e-ISSN : 2600-8793

Available online at
https://jcrinn.com/
https://crinn.conferencehunter.com/

Journal of Computing Research and Innovation 11(1) 2026

Journal of Computing Research and Innovation

IoT-Enabled Temperature Pattern Classification Model for Sustainable Applications

Noor Hafizah Khairul Anuar[1*], Masmaria Abdul Majid[2], Zahari Abu Bakar[3], Norlina Mohd Zain[4], Norhalida Othman[5]

[1,2,3,4,5]*Faculty of Electrical Engineering, Universiti Teknologi MARA (UiTM) Johor Branch, Pasir Gudang Campus, 81750 Masai, Johor, Malaysia.*

ARTICLE INFO

Article history:
Received 26 May 2025
Revised 5 September 2025
Accepted 26 October 2025
Online first
Published 1 March 2026

Keywords:
IoT
Environmental Monitoring
Random Forest
Classification Model
Google Cloud Storage
Temperature Pattern

DOI:
10.24191/jcrinn.v11i1.534

ABSTRACT

This study presents the design and implementation of an IoT-based environmental monitoring system that integrates real-time data collection with classification analytics to support sustainable decision-making. In response to the increasing need for accessible and affordable monitoring tools, the system utilizes a NodeMCU ESP8266 microcontroller, paired with DHT22 and raindrop sensors, to capture temperature, humidity, and rainfall status at 30-second intervals. Data is transmitted wirelessly and stored on Google Sheets, enabling cloud-based visualization and analysis. A Random Forest classifier was applied to categorize temperature conditions into low, medium, and high ranges based on humidity and rain status to derive actionable insights from the collected data. Model performance produces overall accuracy of 65.9% revealed a strong ability to detect high-temperature conditions, with rain status identified as the most influential predictor. However, challenges such as class imbalance and limited prediction of low-temperature conditions were observed. Recommendations include enhancing the model with balanced datasets, time-based feature engineering, and considering regression models for more granular forecasting. This system demonstrates a scalable and adaptable approach to environmental monitoring, suitable for educational, research, and field applications in data-driven sustainability efforts.

1. INTRODUCTION

Global temperatures are likely to exceed the 1.5°C warming threshold in 2024 with a 76% chance of setting a record (Dunstone et al., 2022). As a result, 68% of environment-related Sustainable Development Goals (SDGs) lack sufficient data, making it difficult to make informed decisions for sustainability (Campbell et al., 2020). IoT-based real-time environmental monitoring systems with predictive and classification analytics are urgently needed to improve data accuracy and support sustainable decision-making. However, existing environmental monitoring solutions face challenges such as high costs, limited accessibility, and

[1*] Corresponding author. *E-mail address*: noorhafizah2575@uitm.edu.my
https://doi.org/10.24191/jcrinn.v11i1.534

scalability issues. Many traditional systems are bulky, require significant power, and lack predictive capabilities, making them inefficient for real-time applications in remote or resource-limited areas. Additionally, the absence of an integrated and affordable system that can continuously monitor multiple environmental parameters has created a gap in sustainability-driven decision-making. This project aims to address these limitations by developing an IoT-based environmental monitoring system that is compact, affordable, and capable of monitoring real data. As an initial system, this project uses a temperature and raindrop sensor device to establish the whole system architecture. Due to its scalability and friendliness, this system is highly adaptable and can be integrated with additional parameters or sensors. As a result of its flexibility, the system can monitor a broader range of environmental parameters and benefit various applications.

1.1 Literature review

Table 1 presents recent advancements in IoT-based environmental monitoring that integrate artificial intelligence, predictive analytics, and real-time data collection. The studies emphasize the growing capability of machine learning and deep learning techniques to enhance data-driven decision-making in environmental systems. For instance, several works demonstrate how AI integration has significantly improved the accuracy and adaptability of environmental predictions (Arabelli et al., 2024; Khan et al., 2024). In the urban area, many applications are focused on sustainability in transportation and city planning (Rozhdestvenskiy & Poornima, 2024). However, most studies lack real-world deployment validation and overlook crucial aspects such as system scalability, energy efficiency, and data security (Barnett & Riley, 1995; Kychkin et al., 2022). Based on literature, holistic, secure, and scalable solutions are needed to support environmental monitoring infrastructures' long-term reliability and sustainability.

Table 1. Summary of Research conducted in environmental monitoring utilizing predictive modeling approaches

Reference	Year	Method Applied	Key Contributions	Identified Research Gaps
Barnett and Riley (1995)	2024	IoT sensors, AI-based analytics, real-time monitoring	Discusses IoT's role in real-time environmental monitoring, predictive analytics, and challenges.	Limited focus on scalability and energy efficiency in IoT deployment.
Arabelli et al. (2024)	2024	Machine learning, deep learning, IoT sensors	Explores AI integration in IoT-based environmental monitoring for real-time predictive analytics.	Lack of real-world validation for AI-driven predictive analytics in environmental monitoring.
Rozhdestvenskiy and Poornima (2024)	2024	Predictive modeling, IoT-based data collection, urban analytics	Uses IoT and predictive analytics to enhance urban sustainability and transportation.	More research is needed on integrating IoT data with urban planning strategies.
Khan et al. (2024)	2024	GRU-Auto encoder, AI, IoT-connected optical systems	Combines AI and IoT for real-time adaptive environmental data analysis.	Challenges in real-time adaptability and accuracy of predictive analytics models.
Kychkin et al. (2022)	2022	IoT platform integration, predictive analytics, ML models	Develops an IoT-based predictive analytics system for continuous emissions monitoring.	Security and reliability concerns in IoT-based predictive emissions monitoring.
Blagojevic and Ristic (2015)	2015	Decision-making framework,	Analyzes decision-making in IoT-based environmental monitoring, emphasizing	Limited frameworks for multi-parameter decision-making under uncertainty.

https://doi.org/10.24191/jcrinn.v11i1.534

		environmental monitoring analytics	predictive-preventive approaches.	
Nandhini et al. (2023)	2023	Sensor fusion, IoT network integration	Presents IoT-based sensor fusion techniques for improved environmental monitoring accuracy.	Scalability and robustness issues in sensor fusion techniques for IoT monitoring.
Doraswamy et al. (2024)	2024	Deep learning framework, IoT dashboard, multimodal data analysis	Develops a deep learning framework for analyzing IoT-generated environmental data.	Optimization of multi-modality data processing techniques for sustainability monitoring.
Panduman et al. (2024)	2024	AI techniques, predictive analytics, IoT data integration	Surveys AI techniques used in IoT applications for real-time environmental monitoring and analytics.	Lack of standardized AI integration frameworks for IoT in environmental analytics.
Arya et al. (2023)	2023	IoT environmental monitoring, smart sustainability frameworks	Evaluates IoT-enabled monitoring systems for sustainability, focusing on air and water quality.	Challenges in IoT system interoperability and security for sustainability applications.

2. SYSTEM DESIGN AND IMPLEMENTATIONS

Fig. 1 shows the general block diagram of the monitoring system utilizing the input, output, and the microcontroller. The NodeMCU ESP8266 was selected for its integrated Wi-Fi module, cost-effectiveness, user-friendly, reliability, and operates on low power, making it highly suitable for IoT-based applications. The input section is flexible and adaptable to various sensor types, depending on the application. In this project, the DHT22 sensor (for measuring temperature and humidity) and a rain sensor are used for measuring the status of rain. For the output section, the project employs Google Sheets as the cloud-based data logging platform. It is chosen due to its structured format, easy accessibility, real-time update capabilities, and the ability to support data visualization, sharing, and storage without requiring a complex setup. This makes it ideal for beginners and educational purposes while still being functional for practical applications.

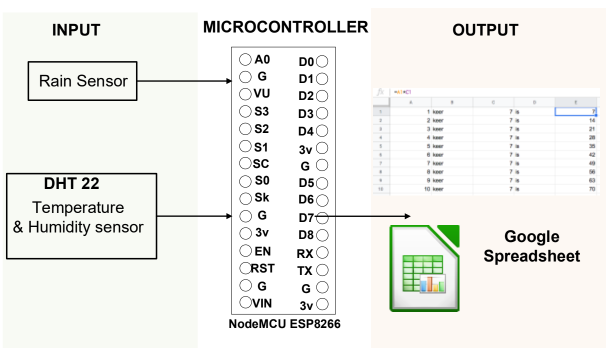


Fig. 1. The general block diagram consists of the input and output architecture of the monitoring system

Fig. 2 shows the process flow for real-time data acquisition and cloud integration of the system. The process begins with initializing a Google Spreadsheet, where relevant columns such as ‘DATE’, ‘TIME’, and sensor readings are defined. Next, a Google Apps Script is created to handle ‘HTTP GET’ requests,

https://doi.org/10.24191/jcrinn.v11i1.534

enabling the extraction of variables and timestamped data for updating the spreadsheet. On the hardware side, the NodeMCU ESP8266 is programmed using the Arduino IDE. Necessary libraries are imported, and the Wi-Fi connection is established to enable internet-based data transmission. Sensor readings are captured periodically through the millis() function, which provides non-blocking timing. This function enables customizable data sampling intervals. The microcontroller continuously checks if the specified interval has elapsed (currentMillis - previousMillis >= interval). Once true, the sensor data is processed and pushed to the Google Spreadsheet via a formatted GET request. This loop ensures asynchronous and consistent environmental data logging, suitable for lightweight IoT applications.

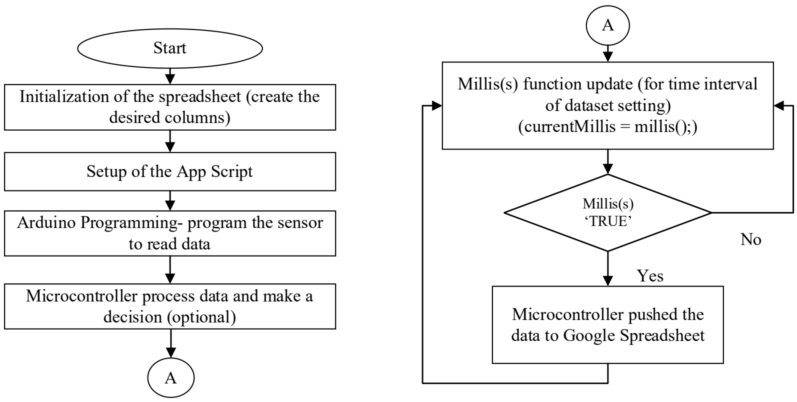


Fig. 2. The flow of overall system operations

3. DATASET COLLECTION AND INTERPRETATIONS

Fig. 3 shows the data set spanning from 9 January 2025 to 23 January 2025 using the designed IoT-based monitoring system installed on the balcony area of a residential house in Bandar Seri Alam, Johor Bharu. The data entries were recorded at 30-second intervals. Temperature readings are recorded in degrees Celsius (°C), while humidity levels are expressed as relative humidity in percent (%RH). Meanwhile, the rain status is encoded as a binary value, where '1' indicates no rain and '0' indicates rain at the time of data capture.

	A	B	C	D	E
1	DATE	TIME	TEMPERATURE	HUMIDITY	RAIN
2	1/9/2025	21:51:18	27.6	1	1
3	1/9/2025	22:21:19	27.2	13.9	1
4	1/9/2025	22:51:21	27	22	1
5	1/9/2025	23:21:23	26.9	38.9	1
6	1/9/2025	23:51:26	27.2	39.8	1
7	1/10/2025	0:21:26	28.2	10.3	1
8	1/10/2025	0:51:27	27.8	13.5	1
9	1/10/2025	1:21:29	27.7	11.7	1
10	1/10/2025	1:51:30	27.6	7.6	1
11	1/10/2025	2:21:32	27.5	15	1
12	1/10/2025	2:51:33	27.5	22.2	1
13	1/10/2025	16:02:44	25.6	99.9	0
14	1/10/2025	16:32:46	26.2	99.9	0
15	1/10/2025	17:02:48	26.5	99.9	0
16	1/10/2025	17:32:50	26.5	99.9	0
17	1/10/2025	18:02:51	26.6	99.9	0

Fig. 3. Example of Google spreadsheet interface of the dataset utilized in this study

Fig. 4 shows the temperature (°C) and humidity (%RH) fluctuation from January 9 to January 23, 2025, based on interpolated IoT sensor data. The left y-axis corresponds to temperature values (in red), while the right y-axis represents humidity levels (in blue). According to the data, temperature and humidity have a

https://doi.org/10.24191/jcrinn.v11i1.534

clear inverse relationship where the humidity tends to decrease as the temperature increases. These patterns align with known meteorological behavior, where warmer air holds more moisture but often results in lower relative humidity percentages.

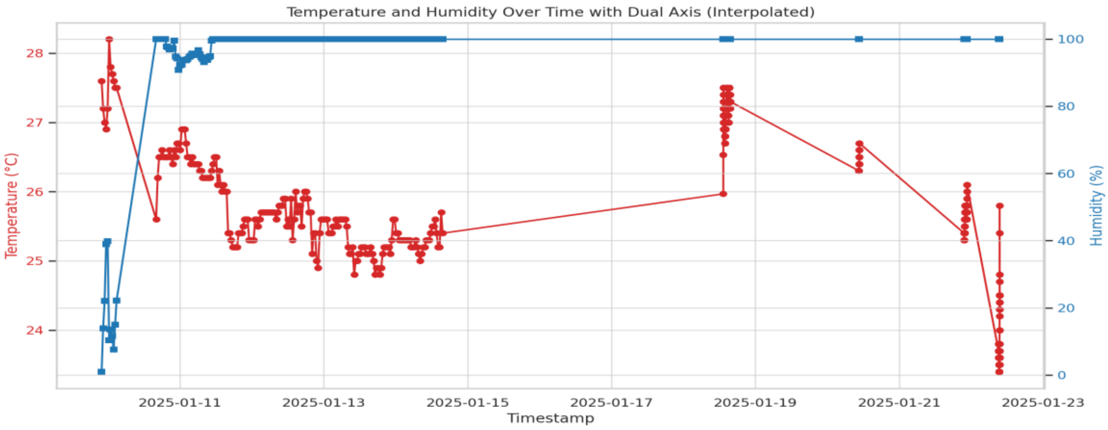


Fig. 4. The interpolated relationship between temperature and humidity

Fig. 5 provides a visual summary of the rainfall status distribution over the monitoring period. The dataset consists of binary rain status values, where '0' represents rain and '1' represents no rain. According to the chart, most observations (~490 entries) occurred during non-rain conditions, while approximately 220 entries were recorded during rain events. This imbalanced distribution highlights that dry conditions were more frequent during the observation window. The skewed nature of the data is an important consideration for predictive modeling or classification tasks, especially when applying supervised learning methods.

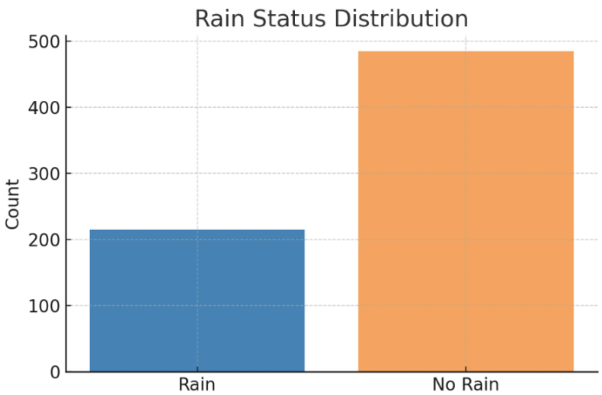


Fig. 5. The bar chart of total distribution data for rain and no rain conditions

4. CLASSIFICATION MODEL USING THE RANDOM FOREST ALGORITHM

In this study, the Random Forest algorithm was utilized as the primary classification model due to its robustness and versatility when it comes to handling structured environmental data. By combining multiple

https://doi.org/10.24191/jcrinn.v11i1.534

decision trees and aggregating their outputs, Random Forest improves predictive accuracy while minimizing overfitting risk. As in this study, it is particularly effective when there is no linear relationship between features in a small or medium-sized dataset. Additionally, Random Forest provides an integrated built-in feature importance analysis to identify significant predictors of the model output. In this experiment, Random Forest classifiers were used to classify temperature patterns according to environmental conditions consisted of temperature, relative humidity, and rain status, as temperature was set as an output of the classification models. The task categorizes temperature readings into three discrete classes as detailed in Table 2.

Table 2: Detailed data labelling for temperature features

Class Label	Temperature Category	Details
Class 0	Low	Temperature ≤ 25 ºC
Class 1	Medium	25 ºC ≤ Temperature ≤ 27 ºC
Class 2	High	Temperature ≥ 27 ºC

The next step is data preprocessing, where all missing values are interpolated to maintain the continuity of the time series dataset. The categorical rain status (0 for rain, 1 for no rain) was included as a binary feature. The indexing of time and date was removed as it is considered non-informative for classification without temporal encoding. Afterwards, the selected features were set for inputs and outputs before the data was divided into training and test sets of 80:20 percent, respectively. Model training was conducted using the scikit-learn library to train a Random Forest classifier. Then, to avoid overfitting, the hyperparameters were tuned, such as the number of trees ('n_estimators') and maximum depth ('max_depth'). Lastly, precision, recall, F1-score, and total accuracy were used to evaluate the model's performance, as shown in Table 3.

Table 3: The model performance using the Random Forest classifier

Class Label	Precision	Recall	F1-Score	Support
Class 0	0.00	0.00	0.00	29
Class 1	0.84	0.62	0.71	98
Class 2	0.57	0.93	0.70	84

As a result, the model's overall accuracy of 65.9% suggests a decent capacity for classification. The precision, recall, and F1-score in Class '0' return to 0.00 values due to an imbalance and underrepresentation for low temperature data; as a result, the model was unable to identify patterns that set this group apart. Next, the precision, recall, and F1-score in Class '1' yield values of 0.84, 0.62, and 0.71, respectively demonstrated a strong balance and precision, although the model's intermediate recall suggests that it makes accurate predictions but occasionally overlooks real-world occurrences. Finally, Class '2' gives a high recall value at 0.93. Although the precision is lower (0.57), the high recall suggests that the model is particularly sensitive to identifying hot temperature periods, occasionally mislabeling other classes as high.

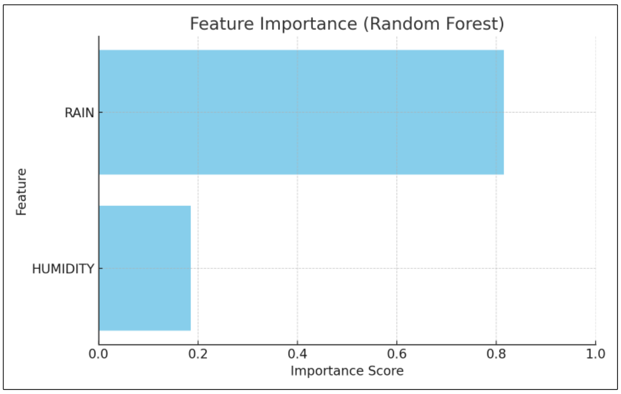
5. THE FEATURE IMPORTANCE AND CONFUSION MATRIX ANALYSIS

Fig. 5(a) shows the feature importance analysis of the Random Forest classifier, revealing that rain status is the most influential predictor in determining the temperature class, contributing approximately 81.5% to the model's decision-making process. In contrast, humidity contributed only 18.5%, indicating a significantly less impact on the classification outcome. This suggests that the model relies heavily on the presence or absence of rain to distinguish between temperature classes. However, this reliance introduces

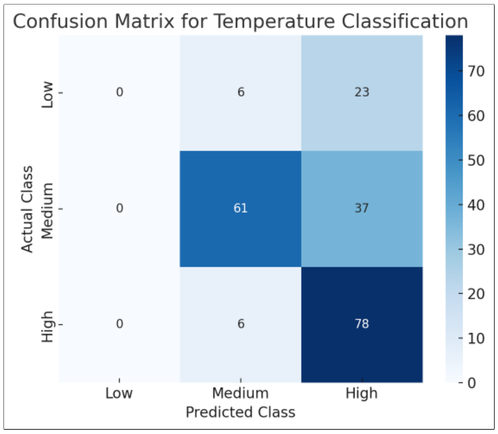
https://doi.org/10.24191/jcrinn.v11i1.534

a limitation in the model's generalizability, especially when temperature variation is not strongly associated with rain patterns. Moreover, the model struggles to predict low-temperature events (Class 0), likely due to class imbalance, as the dataset contains relatively fewer samples of low-temperature conditions. As a result, the model fails to capture the unique patterns associated with this class, leading to poor classification performance. To improve this, it is recommended to augment the dataset with more low-temperature samples or apply balancing techniques such as oversampling, undersampling, or synthetic data generation.

Fig. 5(b) shows the confusion matrix distribution of actual versus predicted temperature classes across three categories: Low, Medium, and High. The model struggled significantly with predicting the 'Low' temperature class, with 0 correct predictions. All instances from the Low class were misclassified as either Medium or High. For the Medium temperature class, the model correctly identified 61 instances but misclassified 37 samples as High. The High-temperature class shows the most promising performance, with 78 out of 84 instances correctly predicted, indicating a high recall rate for this category.



(a)



(b)

Fig. 5: (a) The feature importance plot between humidity and rain status in the model classification (b)The Confusion matrix distribution between actual versus predicted classes across three temperature categories.

6. CONCLUSION AND RECOMMENDATIONS

In conclusion, the classification model showed reasonable performance in identifying high-temperature conditions but performed poorly in predicting low-temperature instances, primarily due to class imbalance. The analysis also revealed that rain status was the most influential feature, indicating the model's heavy reliance on rainfall patterns in making predictions.

To enhance the model's effectiveness, several recommendations are proposed. First, the dataset should be balanced to ensure adequate representation of all temperature classes, which can be achieved through oversampling techniques like SMOTE or by undersampling the dominant classes. Second, additional features should be engineered, such as incorporating time-based variables or lagged values to better capture temporal dependencies. Third, alternative classification algorithms such as Gradient Boosting or Support Vector Machines (SVM) should be explored to improve predictive performance and model robustness. Lastly, instead of classifying temperature into discrete categories, the problem could be reframed as a regression task to predict continuous temperature values, allowing for more precise and detailed environmental monitoring.

https://doi.org/10.24191/jcrinn.v11i1.534

7. ACKNOWLEDGEMENTS/FUNDING

The authors would like to acknowledge the support of Universiti Teknologi MARA (UiTM), Johor Branch, Pasir Gudang Campus, and Faculty of Electrical Engineering for supporting this study.

8. CONFLICT OF INTEREST STATEMENT

The authors agree that this research was conducted in the absence of any self-benefits, commercial or financial conflicts and declare the absence of conflicting interests with the funders.

9. AUTHORS' CONTRIBUTIONS

Noor Hafizah Khairul Anuar: Project coordination, manuscript writing, and overall supervision; **Norhalida Othman**: Data analysis, validation of system functionality, and manuscript editing; **Masmaria Abdul Majid**: Literature review, system testing, and result interpretation; **Norlina Mohd Zain**: Technical documentation, graphical representation, and formatting; **Zahari Abu Bakar**: System design, hardware integration, and prototype development.

10. REFERENCES

Arabelli, R., Boddepalli, E., Buradkar, M., Goriparti, S., & Chakravarthi, M. K. (2024). IoT-enabled environmental monitoring system using AI. In *2024 International Conference on Advances in Computing, Communication and Applied Informatics (ACCAI)* (pp. 1–6). IEEE. https://doi.org/10.1109/ACCAI61061.2024.10602131

Arya, L., Sharma, Y. K., & Kumar, R. (2023). Towards a greener tomorrow: IoT-enabled smart environment monitoring systems. In *2023 International Conference on Advances in Computation, Communication and Information Technology (ICAICCIT)* (pp. 1112–1117). IEEE. https://doi.org/10.1109/ICAICCIT60255.2023.10465814

Barnett, V., & Riley, J. (1995). Statistics for environmental change. *Experimental Agriculture*, *31*(2), 117–130. https://doi.org/10.1017/S0014479700025217

Blagojević, M., & Ristić, D. (2015). Some aspects of decision making in a system of environmental monitoring. *Facta Universitatis, Series: Working and Living Environmental Protection*, *12*(1), 31–41.

Campbell, J., Neuner, J., See, L., Fritz, S., Fraisl, D., Espey, J., & Kim, A. (2020). The role of combining national official statistics with global monitoring to close the data gaps in the environmental SDGs. *Statistical Journal of the IAOS*, *36*(2), 443–453. https://doi.org/10.3233/sji-200648

Doraswamy, B., Krishna, K. L., & Tarigonda, H. (2024). IoT-generated multi-modality data analysis using a deep learning framework for managing sustainability in smart environments. *International Research Journal of Multidisciplinary Scope*, *5*(1), 12–20. https://doi.org/10.47857/irjms.2024.v5i1.0123

Dunstone, N., Smith, D., Atkinson, C., Colman, A., Folland, C., Hermanson, L., Ineson, S., Killick, R., Morice, C., Rayner, N., Seabrook, M., & Scaife, A. (2024). Will 2024 be the first year that global temperature exceeds 1.5°C? *Atmospheric Science Letters*, *25*(4), Article e1254. https://doi.org/10.1002/asl.1254

Khan, S. A., Kalifullah, A. H., Ibragimova, K., Singh, A. K., Muniyandy, E., & Rachapudi, V. (2024).

Integrating AI and IoT in advanced optical systems for sustainable energy and environment monitoring. *International Journal of Advanced Computer Science and Applications*, *15*(5), 1013–1022. https://doi.org/10.14569/IJACSA.2024.01505123

Kychkin, A., Gorshkov, O. V., & Kukarkin, M. (2022). Predictive models integration with an environmental monitoring IoT platform. *Journal of Applied Informatics*, *17*(4), 61–73. https://doi.org/10.37791/2687-0649-2022-17-4-61-73

Nandhini, B., Karthika, S., Karthiyayini, S., & Krishnaveni, K. (2023). Sensor fusion techniques for enhanced environmental monitoring in IoT. In *2023 International Conference on Sustainable Communication Networks and Application (ICSCNA)* (pp. 445–450). IEEE. https://doi.org/10.1109/ICSCNA58489.2023.10370221

Panduman, Y., Funabiki, N., Fajrianti, E. D., Fang, S., & Sukaridhoto, S. (2024). A survey of AI techniques in IoT applications with use case investigations in the smart environmental monitoring and analytics in real-time IoT platform. *Information*, *15*(3), 153. https://doi.org/10.3390/info15030153

Rozhdestvenskiy, O. I., & Poornima, E. (2024). Enabling sustainable urban transportation with predictive analytics and IoT. *MATEC Web of Conferences*, *392*, 01179. https://doi.org/10.1051/matecconf/202439201179