
AWS BSD Serpong	0.91	0.92	-

Table V shows the correlation between air humidity at different AWS locations. The correlation between air humidity at different AWS locations shows very strong values between each other. This correlation also indicates that air humidity conditions in the South Tangerang area are fairly homogeneous. [21].

B. *Air Temperature Imputation*

The third AWS air temperature parameters correlate very strongly with each other. Air temperature imputation modelling utilises the same parameters from nearby AWS locations. Figure 3 shows the GRU model architecture for AWS air temperature imputation.

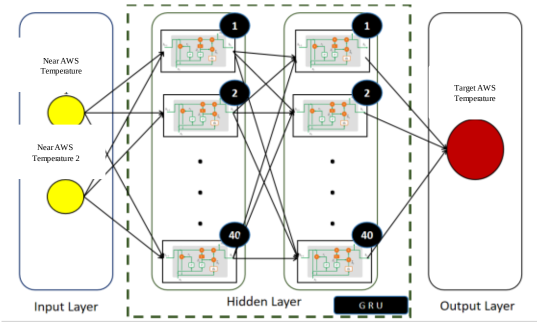

Figure 3. GRU Model Design for AWS Air Temperature Imputation

GRU hyperparameters were designed identically to the LSTM hyperparameters with two hidden layers and 40 neurons in each layer, so that they could be compared with each other. The activation function, batch size, and number of epochs were also the same as the LSTM. Table VI shows the results of AWS air temperature imputation based on GRU.

TABEL VI
AIR TEMPERATURE IMPUTATION USING GRU

Target Location	Percentage of Missing Data	R	RMSE	MAPE
AWS BSD Serpong	5%	0.97	0.60	1.62
	15%	0.98	0.70	1.89
	25%	0.98	0.70	1.95
	35%	0.98	0.71	1.97
	50%	0.97	0.73	2.03
AWS Golf Modern	5%	0.96	0.59	1.64
	15%	0.97	0.63	1.54
	25%	0.97	0.65	1.61

	35%	0.97	0.66	1.68
	50%	0.97	0.67	1.74
AWS Staklim Banten	5%	0.96	0.70	1.65
	15%	0.97	0.76	1.90
	25%	0.97	0.72	1.87
	35%	0.97	0.71	1.85
	50%	0.97	0.72	1.90

Based on Table VI, the GRU algorithm is also capable of producing an optimal AWS air temperature imputation model. The GRU imputation data correlates strongly with the actual AWS air temperature sensor data, with a correlation coefficient (R) value >0.9. In terms of RMSE values, the average standard deviation of the AWS BSD Serpong, AWS Golf Modern and AWS Staklim Banten errors were 0.051^{0}C, 0.032^{0}C and 0.021^{0}C, respectively. These three values also meet the precision criteria in the WIGOS Monitoring Data Quality Control document (2018). If the standard deviation of the *error* is converted into the standard uncertainty of *repeatability*, the measurement uncertainties of AWS BSD Serpong, AWS Golf Modern and AWS Staklim Banten are 0.010^{0}C, 0.006^{0}C and 0.004^{0}C, respectively. These values still meet the requirements of WMO No. 8 with a measurement uncertainty limit for air temperature of 0.2^{0}C [15].

C. Air Humidity Imputation

The third AWS air humidity parameters correlate very strongly with each other. Air humidity imputation modelling utilises the same parameters from nearby AWS locations. Figure 4 shows the GRU model architecture for AWS air humidity imputation.

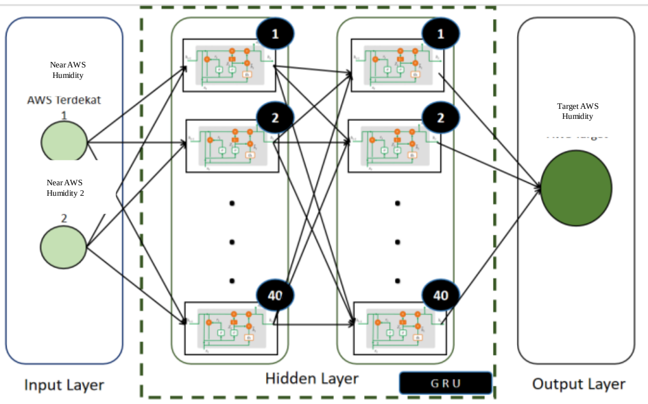

Figure 4. GRU Model Design for AWS Air Humidity Imputation

GRU hyperparameters were designed identically to the LSTM hyperparameters with two hidden layers and 40 neurons in each layer, so that they could be compared with each other. The activation function, batch size and number of epochs were also the same as for the LSTM. Table VII shows the results of AWS air humidity imputation based on GRU.

TABEL VII
AIR HUMIDITY IMPUTATION USING GRU

Target Location	Percentage of Missing Data	R	RMSE	MAPE
AWS BSD Serpong	5%	0.97	2.64	2.41
	15%	0.97	3.23	2.84
	25%	0.98	3.14	2.97
	35%	0.98	3.10	2.90
	50%	0.98	3.03	2.77
AWS Golf Modern	5%	0.94	3.68	3.60
	15%	0.95	4.17	3.99
	25%	0.95	4.32	4.23
	35%	0.95	4.22	4.11
	50%	0.95	4.19	4.02
AWS Staklim Banten	5%	0.94	4.73	4.09
	15%	0.95	5.12	5.11
	25%	0.95	5.16	5.47
	35%	0.94	5.09	5.36
	50%	0.95	4.86	4.99

Based on Table VII, overall, the GRU algorithm was able to produce a good AWS air humidity imputation model. The GRU imputation data correlated strongly with the actual AWS air humidity sensor data. This was indicated by a correlation coefficient (R) value >0.9. The average RMSE values for AWS BSD Serpong, AWS Golf Modern and AWS Staklim Banten are 3%, 4.1% and 4.9% respectively. These three RMSE values still meet the trueness criteria in the WIGOS Monitoring Data Quality Control document (2018) for surface air humidity parameters, because the error value is <10%. Furthermore, the average standard deviation error values for AWS BSD Serpong, AWS Golf Modern and AWS Staklim Banten are 0.23%, 0.25% and 0.18%, respectively. If the standard deviation of the *error* is converted into the standard repeatability uncertainty (divided by the number of RMSE data per location n=5), the measurement uncertainty values for RH AWS BSD Serpong, AWS Golf Modern and AWS Staklim Banten are 0.046%, 0.05% and 0.036%, respectively. These values still meet the requirements of WMO No. 8 regarding surface air humidity measurements with a maximum measurement uncertainty limit of 3% [15].

D. Discussion

AWS imputation model using GRU has varying performance results in each AWS location. Performance is reviewed based on the accuracy, computation duration and memory capacity of each algorithm. Accuracy performance includes correlation coefficients, RMSE and MAPE. The correlation coefficient per parameter is obtained based on the average correlation coefficient of all AWS locations. The GRU imputation model's correlation coefficient values are strong for all AWS parameters with R>0.7. The evaluation results indicate a notable difference in imputation accuracy between the two parameters, where the Mean Absolute Percentage Error (MAPE) for air humidity (3.92%) is higher compared to air temperature (1.79%). This discrepancy is attributed to the inherent physical characteristics of humidity data, which exhibits higher temporal variability and more rapid fluctuations than temperature. Air temperature typically follows a smooth diurnal cycle driven by solar radiation, making it easier for the GRU model to learn and predict its patterns. In contrast, relative humidity is highly sensitive to sudden microclimatic changes—such as brief rainfall events, wind gusts, or cloud cover—resulting in a more stochastic data profile that presents a greater challenge for the imputation model. Nevertheless, the error rates for both parameters remain well within acceptable operational limits. In terms of RMSE, the GRU model has good speed in learning and training processes, as seen from the fairly fast training process loss graph [22].

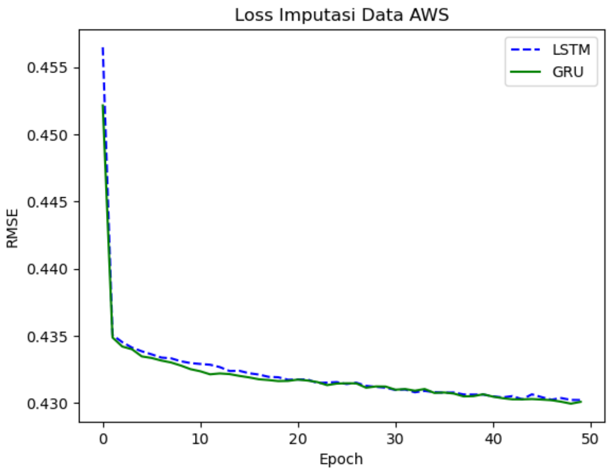

Figure 5. Graph of Loss and Epoch Model

To evaluate the training stability and compare the proposed model against a standard baseline, we analyzed the learning curve of the GRU algorithm versus the Long Short-Term Memory (LSTM) model. Figure 5 illustrates the RMSE loss reduction over 50 epochs.

As shown in the graph, the GRU model (green line) demonstrates a rapid convergence rate, effectively reducing the error within the first 5 epochs, similar to the LSTM (blue dashed line). Notably, the GRU achieves this performance with a simpler architecture. The curve flattens and remains stable towards the 50 epoch without showing signs of erratic oscillation or divergence, indicating that the model has reached an optimal state without severe overfitting. This visual evidence supports the decision to utilize GRU, as it matches the learning capability of LSTM while maintaining the computational efficiency required for AWS. The epoch computation duration for air temperature is 6 seconds and for air humidity is 11 seconds. This computation duration is due to the simple GRU neuron cell architecture, which consists of two gates, namely: update and reset, and only has one hidden state [9]. The simplification of this neuron cell architecture results in a lighter computation duration using 80.18 kb of memory [9].

IV. Conclusion

Based on the results of testing and simulation of *missing data* of 5%-50%, the GRU model was able to optimally impute AWS air temperature and humidity data. The GRU correlation coefficients achieved in AWS air temperature imputation were 0.971 and 0.957, respectively. The GRU MAPE achieved in AWS air temperature imputation was 1.79% and in air humidity imputation was 3.92%. Overall, the imputation model still meets the WMO document requirements related to air temperature and humidity measurements. The GRU's computing process and memory capacity are very efficient, as evidenced by the short GRU epoch computing duration and lower GRU memory capacity of 80.18 kb.

References

[1] T. T. K. Tran, S. M. Bateni, S. J. Ki, and H. Vosoughifar, “A Review of Neural Networks for Air Temperature Forecasting,” *Water (Basel).*, vol. 13, no. 9, p. 1294, May 2021, doi: 10.3390/w13091294.

[2] S. Naderi, S. Heslop, D. Chen, I. MacGill, and G. Pignatta, “Consumer cost savings, improved thermal comfort, and reduced peak air conditioning demand through pre-cooling in Australian housing,” *Energy Build.*, vol. 271, p. 112172, Sep. 2022, doi: 10.1016/j.enbuild.2022.112172.

[3] A. F. Radjab, H. Akib, Rifdan, and Jasruddin, “Partnership in Weather Observation using the Crowdsourcing Method,” *IOP Conf. Ser. Earth Environ. Sci.*, vol. 499, no. 1, p. 012019, Jun. 2020, doi: 10.1088/1755-1315/499/1/012019.

[4] A. Furqon, N. Ananda, and S. Soekirno, “Development of Vessel Automatic Weather Station Using Smart Weather Sensor Based on Internet of Things,” in *2025 International Conference on Smart-Green Technology in Electrical and Information Systems (ICSGTEIS)*, IEEE, Oct. 2025, pp. 123–127. doi: 10.1109/ICSGTEIS68532.2025.11284441.

[5] A. Ariffudin and P. Musa, “Analisa sistem komunikasi data berbasis Internet of Things (IoT) menggunakan metode PIECES pada Sistem Pengamatan Cuaca Otomatis di Badan Meteorologi Klimatologi dan Geofisika (BMKG),” *Jurnal Meteorologi dan Geofisika*, vol. 23, no. 2, p. 81, Jun. 2022, doi: 10.31172/jmg.v23i2.831.

[6] A. Kurniawan, “Evaluasi Pengukuran Curah Hujan Antara Hasil Pengukuran Permukaan (AWS, HELLMAN, OBS) dan Hasil Estimasi (Citra Satelit =GSMaP) Di Stasiun Klimatologi Mlati Tahun 2018,” *Jurnal Geografi, Edukasi dan Lingkungan (JGEL)*, vol. 4, no. 1, pp. 1–7, Jan. 2020, doi: 10.29405/jgel.v4i1.3797.

[7] M. Vélez-Nicolás, S. García-López, V. Ruiz-Ortiz, S. Zazo, and J. L. Molina, “Precipitation Variability and Drought Assessment Using the SPI: Application to Long-Term Series in the Strait of Gibraltar Area,” *Water (Basel).*, vol. 14, no. 6, p. 884, Mar. 2022, doi: 10.3390/w14060884.

[8] A. Gharehbaghi, R. Ghasemlounia, F. Ahmadi, and M. Albaji, “Groundwater level prediction with meteorologically sensitive Gated Recurrent Unit (GRU) neural networks,” *J. Hydrol. (Amst).*, vol. 612, p. 128262, Sep. 2022, doi: 10.1016/j.jhydrol.2022.128262.

[9] D. M. Asad, M. Shahriar, and A. H. Chowdhury, “Load Forecasting for Dhaka City Using RNN, LSTM, and GRU Architectures with Meteorological and Temporal Data,” in *2025 International Conference on Electrical, Computer and Communication Engineering (ECCE)*, IEEE, Feb. 2025, pp. 1–6. doi: 10.1109/ECCE64574.2025.11013779.

[10] J. Liu *et al.*, “Assessment of Flood Susceptibility Using Support Vector Machine in the Belt and Road Region,” May 25, 2021. doi: 10.5194/nhess-2021-80.

[11] C. Chen *et al.*, "Forecast of rainfall distribution based on fixed sliding window long short-term memory," *Engineering Applications of Computational Fluid Mechanics*, vol. 16, no. 1, pp. 248–261, 2022, doi: 10.1080/19942060.2021.2009374.

[12] H. S. Wicaksana, S. Winarko, and D. Kurniadi, "Treatment on Missing Data of Wind Speed Measurement in Automatic Weather Station Using Multivariate Imputation Chained Equation," in *2023 8th International Conference on Instrumentation, Control, and Automation (ICA)*, IEEE, Aug. 2023, pp. 81–85. doi: 10.1109/ICA58538.2023.10273106.

[13] R. Pahlepi, S. Soekirno, and H. S. Wicaksana, "Solar Radiation Intensity Imputation in Pyranometer of Automatic Weather Station Based on Long Short Term Memory," *Ultima Computing : Jurnal Sistem Komputer*, pp. 35–40, Dec. 2023, doi: 10.31937/sk.v15i2.3348.

[14] C. A. Fiebrich, C. R. Morgan, A. G. McCombs, P. K. Hall, and R. A. McPherson, "Quality Assurance Procedures for Mesoscale Meteorological Data," *J. Atmos. Ocean. Technol.*, vol. 27, no. 10, pp. 1565–1582, Oct. 2010, doi: 10.1175/2010JTECHA1433.1.

[15] WMO, "Guide to Instruments and Methods of Observation WMO-No. 8," Geneva, 2021.

[16] S. Soekirno *et al.*, "Design of drought early warning system based on standardized precipitation index prediction using hybrid ARIMA-MLP in Banten province," *IAES International Journal of Artificial Intelligence (IJ-AI)*, vol. 13, no. 2, p. 1878, Jun. 2024, doi: 10.11591/ijai.v13.i2.pp1878-1887.

[17] G. Akbar, P. Prajitno, Ariffudin, and N. Ananda, "Multivariate Imputation Chained Equation on Solar Radiation in Automatic Weather Station," *Jurnal Penelitian Pendidikan IPA*, vol. 10, no. 7, pp. 3633–3639, Jul. 2024, doi: 10.29303/jppipa.v10i7.7679.

[18] 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 Comput. Sci.*, vol. 7, pp. 1–24, 2021, doi: 10.7717/PEERJ-CS.623.

[19] W.-C. Lin and C.-F. Tsai, "Missing value imputation: a review and analysis of the literature (2006–2017)," *Artif. Intell. Rev.*, vol. 53, no. 2, pp. 1487–1509, Feb. 2020, doi: 10.1007/s10462-019-09709-4.

[20] Z. S. Venter, T. Chakraborty, and X. Lee, "Crowdsourced air temperatures contrast satellite measures of the urban heat island and its mechanisms," 2021. [Online]. Available: https://www.science.org

[21] F. R. G. Cruz, M. G. Binag, M. R. G. Ga, and F. A. A. Uy, "Multi-Layered Artificial Neural Network Flood Prediction System with Rain Gauge, Temperature Humidity Pressure Sensor, Ultrasonic Sensor, Soil Moisture Sensor and Anemometer," in *2019 IEEE 11th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment, and Management (HNICEM)*, IEEE, Nov. 2019, pp. 1–6. doi: 10.1109/HNICEM48295.2019.9072733.

[22] G. Teshome Nigatu, B. A. Abebe, B. Grum, M. G. Kebedew, and E. M. Semane, "Investigation of flood incidence causes and mitigation: Case study of Ribb river, northwestern Ethiopia," *Natural Hazards Research*, vol. 3, no. 3, pp. 408–419, Sep. 2023, doi: 10.1016/j.nhres.2023.04.009.