Enhancing Flood Severity Prediction through a Combined Threshold-Based Alert Algorithm and Random Forest Classifier

Aksshat Govind
Department of Computer Science
KIET Group of Institutions
Delhi-NCR, Ghaziabad
aksshat.2125cs1189@kiet.edu

Aishwarya Gupta
Department of Computer Science
KIET Group of Institutions
Delhi-NCR, Ghaziabad
aishwarya.2125cs1148@kiet.edu

Aviral Katiyar
Department of Computer Science
KIET Group of Institutions
Delhi-NCR, Ghaziabad
aviral.2125cs1067@kiet.edu

Chirag Tyagi
Department of Computer Science
KIET Group of Institutions
Delhi-NCR, Ghaziabad
chirag.2125cs1202@kiet.edu

Sreesh Gaur
Department of Computer Science
KIET Group of Institutions
Delhi-NCR, Ghaziabad
sreesh.gaur@kiet.edu

***Abstract*—Flooding is a significant and recurrent is- sue that poses severe risks to communities, economies, and environments. Traditional flood prediction methods, which often rely on fixed thresholds, may lack adaptability to changing conditions. Meanwhile, machine learning (ML) models like Random Forest Classifiers (RFCs) offer advanced predictive capabilities, but may not provide straightforward alert mechanisms. This research proposes a novel approach that integrates a Threshold-Based Alert Algorithm (TBA) with an RFC to enhance flood severity prediction. The combined model leverages the strengths of both methodologies: immediate, interpretable alerts from the TBA and adaptive, data-driven insights from the RFC. Using historical and real-time data from river basins in Wisconsin and New York, the hybrid model demonstrates improved accuracy and reliability over traditional meth- ods and standalone RFCs. The results indicate that this integrated approach offers a more effective tool for flood management and disaster mitigation.**

Keywords—Flood forecasting, threshold-based alert algorithm (TBA), random forest classification (RFC), machine learning (ML), hybrid models, flood management, data integration, predictive analytics, environmental data, disaster risk reduction.

I. Introduction

Flooding represents a major threat to infrastructure, agriculture, and human safety, emphasizing the critical need for accurate and timely flood prediction systems to mitigate associated risks [1]. Traditionally, flood prediction relies on predefined thresholds that, while straightforward, can be limited in their ability to adapt to dynamic and evolving conditions. These threshold-based methods, although useful for generating immediate alerts, often lack the flexibility required to handle complex and variable flood scenarios [2].

In contrast, machine learning (ML) models have demonstrated strong predictive power in time-series applications, as evidenced in applications like radon level forecasting [3]. Studies by Gao et al. [4] and Islam et al. [5] highlight the efficacy of machine learning models in weather forecasting and precipitation prediction. However, while these models excel in refining predictions through sophisticated data analysis, they may fall short in providing immediate, interpretable alerts, which are crucial for real-time decision-making [6].

This paper proposes a hybrid approach that combines the strengths of both traditional and modern methodologies. We integrate a Threshold-Based Alert Algorithm (TBA) to generate initial flood alerts based on predefined conditions, alongside a Random Forest Classifier (RFC) for enhanced flood severity classification. The goal of this integration is to improve prediction accuracy, offer actionable insights, and provide a more comprehensive tool for flood management. By leveraging the immediate, interpretable alerts from the Threshold-Based Alert Algorithm and the advanced analytical capabilities of the Random Forest Classifier, this approach aims to bridge the gap between traditional flood forecasting and modern machine learning techniques. Ultimately, this model enhances flood prediction and response strategies, providing an adaptable and efficient solution for real-time flood management.

II. Related Work

Numerous studies have advanced flood prediction through machine learning and threshold based approaches, yet challenges remain in balancing immediacy with precision. The author of [6] addressed climate impact on precipitation and temperature using Random Forest (RF) algorithms with seasonal bias correction, improving accuracy in dynamic forecasting but lacking real-time alert simplicity. The author of [4] implemented a hybrid spatial temporal model for sea surface temperature, highlighting the benefits of multiscale analysis for short-term and long-term forecasting, though not directly related to actionable flood alerts. Similarly, the author of [5] leveraged location-agnostic deep learning models for precipitation, aiming for general applicability but limited by model transparency in immediate alerting scenarios.

979-8-3315-3038-9/25/$31.00 ©2025 IEEE

Meanwhile, traditional hydrological models, like those examined in [7], [8], and [9], have enhanced satellite data assimilation for European Centre for Medium-Range Weather Forecasts (ECMWF) forecasts, underscoring the critical role of accurate moisture data. However, these models tend to be complex and less adaptable for instant flood alerts. Studies on threshold-based systems in [9] have been fundamental for flood alerts, although with limited adaptability to evolving environmental conditions.

The proposed integration of a Threshold-Based Alert Algorithm (TBA) with a Random Forest Classifier (RFC) aims to bridge these gaps by offering immediate, inter- pretable alerts alongside adaptable, data-driven severity classification. This hybrid model provides a refined tool that leverages real-time data for accuracy and responsive- ness, enhancing the utility of flood forecasting systems.

III. Proposed Method

A. Overview

In this section, we propose a novel approach to flood severity prediction that integrates a Threshold-Based Alert Algorithm (TBA) with a Random Forest Classifier (RFC). This combined method aims to leverage the immediate actionable alerts provided by the Threshold-Based Alert Algorithm and the refined predictive capa- bilities of the Random Forest Classifier to enhance flood severity prediction. The concept of thresholding, essential in digital image processing for accurate segmentation [10], is similarly critical here for flood prediction.

In the Threshold-Based Alert Algorithm, thresholds are applied to meteorological and hydrological data, generating initial flood alerts that respond to environmental conditions. These thresholds work similarly to segmentation in digital image processing, where a global or adaptive local threshold can be applied for different scenarios [10]. By setting precise thresholds, the TBA component ensures immediate response capabili- ties, while the Random Forest Classifier offers detailed classification to improve overall prediction accuracy.

The proposed methodology involves a two-step process; initial alert generation through threshold-based rules and refined severity classification using machine learning.

B. Threshold-Based Alert Algorithm (TBA)

The Threshold-Based Alert Algorithm (TBA), as illustrated in Fig. 1, operates by applying predefined thresholds to real-time meteorological and hydrological data, generating initial flood alerts. These thresholds are determined on the basis of historical flood events, expert judgment, and regional guidelines. Key parameters may include rainfall intensity, river stage levels, and soil moisture content.

1. Define thresholds T_i for flood indicators X_i based on historical data and expert guidelines. For example, if X_1 represents rainfall intensity, and X_2 represents river stage, set thresholds T_1 and T_2 such that:

$$\boldsymbol{T_i = \textit{Critical value for } X_i(\textit{e.g., } X_i \textit{ exceeds } T_i)} \tag{1}$$

2. Continuously monitor real-time data $x_i(t)$ for each indicator. Compare the observed data with the thresholds. For each indicator X_i:

$$\boldsymbol{Alert_i(t) = 1 \textit{ if } x_i(t) > T_i, 0 \textit{ otherwise}} \tag{2}$$

3. Generate an overall alert based on whether any of the indicators exceed their thresholds. If any $Alert_i(t) = 1$, trigger a flood alert:

$$\boldsymbol{Flood\ Alert(t) = max(Alert_1(t), Alert_2(t), .., Alert_n(t))} \tag{3}$$

Alerts can be categorized into levels depending on the magnitude of exceedance.

4. Periodically update thresholds T_i based on new data. This involves recalibrating thresholds by ana- lyzing historical flood events and adjusting them to improve accuracy.

```
//Define thresholds Ti for each flood indicator Xi based on historical data
T1 = Critical value for X1 (e.g., rainfall intensity)
T2 = Critical value for X2 (e.g., daily runoff)
...
Tn = Critical value for Xn

//Monitoring and Alert Generation
For each timestamp t (real-time data):
    if t:
        For each flood indicator Xi:
            // Compare observed data with threshold
            if xi(t) > Ti:
                Alerti(t) = 1  // Trigger alert if threshold is exceeded
            else:
                Alerti(t) = 0  // No alert

        // Generate overall flood alert
        FloodAlert(t) = max(Alert1(t), Alert2(t), ..., Alertn(t))

    else:
        continue
```

Fig. 1. Threshold-Based Alert Algorithm

C. Random Forest Classifier (RFC)

The Random Forest Classifier (RFC), illustrated in Fig. 2, enhances the accuracy of flood severity predictions by leveraging historical flood records, mete- orological data (e.g., rainfall, temperature, humidity), and hydrological information (e.g., river discharge, soil moisture). The data undergoes preprocessing to normalize values, address missing entries, and select relevant features. This ensures that the Random Forest Classifier model is trained on high-quality data, optimizing its performance in predicting flood severity.

Random Forest algorithms combine multiple decision tree predictors, with each tree relying on a randomly sampled subset of features, as outlined by [11]. This technique reduces error rates and enhances robustness against noise, providing a distinct advantage over other classifiers such as Adaptive Boosting (AdaBoost). The generalization error of the Random Forest Classifier converges to a limit as the number of trees increases, depending on the individual strength of each tree and their mutual correlation. The random feature selection at each node split yields a balanced model that effectively handles diverse environmental variables crucial for flood prediction. Additionally, the internal error estimates in the Random Forest Classifier monitor strength and correlation among trees, enhancing the model's reliability and enabling the measurement of variable importance—a critical factor in assessing flood-related parameters.

1. Collect historical data (x_i,y_i), where x_i represents the feature set and y_i represents the flood severity class. Preprocess the data by normalizing features, handling missing values, and splitting the dataset into training and testing subsets.

2. Train the RFC model using the dataset with N samples. The Random Forest algorithm constructs T decision trees by randomly selecting subsets of features and data for each tree. For a given feature set x, the RFC output is the majority vote across all decision trees:

$$\mathbf{RFC(x) = (\{Tree_1(x),\ Tree_2(x),....,Tree_T\ (x)\})} \quad (4)$$

3. Apply the trained RFC to new data x. Each decision tree in the forest predicts a flood severity class, and the final classification is determined by aggregating these predictions:

$$\mathbf{Severity\ Prediction = mode(Tree_1(x),Tree_2(x),.,Tree_T(x))} \quad (5)$$

Alerts can be categorized into levels depending on the magnitude of exceedance.

4. Assess the RFC model using performance metrics such as accuracy A, precision P, recall R, and $F1$ score. For classification:

$$\mathbf{A = (TP + TN)/(FP + FN + TP + TN)} \quad (6)$$

$$\mathbf{P = (TP)/(FP + TP)} \quad (7)$$

$$\mathbf{R = (TP)/(FN + TP)} \quad (8)$$

$$\mathbf{A = (2.P.R)/(P + R)} \quad (9)$$

Where *TP* is true positives, *TN* is true negatives, *FP* is false positives, and *FN* is false negatives.

D. Integration of Threshold-Based Alerting (TBA) and Random Forest Classifier (RFC)

The integration of the TBA and RFC combines the strengths of both methods. The TBA provides immediate alerts based on predefined thresholds, offering quick, actionable warnings. The RFC then refines these alerts by providing detailed flood severity classifications based on a comprehensive analysis of historical and real-time data.

1. Collect real-time data for both TBA and RFC components. Ensure data consistency and accuracy.

2. Use TBA to generate initial alerts based on predefined thresholds T_i. This provides early warnings:

$$\mathbf{Initial\ Alert(t) = Flood\ Alert(t)\ from\ TBA} \quad (10)$$

3. Prepare data for RFC by augmenting it with additional historical records and features. This involves

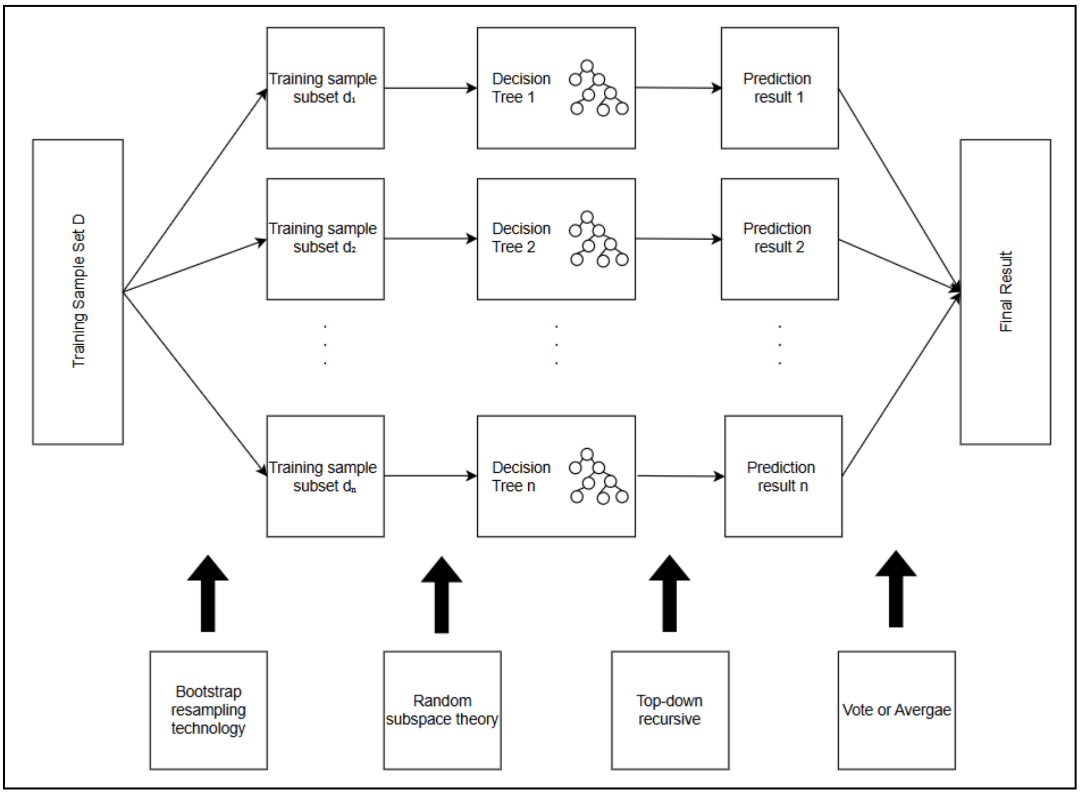

Fig. 2. Main Structure of Random Forest

techniques like data normalization and feature extraction.

4. Use the RFC model to classify the severity of the flood risk based on augmented data:

$$\textbf{Severity Classification} = \textbf{RFC(x)} \tag{11}$$

5. Integrate results from TBA and RFC. Use TBA alerts as initial warnings and RFC for detailed severity classification. Communicate the combined results to stakeholders.

$$\textbf{Combined Result} = \textbf{RFC(x) with Flood Alert(t)} \tag{12}$$

IV. Study Area & Dataset

A. Study Area

The study focuses on two significant river basins in the United States: the Wisconsin River Basins and the New York State River Basins, as illustrated in Fig. 3.

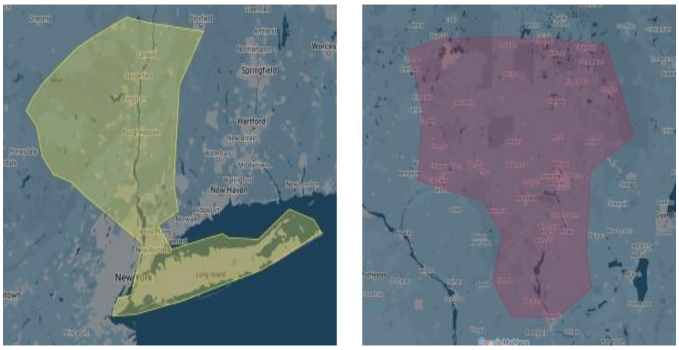

Fig. 3. Hudson Basin and Delaware Basin(New York)-*Left*; Central Wisconsin Basin and Fox River Basin(Wisconsin)-*Right*

Wisconsin River Basins: Located in the state of Wisconsin, this basin encompasses a diverse range of topographical and climatic features, characterized by a mixture of flatlands, hills, and valleys. Wisconsin has a long history of stakeholder engagement in watershed-based resource management, with an increase in river organization formation since the 1990s, particularly in southeastern regions like the Fox River watershed. This shift reflects the growing recognition of the need to manage water resources on a watershed scale, as highlighted in studies comparing lake and river organizations [12]. The basin experiences a continental climate with cold winters and warm, humid summers, with average annual temperatures ranging from 5°C to 10°C and significant precipitation primarily from spring to autumn. The Wisconsin River and its tributaries play a crucial role in regional water management and flood dynamics.

New York State River Basins: This area includes several prominent river basins, such as the Hudson River Basin and the Finger Lakes region. New York's water supply infrastructure faces critical repair needs, with significant water losses due to leaks in the aqueducts of New York City's supply system, which has led to issues like chronic flooding in some communities. This challenge underscores the importance of integrated regional water management [13]. The topography in New York varies from the mountainous Adirondacks to the fertile Mohawk Valley. The climate is predominantly humid continental, with cold winters and warm, humid summers; average annual temperatures range from 7°C to 12°C, and precipitation is well-distributed throughout the year. This complex water system, including numerous lakes and rivers, heightens the region's flood risk profile and reinforces the need for coordinated watershed management practices.

B. Dataset

This research utilizes two distinct datasets, TABLE I: the Wisconsin dataset and the New York dataset. Each dataset includes a range of measurements and features essential for flood prediction modeling. The data collected are organized into standardized columns following the format '*Basins_Parameter_Unit*', where Basins represent the specific water basin (e.g., Wind Lake or Albany), Parameter refers to the specific type of mea- surement (e.g., Gage Height or Temperature), and Unit represents the measurement unit (e.g., feet or Celsius). This column naming format ensures consistency and clarity across the datasets. The details of these datasets are presented visually below.

Wisconsin Dataset, TABLE II — This dataset encompasses key water measurements and the following features were found to be most relevant for flood prediction: Gage height and Gage height for Tailwater.

New York Dataset, TABLE III — This dataset includes similar features but specific to New York's river basins and water infrastructure, with relevant features like: minimum, maximum and mean temperatures.

Both datasets are structured to ensure consistency and facilitate integration during analysis, with column names such as A_Temp_degC or A_GH_ft, representing the temperature in Celsius for Wind Lake Basin (Area A) and the Gage Height in feet for Albany Basin (Area A), respectively.

Table I. Comparison of River Basin Data

<table>
<tr><th colspan="6">RIVER BASINS</th></tr>
<tr><th colspan="3">WISCONSIN</th><th colspan="3">NEW YORK</th></tr>
<tr><td colspan="3">Wind Lake, Waterford, Shawano, Westfield, Green Lake, Stoughton, Green Leaf</td><td colspan="3">Albany, Neversink, Kelsey, Lew Beach, Schoharie Reservoir, Skaneateles,Port Henry, Rouses Point, Dunham Basin,Cranberry Lake</td></tr>
<tr><th>Feature Importance</th><th>Threshold</th><th>Flood Severity</th><th>Feature Importance</th><th>Threshold</th><th>Flood Severity</th></tr>
<tr><td>A_GH_ft_Rolling_Mean_3: 80.7%</td><td rowspan="3">minThreshold: 5th percentile
maxThreshold: 95th percentile</td><td>Low: A_GH_ft < minThreshold</td><td>H_MinT_degC: 78.3%</td><td rowspan="3">minThreshold: 8th percentile
maxThreshold: 92nd percentile</td><td>Low: J_D-IGH_ft < minThreshold</td></tr>
<tr><td>A_GH_ft_Lag1: 10.0%</td><td>Moderate: minThreshold ≤ A_GH_ft < maxThreshold</td><td>C3_MaxT_degC: 15.2%</td><td>Moderate: minThreshold ≤ J_D-IGH_ft < maxThreshold</td></tr>
<tr><td>A_GHTW_ft_B_D_SPinches_Interaction: 4.5%</td><td>High: A_GH_ft ≥ maxThreshold</td><td>H_MeanT_degC: 4.8%</td><td>High: J_D-IGH_ft ≥ maxThreshold</td></tr>
</table>

Table II. Wisconsin Dataset (Table truncated for brevity)

Date	*A_GH_ft*	*A_GHTW_ft*	*A1_D-SP_inches*	*A2_D-SP_inches*	…	*G_MeanRH_%*
2023-08-29	8.3	4.8	0.2	0.3	…	84.6
2023-08-30	8.3	4.8	0.2	0.3	…	84.6
2023-08-31	8.3	4.8	0.2	0.3	…	84.6
2023-08-01	8.3	4.8	0	0.1	…	67.7
2023-08-02	8.3	4.8	0	0.1	…	67.7
…	…	…	…	…	…	…
2024-08-29	8.4	5	1.5	0.8	…	74.6

Table III. New YorkDataset (Table truncated for brevity)

Date	*A1_MaxT_degC*	*A1_MinT_degC*	*A1_MeanT_degC*	*A2_MaxT_degC*	…	*J_D-IGH_ft*
2023-08-28	15.4	14.4	14.8	14.8	…	16.3
2023-08-29	15.4	14.4	14.8	14.8	…	16.3
2023-08-30	15.4	14.4	14.8	14.8	…	16.3
2023-08-31	15.4	14.4	14.8	14.8	…	16.3
2023-08-01	15.4	14.4	14.8	14.8	…	16.3
…	…	…	…	…	…	…
2024-08-28	27.9	26.7	27.1	27.8	…	16.7

C. Output

The combined Threshold-Based Alert (TBA) and Random Forest Classifier (RFC) model outperforms several traditional and machine learning models in flood severity prediction. In Wisconsin, it achieved 84% accuracy with an MSE of 0.02, and in New York, it achieved 98% accuracy with an MSE of 0.004, in Fig.4.

```
Overall Accuracy (via 5-fold cross-validation): 0.8482
Overall Accuracy (via 10-fold cross-validation): 0.9807
```

Fig. 4. Model accuracies

V. Results

A. Model Performance

Testing parameters were evaluated, and accuracy scores were calculated for both datasets. The results are illustrated in Fig.5, Fig.6, and Fig.7. Fig.5 presents a bar chart of the monthly Proportional Accuracy for New York in 2024, showing fluctuating accuracy with a significant drop in April to 0.63, while other months maintain values around 0.9. Fig.6 displays a similar bar chart for Wisconsin in 2024, reflecting the same trends observed in New York. Fig.7 shows the ROC curve for three classes, highlighting the classifier's performance across various thresholds, with higher curves signifying better model performance.

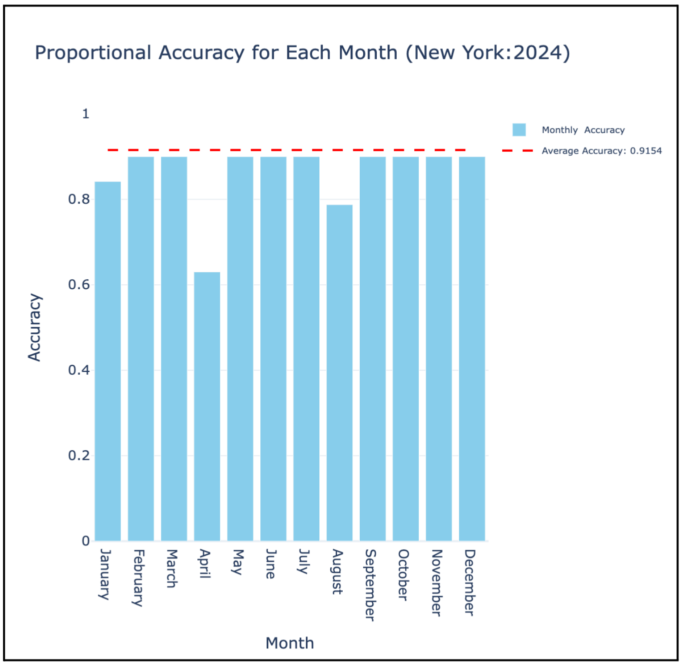

Fig. 5. Accuracy Scores month-wise of basins in New York

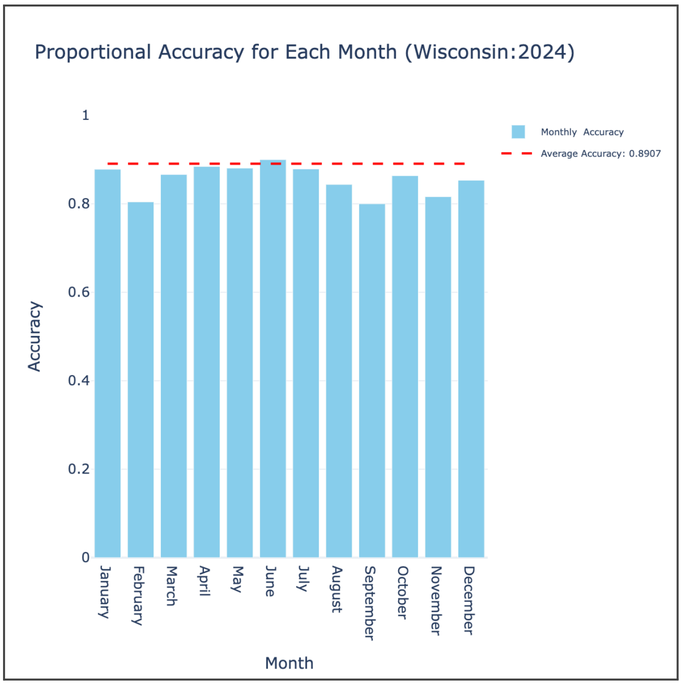

Fig. 6. Accuracy Scores month-wise of basins in Wisconsin

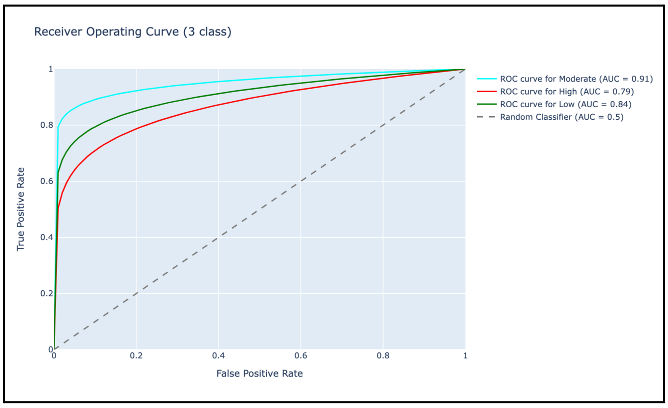

Fig. 7. Receiver Operating Characteristic (ROC) Curve for three classes

B. *Threshold Calculation*

Flood severity thresholds were computed from historical data distributions. These thresholds provided initial flood alerts that were further refined using the RFC, enhancing the model's precision.

C. *Model Comparison*

The integrated TBA and RFC approach outperformed traditional threshold-based methods and standalone RFC models. Compared to traditional threshold-based models (70-80% accuracy) [14] [15], Decision Trees (82%) [16], and Support Vector Machines (83%) [17], the TBA-RFC model significantly improves accuracy. Although deep learning models like RNNs (90%) [18] and CNNs (87%) [19] achieve high accuracy, they require large datasets and computational resources, while the TBA-RFC model performs efficiently with fewer resources. Hybrid models achieve 85% [20] accuracy, but the TBA-RFC model offers a more integrated and effective solution. Therefore, the TBA-RFC model delivers higher accuracy and computational efficiency, making it a superior choice for flood severity prediction.

Conclusion

This study demonstrates that the TBA-RFC model is highly effective in flood severity prediction, offering superior accuracy and computational efficiency compared to traditional and other machine learning models. The model achieved an impressive 98% accuracy in New York with an MSE of 0.004, and 84% accuracy in Wisconsin with an MSE of 0.02, as shown in Fig. 4. Its performance consistently outperforms standalone threshold-based models and RFC techniques, with monthly proportional accuracy maintaining strong results across datasets. While deep learning models like RNNs and CNNs require large datasets and substantial computational resources, the TBA-RFC model provides comparable results with fewer resources. Achieving higher accuracy than traditional methods, such as threshold-based models, Decision Trees, and Support Vector Machines, the TBA-RFC model proves to be an integrated and efficient solution for flood severity prediction. In conclusion, the TBA-RFC model represents a significant advancement by combining high accuracy with low computational cost, offering a promising tool for transforming disaster management practices. Future work could explore expanding its use to other regions and disaster types, further solidifying its impact on predictive analytics in critical scenarios.

References

[1] Intergovernmental Panel on Climate Change (IPCC) (2007), Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the In- tergovernmental Panel on Climate Change, edited by S. Solomon et al., Cambridge Univ. Press, Cambridge, U. K.

[2] Douville, H., P . Viterbo, J.-F . Mahfouf, and A. C. M. Beljaars (2000), Evaluation of the optimum interpolation and nudging techniques for soil moisture analysis using FIFE data, Mon. Weather Rev., 128, 1733–1756, doi:10.1175/1520- 0493(2000)128¡1733:EOTOIA¿2.0.CO;2.

[3] Valdo Mpinga, A. M. Rosado Da Cruz, Se´rgio Ivan Lopes, ”Forecasting Short-Term Indoor Radon: A Machine Learning Approach Using LSTM Networks”, 2023 18th Iberian Conference on Information Systems and Technologies (CISTI), pp.1-6, 2023.

[4] G. Gao, B. Yao, Z. Li, D. Duan and X. Zhang, ”Forecasting of Sea Surface Temperature in Eastern Tropical Pacific by a Hybrid Multiscale Spatial–Temporal Model Combining Error Correction Map,” in IEEE Transactions on Geoscience and Re- mote Sensing, vol. 62, pp. 1-22, 2024, Art no. 4204722, doi: 10.1109/TGRS.2024.3353288.

[5] M. S. Islam, M. S. Rahman, M. S. Ul Haque, F . A. Tumpa, M. S. Bin Hossain and A. A. Arabi, ”Location Agnostic Adaptive Rain Precipitation Prediction using Deep Learning, ” 2023 IEEE 9th International Women in Engineering (WIE) Conference on Electrical and Computer Engineering (WIECON-ECE), Thiruvananthapuram, India, 2023, pp. 148-153, doi: 10.1109/WIECON-ECE60392.2023.10456479.

[6] T. Tang, T. Liu and G. Gui, ”Forecasting Precipitation and Temperature Evolution Patterns Under Climate Change Using a Random Forest Approach With Seasonal Bias Correction,” in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 17, pp. 12609-12621, 2024, doi: 10.1109/JSTARS.2024.3425639.

[7] Andersson, E., et al. (2005), Assimilation and modeling of the hydrological cycle in the ECMWF forecasting system, Bull. Am. Meteorol. Soc., 86, 387–402, doi:10.1175/BAMS-86-3-387.

[8] Andersson, E., E. Ho´lm, P . Bauer, A. Beljaars, G. A. Kelly, A. P . McNally, A. J. Simmons, J.-N. The´paut, and A. M. Tompkins (2007), Analysis and forecast impact of the main humidity observing systems, Q. J. R. Meteorol. Soc., 133, 1473–1485.

[9] Dee, D. P ., and S. Uppala (2009), Variational bias correction of satellite radiance data in the ERA-Interim reanalysis, Q. J. R. Meteorol. Soc., 135, 1830–1841, doi:10.1002/qj.493.

[10] Niu, Zuodong & Li, Handong. (2019). Research and analysis of threshold segmentation algorithms in image processing. Journal of Physics: Conference Series. 1237. 022122. 10.1088/1742-6596/1237/2/022122.

[11] Cutler, Adele & Cutler, David & Stevens, John. (2011). Random Forests. 10.1007/978-1-4419-9326-7 5.

[12] Slawski, Thomas. (2013). River Basin Organization development in southeastern Wisconsin. Lakes & Reservoirs Research & Management. 18. 53-65. 10.1111/lre.12014.

[13] Negro, S. & Porter, K.S.. (2009). Water stress in New York State: The regional imperative?. 20.5-16.

[14] Jha, M. K., et al., ”Flood forecasting using threshold-based approach,” *Journal of Hydrology*, 2012. International Conference on Flood Prediction, 2012.

[15] Van Oort, B., et al., ”Threshold-based flood prediction models for the UK,” *Hydrology and Earth System Sciences*, 2016.European Geophysical Union General Assembly, 2016.

[16] Xia, J., et al., ”Flood prediction using decision trees,” *Environ- mental Modelling Software*, 2015. International Conference on Environmental Modelling, 2015.

[17] Nikolopoulos, N., et al., ”Flood forecasting with support vector

machines," *Hydrological Sciences Journal*, 2019.International Conference on Hydrological Sciences, 2019.

[18] Shi, Z., et al., "Flood prediction using recurrent neural networks," *Water Resources Research*, 2019. IEEE International Conference on Neural Networks, 2019.

[19] Wang, Y., et al., "Convolutional neural networks for flood pre- diction using satellite data," *Journal of Hydrology*, 2020. Inter- national Conference on Machine Learning for Earth Sciences, 2020.

[20] Rashid, M., et al., "Hybrid threshold-based and machine learning model for flood prediction," *Environmental Modelling Software*, 2018. Proceedings of the International Conference on Environ- mental Science and Technology, 2018.