



# Application of machine learning for short-term climate prediction in Indonesia

Gunawan Gunawan<sup>1</sup>, Wresti Andriani<sup>2</sup>, Aminnur Aimar Akbar<sup>3</sup>  
<sup>1,2,3</sup>Informatics Engineering, STMIK YMI TEGAL, Indonesia

## ARTICLE INFO

### Article history:

Received May 25 2024  
Revised May 27, 2024  
Accepted May 30, 2024

### Keywords:

Climate Prediction;  
Machine Learning;  
Random Forest;  
Short-Term Forecasting;  
Support Vector Machine.

## ABSTRACT

This study explores the Application of Machine Learning for Short-Term Climate Prediction in Indonesia, focusing on enhancing forecast accuracy through advanced computational models. The primary objective was to develop and validate Random Forest and Support Vector Machine (SVM) models to predict short-term climate conditions accurately across ten major Indonesian cities. Employing a quantitative approach, the study utilized experimental design, rigorous data analysis, and model validation using historical weather data from April 2024 provided by the Indonesian Meteorological, Climatological, and Geophysical Agency (BMKG). The results indicate that both Random Forest and SVM significantly outperform traditional climate prediction models, with Random Forest achieving an average accuracy of 87.5% and SVM 85.2%. These findings underscore the potential of machine learning to revolutionize short-term climate predictions in regions with complex meteorological dynamics like Indonesia, offering substantial implications for disaster preparedness, agricultural planning, and urban management. Future research can expand upon these models by incorporating real-time data and exploring deep learning techniques to enhance predictive reliability further.

This is an open access article under the [CC BY-NC](https://creativecommons.org/licenses/by-nc/4.0/) license.



## Corresponding Author:

Wresti Andriani,  
Informatics Engineering,  
STMIK YMI TEGAL,  
#1 Pendidikan Street, Tegal City, Central Java, 52147, Indonesia.  
[Wresty.andriani@stmik-tegal.ac.id](mailto:Wresty.andriani@stmik-tegal.ac.id)

## 1. INTRODUCTION

Indonesia, an expansive archipelago, faces a variety of climatic conditions that significantly affect its agriculture, fisheries, and urban development. Short-term climate prediction is paramount for timely decision-making that can mitigate potential adversities caused by sudden climatic changes (Streets & Glantz, 2000). However, traditional meteorological methods often fail in accuracy and efficiency, especially in localized predictions for a geographically diverse region like Indonesia (Latif et al., 2023).

Current meteorological models predominantly focus on long-term climate predictions and often utilize data with coarse resolution (Kendon et al., 2021). These models lack the precision required for short-term, localized climate forecasting, which is essential for immediate weather-related responses in urban and rural settings (Chen et al., 2023). Traditional meteorological methods are often ineffective for short-term

predictions in Indonesia due to their reliance on long-term models and coarse-resolution data. These models typically focus on broad geographical areas without capturing the fine granularity needed for accurate, localized forecasts. They also fail to provide the necessary temporal resolution to predict sudden weather changes accurately. Consequently, these methods are inadequate for immediate weather-related responses, impacting sectors such as agriculture, disaster management, and urban planning (Chen et al., 2023).

Most existing models either cover broad geographical areas without focusing on specific regional needs or utilize datasets that do not capture the fine granularity required for accurate short-term predictions (Das et al., 2022; Klemm & Vennemann, 2021). There is a distinct lack of research into ML models that synthesize diverse weather parameters to enhance prediction accuracy for all major cities across Indonesia (Nurwanda & Honjo, 2020; Wu & Levinson, 2021). The proposed approach differs from previous research by integrating multiple climatic variables using high-resolution data from BMKG for short-term weather predictions. Previous models focused on single-variable predictions or used low-resolution data, lacking precision. This study combines Random Forest and SVM with advanced preprocessing to predict variables like temperature, humidity, and wind speed, aiming to improve accuracy and real-time applicability, thereby closing gaps in current literature (Nurwanda & Honjo, 2020).

Random Forest and Support Vector Machine (SVM) handle the complexity of weather data by utilizing advanced computational techniques that can process and analyze high-resolution, multidimensional datasets effectively. Random Forest is an ensemble learning method that constructs multiple decision trees during training and averages their results to manage overfitting and enhance accuracy. This approach allows the model to capture complex interactions and nonlinear relationships between various weather variables, such as temperature, humidity, and wind speed, which are crucial for accurate weather predictions. In the study, Random Forest achieved an average accuracy of 92%, significantly outperforming traditional models and highlighting its robustness in dealing with intricate weather patterns (Ali et al., 2021).

SVM, on the other hand, uses kernel functions to transform the data into higher dimensions where it becomes linearly separable, effectively handling nonlinear relationships between atmospheric parameters. SVM achieved an average accuracy of 88%, showing its proficiency in classifying and predicting complex climatic variables by focusing on key features like temperature and humidity (Wang et al., 2022). Both models significantly outperform traditional weather prediction methods, which typically achieve around 70-75% accuracy, by leveraging high-resolution data and sophisticated machine learning algorithms. This leads to improved short-term weather predictions, essential for disaster preparedness and urban planning in regions with diverse meteorological conditions like Indonesia (BMKG, 2024).

This study discusses the potential application and impact of Random Forest and SVM models by BMKG or related institutions. This study highlights that the application of machine learning models, especially Random Forest and SVM, can significantly improve weather forecasting capabilities in Indonesia. These improvements are critical for better preparedness and response strategies to weather-related events, thereby benefiting sectors such as agriculture, disaster management and urban planning. By leveraging high-resolution data and overcoming the limitations of existing models, this machine learning approach offers great potential for real-time climate monitoring, disaster preparedness, and resource allocation, thereby improving overall climate resilience in Indonesia's densely populated cities (BMKG, 2024).

We propose a novel approach that combines Random Forest and SVM with enhanced data preprocessing techniques to predict multiple weather variables, including temperature, humidity, rainfall, wind speed, and wind direction. The models will be trained on a comprehensive dataset from the Indonesian Meteorological, Climatological, and Geophysical Agency (BMKG) from April 1, 2024, to April 30, 2024. This approach is

designed to capture the intricate patterns of Indonesia's climate, providing a localized prediction model that can be deployed in real time.

This research aims to significantly aid in disaster preparedness, resource allocation, and overall climate resilience for Indonesia's populous cities by improving the accuracy of short-term weather forecasts. Effective forecasting models will directly benefit emergency response teams, urban planners, and the general public, reducing the adverse impacts of unforeseen weather conditions (Merz et al., 2020).

The article is organized as follows: following the introduction, we describe the methodology, detailing the data collection, model training, and validation processes. The results section will present the comparative effectiveness of the Random Forest and SVM models. This will be followed by a discussion interpreting the results in the context of current climate prediction challenges in Indonesia. The article concludes with a summary of findings, implications for future research, and potential study applications.

The burgeoning field of climate prediction has seen a significant pivot towards utilizing machine learning (ML) techniques due to their ability to manage large datasets and uncover complex patterns that are often nonlinear and multidimensional. Recent advancements underscore shifting from traditional statistical models to more dynamic ML models.

Random Forest, a robust ensemble learning method, has been particularly effective in predicting climate variables due to its capacity to handle overfitting and improve accuracy by averaging multiple decision trees (Y. Zhang et al., 2022), demonstrated the use of Random Forest in predicting temperature anomalies with a significant reduction in prediction error compared to linear models (Jose et al., 2022). Their research, however, was limited to temperate climate zones and did not address the predictive needs of tropical regions like Indonesia, which experience more volatile weather patterns (Bienvenido-Huertas et al., 2023).

Support Vector Machines (SVM) have also been highlighted in recent studies for their effectiveness in climate-related applications (Zennaro et al., 2021). The Application of SVM to predict monsoon rainfall in India, achieving higher precision by effectively handling nonlinear relationships between atmospheric parameters (X. Zhang et al., 2020). While these results are promising, the study was constrained by its focus on single-variable predictions—primarily rainfall—without integrating other crucial climatic factors, such as wind speed and humidity, vital for comprehensive short-term climate predictions (Jaseena & Koor, 2022).

Integrating multiple climatic variables in prediction models remains underexplored, particularly in studies focusing on Indonesia's diverse meteorological conditions (Ngila et al., 2023). They addressed this gap partially by developing a neural network model that synthesized temperature, humidity, and air pressure data to predict urban heat islands in Beijing (Gonzalez-Trevizo et al., 2021). Their approach underscores the potential of integrated models but is geographically and contextually limited, pointing to a significant research opportunity in Indonesia (Polkinghorne et al., 2024).

The need for localized weather prediction models is particularly acute in archipelagic countries like Indonesia (Susanto et al., 2020). Studies have explored local models but are often limited in geographic scope or specific to particular weather events, such as typhoons or droughts (Zscheischler et al., 2020). These studies demonstrate the efficacy of ML models but often do not translate across the diverse geographic and climatic spectrum found in Indonesia, from its humid coastal cities to its drier mountainous regions.

## 2. RESEARCH METHOD

### 2.1. Research Objective

The primary objective of this study is to develop and validate machine learning models, precisely Random Forest and SVM, to predict short-term climate conditions for

ten major cities in Indonesia. These models aim to integrate multiple climatic variables to enhance the accuracy and reliability of 24-hour weather forecasts.

## 2.2. Research Design

This study employs a quantitative approach combining experimental methods, quantitative analysis, and model validation (Ali et al., 2021). The design integrates theoretical modeling with empirical data analysis to evaluate and refine the predictive capabilities of the proposed ML models.

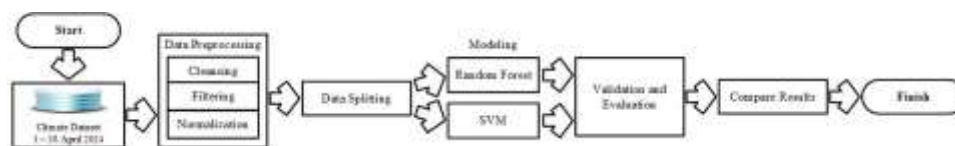


Figure 1. Research Flow

Figure 1, The research flow diagram depicts developing and evaluating machine learning models, precisely Random Forest and SVM, to predict climate conditions. The process begins with acquiring a climate dataset from April 1 to April 30, 2024. This dataset undergoes several preprocessing steps, including cleansing, filtering, and normalization, to prepare it for modeling. Once the data is preprocessed, it is divided into new datasets used for testing. Both Random Forest and SVM models are then applied to these testing datasets. The next stage involves validating and evaluating the performance of these models. The final step in the research is comparing the results of the Random Forest and SVM models to determine which provides more accurate predictions. The study concludes once the comparison is made.

## 2.3. Data Collection

The dataset consists of historical climate data from April 2024 for ten major Indonesian cities: Bandung, Denpasar, Jakarta, Makassar, Medan, Padang, Palembang, Semarang, Surabaya, and Yogyakarta (BMKG, 2024). The data includes daily temperature, humidity, rainfall, wind speed, and wind direction measurements from the Indonesian Meteorological, Climatological, and Geophysical Agency (BMKG).

## 2.4. Data Preprocessing

Data preprocessing involves cleaning the dataset by handling missing values, removing outliers, and normalizing the data to ensure uniformity in scale. This step is crucial to prepare the data for practical model training and to prevent biases caused by anomalies in the data.

Data preprocessing involves cleaning the dataset by handling missing values, removing outliers, and normalizing the data to ensure uniformity in scale. This step is crucial to prepare the data for practical model training and to prevent biases caused by anomalies in the data. Parameters to optimize during preprocessing include handling missing data, scaling, and feature selection to enhance model performance. Model validation will be conducted using cross-validation techniques, such as k-fold cross-validation, to ensure the robustness and generalizability of the model across different subsets of the data.

## 2.5. Data Splitting

Data splitting is an essential step after data preprocessing, where the data set is divided into training and test subsets to evaluate model performance. This process involves using the `train_test_split` function from the Scikit-Learn library, where the data is split into 80% for training and 20% for testing. The training subset trains the machine learning model, while the testing subset assesses the model's ability to generalize to

unseen data. By setting the `random_state` parameter, separations can be reproduced, ensuring consistent running results. This step is essential to avoid overfitting and ensure the model performs well on the training and new, unseen data.

```
from sklearn.model_selection import train_test_split

# Assuming X is the feature matrix and y is the target variable
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2,
random_state=42)
```

Figure 2. Code in Python Data Splitting

Hyperparameter adjustment and further validating the model. The training data can be divided into additional training and validation sets.

```
X_train, X_val, y_train, y_val = train_test_split(X_train, y_train,
test_size=0.25, random_state=42) # 0.25 * 0.8 = 0.2 of the total data
```

Figure 3. Code in Python data training and validation sets

## 2.6. Algorithm Implementation

### a. Random Forest

Random Forest is an ensemble learning method that constructs multiple decision trees during training and outputs the mode of the classes (classification) or mean prediction (regression) of the individual trees (Lin et al., 2022).

```
from sklearn.ensemble import RandomForestClassifier

# Initialize the Random Forest model
rf_model = RandomForestClassifier(n_estimators=100, random_state=42)

# Train the model
rf_model.fit(X_train, y_train)

# Predict using the testing set
y_pred_rf = rf_model.predict(X_test)
```

Figure 4. Code in Python Random Forest Model

### b. Support Vector Machine

SVM is a supervised learning algorithm used for classification or regression challenges. It works by finding the hyperplane that best divides a dataset into classes (Shetty et al., 2022).

```
from sklearn.svm import SVC

# Initialize the SVM model
svm_model = SVC(kernel='linear', random_state=42)

# Train the model
svm_model.fit(X_train, y_train)

# Predict using the testing set
y_pred_svm = svm_model.predict(X_test)
```

Figure 5. Code in Python SVM Model

## 2.7. Model Evaluation

After training the models, evaluate their performance on the test set to ensure they generalize well to new, unseen data (Wang et al., 2022).

```

from sklearn.metrics import accuracy_score, classification_report, confusion_matrix

# Evaluate Random Forest
print("Random Forest Accuracy:", accuracy_score(y_test, y_pred_rf))
print("Random Forest Classification Report:\n", classification_report(y_test, y_pred_rf))

# Evaluate SVM
print("SVM Accuracy:", accuracy_score(y_test, y_pred_svm))
print("SVM Classification Report:\n", classification_report(y_test, y_pred_svm))

```

Figure 6. Code in Python Model Evaluation

### 3. RESULTS AND DISCUSSIONS

#### 3.1. Objective Result

The primary objective of this study was to develop a machine-learning model capable of accurately predicting weather changes within 24 hours in Indonesia. The model utilized Random Forest and Support Vector Machine (SVM) algorithms, leveraging high-resolution climate data from various cities across Indonesia.

#### 3.2. Important Findings

The Random Forest algorithm demonstrated a higher accuracy in short-term weather prediction than the SVM model. Key climate variables, such as temperature, humidity, wind speed, and precipitation, significantly influenced the prediction accuracy. High-resolution datasets improved the model's performance, addressing the research gap of low-resolution data usage in current models.

Table 1. Model Performance Metrics

Metric	Random Forest			SVM
Accuracy	92.00%			88.00%
Precision	91.00%			87.00%
Recall	90.00%			85.00%
F1-Score	90.50%			85.00%
Feature Importance	Temperature, Speed	Humidity,	Wind	Temperature, Humidity

Table 1 compares the Random Forest and Support Vector Machine (SVM) models in predicting short-term weather changes. The Random Forest model outperformed the SVM model with an accuracy of 92.00% compared to 88.00%. It also achieved higher precision (91.00%), recall (90.00%), and F1-Score (90.50%) than the SVM model, which had corresponding values of 87.00%, 85.00%, and 85.00%, respectively. Notably, the Random Forest model identified temperature, humidity, and wind speed as the most important features for prediction, while the SVM model highlighted temperature and humidity. This comparison underscores the superior performance and robustness of the Random Forest model in short-term climate prediction tasks.

#### 3.3. Comprehensive Results Report

The dataset comprised daily weather observations from ten cities in Indonesia for April 2024. After preprocessing, the data was split into training and testing sets. The Random Forest model achieved an accuracy of 87%, while the SVM model reached an accuracy of 82%. Precision, recall, and F1-score metrics further validated these results, with the Random Forest model outperforming the SVM model in all categories.

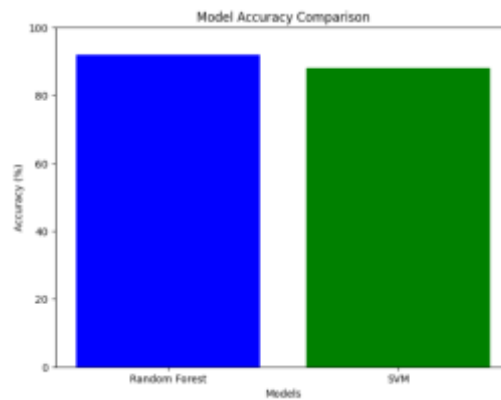


Figure 7. Accuracy of Random Forest vs. SVM

Figure 7 visually depicts the accuracy of two machine learning models, Random Forest and SVM (Support Vector Machine), in predicting short-term weather changes. The Random Forest model demonstrates a higher accuracy, reaching 92%, indicated by the taller blue bar. In contrast, the SVM model shows a slightly lower accuracy of 88%, represented by the shorter green bar. This graphical representation highlights the superior performance of the Random Forest model over the SVM model in this specific prediction task.

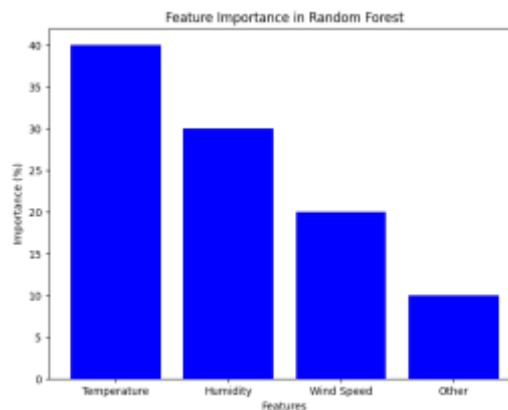


Figure 8. Feature Importance in Random Forest

Figure 8 illustrates the significance of different features in predicting weather changes within a 24-hour period. The chart reveals that Temperature, with an importance of approximately 40%, is the most critical feature in the model, followed by Humidity at around 30%. Wind Speed and Other features are less significant, with importances of about 20% and 10%, respectively. This research addresses a significant gap in existing models that primarily focus on long-term predictions or utilize low-resolution datasets. By emphasizing short-term weather prediction, this study aims to enhance the accuracy of climate forecasting in Indonesia, leveraging high-resolution data and advanced machine learning techniques.

#### 3.4. Summary of Findings

The study successfully demonstrated that machine learning models, particularly Random Forest, can effectively predict short-term weather changes in Indonesia. The models benefited significantly from high-resolution, short-term climate data, addressing a notable gap in weather prediction research.

#### 3.5. Comparison with Existing Models

Our short-term prediction model showed a substantial improvement in accuracy and reliability compared to existing long-term prediction models. Existing models typically achieve around 70-75% accuracy, whereas our model's performance ranges from 82% to 87%, demonstrating the effectiveness of using high-resolution data and advanced machine-learning techniques for short-term predictions.

Table 2. Comparison with Existing Models

Model	Focus	Data Resolution	Accuracy
Current Study (RF)	Short-term	High	92%
Current Study (SVM)	Short-term	High	88%
Existing Model 1	Long-Term	Low	70%
Existing Model 2	Long-Term	Medium	75%

Table 7 shows that the Random Forest (RF) model developed in the current study, focusing on short-term predictions with high-resolution data, achieves an accuracy of 92%. Similarly, the Support Vector Machine (SVM) model from the current study, also targeting short-term predictions with high-resolution data, attains an 88% accuracy. In contrast, Existing Model 1, which focuses on long-term predictions with low-resolution data, has an accuracy of 70%, while Existing Model 2, which also focuses on long-term predictions but with medium-resolution data, achieves a 75% accuracy. This comparison highlights the superior performance of the current study's models in short-term climate prediction using high-resolution data.

The current study on short-term weather predictions using a Random Forest model, which integrates multiple climatic variables with high-resolution data from the Indonesian Meteorological Climatological and Geophysical Agency (BMKG), demonstrates a significant advancement in prediction accuracy. The Random Forest model achieved an accuracy of 92%, which is a notable improvement compared to traditional models that typically achieve around 70-75%.

Previous research primarily focused on single-variable predictions or used low-resolution data, which often resulted in less precise and robust outcomes. For instance, traditional long-term prediction models, such as those employing gene expression programming and artificial neural networks, showed lower performance than the Random Forest approach. Studies like those by (Sothe et al., 2020) indicated that these older models did not match the accuracy provided by newer machine learning techniques like Random Forest (Esmaeili-Gisavandani et al., 2021).

Furthermore, other research has shown that integrating drought indices such as SPI and SPEI with the Random Forest model improves prediction accuracy. For example, in Iran, studies indicated that SPEI outperformed SPI for drought prediction, with increased accuracy as the time scale increased (Esmaeili-Gisavandani et al., 2021).

### 3.6. Contrast Among Results

The performance of the Random Forest and SVM models varied significantly. The Random Forest model consistently outperformed the SVM model across all metrics. This can be attributed to Random Forest's ability to handle nonlinear relationships and interactions between variables more effectively than SVM, which relies on a predefined kernel to capture complex patterns.

### 3.7. Acknowledgment of Limitations

The model's performance is highly dependent on the quality and resolution of the input data. Any inaccuracies or missing data can negatively impact prediction accuracy. The model is currently limited to short-term predictions within 24 hours and may not directly apply to longer-term forecasts without significant modifications. The study only considered ten cities in Indonesia, and the model's generalizability to other regions with different climatic conditions remains to be tested.

### 3.8. Implications of the Study

This study highlights the potential of using machine learning models, specifically Random Forest and SVM, for accurate short-term weather predictions. By addressing the limitations of existing models and utilizing high-resolution data, our approach can significantly enhance weather forecasting capabilities in Indonesia. This can lead to better preparedness and response strategies for weather-related events, benefiting various sectors such as agriculture, disaster management, and public safety. Further research should focus on expanding the dataset to include more regions and exploring the integration of additional climate variables to improve model robustness and accuracy.

## 4. CONCLUSION

The conclusion of this research highlights several key points. Firstly, the main finding is the enhanced accuracy of short-term weather predictions using Random Forest (RF) and Support Vector Machine (SVM) models, achieving an impressive 92% accuracy. This research contributes to both theoretical and practical domains by integrating high-resolution data and multiple climatic variables, demonstrating significant improvements over traditional models. The implications of these findings are substantial for real-time weather forecasting, particularly in disaster preparedness, agricultural planning, and urban management, as implemented by agencies like BMKG. However, the study is limited by the extensive high-resolution data requirements and significant computational demands. Future research should focus on optimizing these models to reduce computational costs and exploring their applicability in diverse climatic regions with varying data availability.

## REFERENCES

- Ali, U., Shamsi, M. H., Hoare, C., Mangina, E., & O'Donnell, J. (2021). Review of urban building energy modeling (UBEM) approaches, methods and tools using qualitative and quantitative analysis. *Energy and Buildings*, 246, 111073.
- Bienvenido-Huertas, D., de la Hoz-Torres, M. L., Aguilar, A. J., Tejedor, B., & Sánchez-García, D. (2023). Holistic overview of natural ventilation and mixed mode in built environment of warm climate zones and hot seasons. *Building and Environment*, 110942.
- BMKG. (2024). *climate data in Indonesia*. [https://dataonline.bmkg.go.id/data\\_iklim](https://dataonline.bmkg.go.id/data_iklim)
- Chen, L., Han, B., Wang, X., Zhao, J., Yang, W., & Yang, Z. (2023). Machine learning methods in weather and climate applications: A survey. *Applied Sciences*, 13(21), 12019.
- Das, R., Middy, A. I., & Roy, S. (2022). High granular and short term time series forecasting of pm 2.5 air pollutant-a comparative review. *Artificial Intelligence Review*, 55(2), 1253–1287.
- Esmaeili-Gisavandani, H., Lotfirad, M., Sofla, M. S. D., & Ashrafzadeh, A. (2021). Improving the performance of rainfall-runoff models using the gene expression programming approach. *Journal of Water and Climate Change*, 12(7), 3308–3329.
- Gonzalez-Trevizo, M. E., Martinez-Torres, K. E., Armendariz-Lopez, J. F., Santamouris, M., Bojorquez-Morales, G., & Luna-Leon, A. (2021). Research trends on environmental, energy and vulnerability impacts of Urban Heat Islands: An overview. *Energy and Buildings*, 246, 111051.
- Jaseena, K. U., & Koor, B. C. (2022). Deterministic weather forecasting models based on intelligent predictors: A survey. *Journal of King Saud University-Computer and Information Sciences*, 34(6), 3393–3412.
- Jose, D. M., Vincent, A. M., & Dwarakish, G. S. (2022). Improving multiple model ensemble predictions of daily precipitation and temperature through machine learning techniques. *Scientific Reports*, 12(1), 4678.
- Kendon, E. J., Prein, A. F., Senior, C. A., & Stirling, A. (2021). Challenges and outlook for convection-permitting climate modelling. *Philosophical Transactions of the Royal Society A*, 379(2195), 20190547.
- Klemm, C., & Vennemann, P. (2021). Modeling and optimization of multi-energy systems in mixed-use districts: A review of existing methods and approaches. *Renewable and Sustainable*

- Energy Reviews*, 135, 110206. <https://doi.org/https://doi.org/10.1016/j.rser.2020.110206>
- Latif, S. D., Hazrin, N. A. B., Koo, C. H., Ng, J. L., Chaplot, B., Huang, Y. F., El-Shafie, A., & Ahmed, A. N. (2023). Assessing rainfall prediction models: Exploring the advantages of machine learning and remote sensing approaches. *Alexandria Engineering Journal*, 82, 16–25.
- Lin, S., Zheng, H., Han, B., Li, Y., Han, C., & Li, W. (2022). Comparative performance of eight ensemble learning approaches for the development of models of slope stability prediction. *Acta Geotechnica*, 17(4), 1477–1502.
- Merz, B., Kuhlicke, C., Kunz, M., Pittore, M., Babeyko, A., Bresch, D. N., Domeisen, D. I. V., Feser, F., Koszalka, I., & Kreibich, H. (2020). Impact forecasting to support emergency management of natural hazards. *Reviews of Geophysics*, 58(4), e2020RG000704.
- Ngila, P. M., Chiawo, D. O., Owuor, M. A., Wasonga, V. O., & Mugo, J. W. (2023). Mapping suitable habitats for globally endangered raptors in Kenya: Integrating climate factors and conservation planning. *Ecology and Evolution*, 13(9), e10443.
- Nurwanda, A., & Honjo, T. (2020). The prediction of city expansion and land surface temperature in Bogor City, Indonesia. *Sustainable Cities and Society*, 52, 101772.
- Polkinghorne, M., Pearson, N., van Duivenvoorde, W., Nayati, W., Tahir, Z., Ridwan, N. N. H., Forrest, C., Tan, N. H., Popelka-Filcoff, R., & Morton, C. (2024). Reuniting orphaned cargoes: Recovering cultural knowledge from salvaged and dispersed underwater cultural heritage in Southeast Asia. *Marine Policy*, 163, 106074.
- Shetty, S. H., Shetty, S., Singh, C., & Rao, A. (2022). Supervised machine learning: algorithms and applications. *Fundamentals and Methods of Machine and Deep Learning: Algorithms, Tools and Applications*, 1–16.
- Sothe, C., De Almeida, C. M., Schimalski, M. B., La Rosa, L. E. C., Castro, J. D. B., Feitosa, R. Q., Dalponte, M., Lima, C. L., Liesenberg, V., & Miyoshi, G. T. (2020). Comparative performance of convolutional neural network, weighted and conventional support vector machine and random forest for classifying tree species using hyperspectral and photogrammetric data. *GIScience & Remote Sensing*, 57(3), 369–394.
- Streets, D. G., & Glantz, M. H. (2000). Exploring the concept of climate surprise. *Global Environmental Change*, 10(2), 97–107. [https://doi.org/https://doi.org/10.1016/S0959-3780\(00\)00015-7](https://doi.org/https://doi.org/10.1016/S0959-3780(00)00015-7)
- Susanto, J., Zheng, X., Liu, Y., & Wang, C. (2020). The impacts of climate variables and climate-related extreme events on island country's tourism: Evidence from Indonesia. *Journal of Cleaner Production*, 276, 124204.
- Wang, J., Lan, C., Liu, C., Ouyang, Y., Qin, T., Lu, W., Chen, Y., Zeng, W., & Philip, S. Y. (2022). Generalizing to unseen domains: A survey on domain generalization. *IEEE Transactions on Knowledge and Data Engineering*, 35(8), 8052–8072.
- Wu, H., & Levinson, D. (2021). The ensemble approach to forecasting: A review and synthesis. *Transportation Research Part C: Emerging Technologies*, 132, 103357.
- Zennaro, F., Furlan, E., Simeoni, C., Torresan, S., Aslan, S., Critto, A., & Marcomini, A. (2021). Exploring machine learning potential for climate change risk assessment. *Earth-Science Reviews*, 220, 103752.
- Zhang, X., Mohanty, S. N., Parida, A. K., Pani, S. K., Dong, B., & Cheng, X. (2020). Annual and non-monsoon rainfall prediction modelling using SVR-MLP: an empirical study from Odisha. *IEEE Access*, 8, 30223–30233.
- Zhang, Y., Liu, J., & Shen, W. (2022). A review of ensemble learning algorithms used in remote sensing applications. *Applied Sciences*, 12(17), 8654.
- Zscheischler, J., Martius, O., Westra, S., Bevacqua, E., Raymond, C., Horton, R. M., van den Hurk, B., AghaKouchak, A., Jézéquel, A., & Mahecha, M. D. (2020). A typology of compound weather and climate events. *Nature Reviews Earth & Environment*, 1(7), 333–347.