**ORIGINAL PAPER**

# Application of random forest (RF) for flood levels prediction in Lower Ogun Basin, Nigeria

O. O. Aiyelokun[1] · O. D. Aiyelokun[2] · O. A. Agbede[1]

## Abstract

This study evaluates the performance of random forest (RF) for predicting flood levels in the Lower Ogun Basin, Southwest Nigeria. Daily flood levels for a period of 36 years (1981 to 2016), recorded at Mokoloki weir, were obtained from the Ogun–Oshun River Basin Development Authority (OORBDA). Descriptive statistics were employed to provide concise information on the flood levels, and trend and autocorrelation assessments were performed using the Mann–Kendall test and the Ljung–Box test, respectively, at 95% confidence level. Antecedent daily flood levels of up to 7 days were selected as input features for the RF model to predict daily flood levels. To develop the RF model, the dataset was divided into train (70%), validation (15%), and test (15%). The performance of the RF model was evaluated using Mean Absolute Error (MAE), coefficient of determination ($R^2$), Nash–Sutcliffe Efficiency Coefficient (NSEC), and Kling-Gupta efficiency (KGE). The study reveals that the highest flood level was 9.5 m, while 75% of the records were less or equal to 7.04 m. The flood level had a significant positive trend (tau = 0.19, 2-sided $p$ value < 0.05) and a significant autocorrelation ($X$-squared = 13,059, df = 1, $p$ value < 0.05). Based on the evaluation criteria, RF is reliable in predicting daily flood levels, having performed well at both the validation (MAE = 0.0484, $R^2$ = 0.9924, NSCE = 0.9924, KGE = 0.9930) and test (MAE = 0.0519, $R^2$ = 0.9943, NSCE = 0.9943, KGE = 0.9948) phases. A high predictive functioning of RF makes it an efficient complementary tool for the development of early warning systems for vulnerable communities in the Lower Ogun Basin.

**Keywords** Early warning systems · Flood risk management · Random forest · Computational modeling

✉ O. O. Aiyelokun
   aiyelokuntobi@gmail.com

1  Department of Civil Engineering, University of Ibadan, Ibadan, Nigeria

2  Olivearc Solutions, Lagos, Nigeria

# 1 Introduction

Floods have been deemed to have the highest potential for destruction compared to other natural catastrophes (Agbde and Aiyelokun 2016). Climate change has significantly expanded the scope of economic losses and the number of individuals impacted by flooding worldwide. A widespread absence of early warning systems and pre-flood information has contributed to flood-induced losses of both life and property (Turner et al. 2014; Agbede et al. 2019). In recent years, the downstream communities of the Ogun River Basin, such as Ibafo, Magoro, Opic, Isheri, and some northern parts of Lagos State, have been ravaged by various flooding events, leading to the loss of life and property. These downstream communities are affected by the increasing severity of annual floods. For instance, major parts of Isheri, Magoro, Opic, Kara, and northern Lagos were inundated by floods between September and October of 2019. Similar occurrences of varying magnitudes were experienced in the following years, up until 2022. It is, therefore, apparent to develop early warning systems to support these communities' preparedness plans and actions before flooding occurs.

Flooding is the temporary state of partial or total deluge of normally dry places by extra water or the unexpected and quick buildup of runoff. Flood occurrences and consequences in recent years have undoubtedly been extraordinary, impacting the lives of hundreds of millions of people worldwide (Nkwunonwo et al. 2016). Worldwide, more people live in cities than in rural areas; it is estimated that about 30% of the world's population lived in cities in 1950, 54% of the world's population lived in cities in 2014, and 66% of the world's population will live in cities by 2050 (United Nations 2014), implying that the urban population is growing. The majority of the once rural areas in the Lower Ogun Basin are rapidly becoming urbanized due to their proximity to Lagos State. As a result of haphazard development in the area, many estates and residential units are being built very close to flood plains and waterways, which increases flood risks. This study seeks to characterize and predict flood levels at Mokoloki weir station, which is about 20 km upstream of the major urban settlements in the study area.

Flood estimations are subject to a variety of sources of uncertainty, such as those induced by climate change, which might have significant effects on the cost and design of hydraulic structures in downstream areas (Aiyelokun et al. 2021a). It is therefore important to develop robust models that can deal with uncertainties and outliers when generalizing a complex phenomenon such as flooding. Artificial Intelligence (AI) and Machine Learning (ML) models are now utilized in flood predictions and early warning systems due to their robustness and their remarkable ability to fit and reproduce complex processes (responses) from many inputs (Mosavi et al. 2018; Yonaba et al. 2021). One of such models is the random forest model (RF). Numerous benefits of RF include its great tolerance for outliers and noise, its difficulties in producing over-fitting phenomena, its capacity to get over the black box concept's restrictions, and its advantages in assessing key qualities (Wu et al. 2020).

Various efforts have been made to incorporate ML algorithms in flood early warning systems and related studies globally. For instance, Felix and Sasipraba (2019) utilized Gradient Boost ML to develop a flood warning system. Their system uses the amount of rainfall and the flood level of nearby water bodies based on satellites and ground applications to predict floods for timely decision-making. Ding et al. (2019) proposed the Spatio–Temporal Attention Long Short-Term Memory (STA-LSTM) for flood forecasting and confirmed that it outperformed the support vector machine (SVM), fully connected network

(FCN), and the original LSTM. Wu et al. (2019) developed a novel SMOTEBoost algorithm to perform flood forecasting in the Changhua River, China. Other works have applied Gaussian Naïve Bayes, hybrid Deep Learning (DL), Fuzzy Logic (FL), Artificial Neural Networks (ANN), Linear Regression, and Convolutional Neural Network (CNN) across the world and in Nigeria (LeCun et al. 2015; Albawi et al. 2017; Karyotis et al. 2019; Aiyelokun et al. 2021b; Rani et al. 2020; Li et al. 2020; Atashi et al. 2022; and Adetunji et al. 2023).

With respect to the applications of RF in flood prediction, Li et al. (2019) posited that RF was able to capture a more realistic characteristic of streamflow and show higher capabilities for streamflow reconstruction in comparison with bagged regression trees (BRT), support vector machines (SVM), and simple linear regression (SLM). In their 2020 study, Li et al. further established the robustness of RF over elastic net regression (ENR) and support vector regression (SVR) in streamflow forecasting for the Three Gorges Reservoir in the Yangtze River Basin, China. Khosravi et al. (2021) evaluated the efficiency of three standalone data-mining algorithms, including RF, M5 Prime (M5P), M5 Rule (M5R), and six hybrid algorithms for daily streamflow prediction. The study revealed that although all the selected models had satisfactory results, the BA-M5P was more efficient in streamflow prediction. More recently, Talukdar et al. (2022) utilized RF for predicting streamflow in the Punarbhaba River basin, Indo–Bangladesh. The contribution of the present study is to expand the existing knowledge on the utilization of RF in flood risk reduction, particularly in the area of flood level predictions at a weir station.

Flood level and discharge simulation are some of the many fields of water resources study where the RF model has been used. Random forest might serve as an alternate strategy to physical and conceptual hydrological models for large-scale hazard assessment in several catchments because of its inexpensive setup and operating costs (Schoppa et al. 2020). The gap in knowledge in the applications of RF for flood simulation in Southwest Nigeria and the strong need to explore simulation tools that can represent complex systems such as water resources systems in a realistic way (Agbede et al. 2019) motivated this study. The specific objectives of the study are to characterize flood levels at Mokoloki weir station, which is located upstream of the urban settlements in Lower Ogun Basin, and to evaluate the performance of RF in the prediction of flood levels, which could serve as a complementary tool in the development of flood early warning systems.

## 2 Description of the study area

The Lower Ogun Basin is the drainage area of the Ogun River Basin, starting from the confluence between the Oyan and Ogun rivers to the south of Abeokuta, as shown in Fig. 1. Geographically, the study area is situated between latitudes 6° 31′ N and 6° 39′ N and longitudes 3° 22′ E to 3° 39′ E, with an approximate catchment area of 361.02 km$^2$ (Odunuga and Raji 2014). The area has a low-lying topography with some evidence of undulating terrain ranging from 0 to 221 m above sea level and two air masses—the tropical maritime air mass and the tropical continental air mass—that influence its climate.
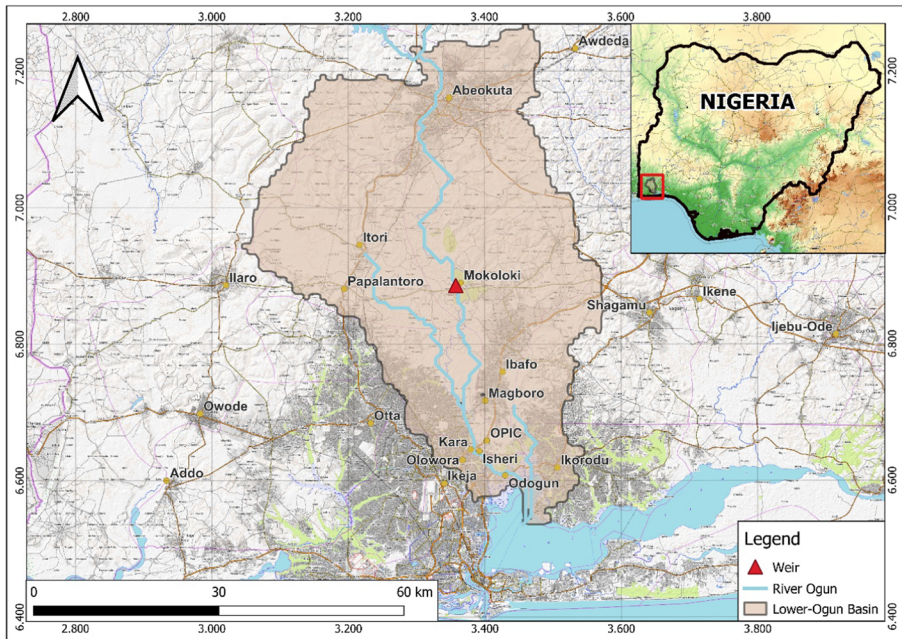
**Fig. 1** Location map of the study area

## 3 Materials and methods

### 3.1 Data acquisition

The daily time series of flood level data recorded at the Mokoloki weir, which spans a period of 36 years (1981–2016), was obtained from the Ogun–Oshun River Basin Development Authority (OORBDA).

### 3.2 Statistical analysis

Descriptive statistical methods were used to derive concise information about the historical flood levels in the study area. The descriptive statistical methods utilized include mean, median, maximum, minimum, 1st quartile, 3rd quartile, interquartile range, standard deviation, coefficient of variation, skewness, and kurtosis. A violin plot was used to represent the distribution of the flood levels as well as to assess outliers in the records. While trend and autocorrelation assessments were conducted using the Mann–Kendall and Ljung–Box tests, respectively, at a 0.05 significant level. Both trend and autocorrelation assessments are important in hydrologic studies and model development. The trend assessment provides an understanding of quantitative long-term changes in the patterns of the flood levels, while autocorrelation gives an impression of temporal dependencies or patterns within the time series of the flood level. The mathematical description of the Mann–Kendall test can be found in Blain (2013), while the description of Ljung–Box test can be found in Ljung and Box (1978).

## 3.3 Random forest (RF)

The Classification and Regression Trees (CART) model is the foundation of RF. The bootstrap resampling approach was used to create a fresh set of training samples from the repeated random K samples that are extracted from a single training sample set N. K decision trees were then constructed, and a random forest was included in the bootstrap sample collection.

With respect to a classification problem, the number of votes a classification tree receives determines how accurate the data is classified, and for a regression problem, all averages of the predictive value of decision trees are taken into consideration as final prediction results (Fig. 2). The works of Boulesteix et al. (2012), Scornet (2017), and Wu et al. (2020) may be consulted for a better understanding of the variation, parameters, and feature-important capabilities of RF and its working process.

Since the flood levels to be predicted are continuous values, a supervised regression framework was adopted in this study. The input combination of flood levels of up to 7 days was considered suitable because antecedent runoff or rainfall of up to 4 days (lag $(t-4)$) has been reported to be effective for runoff prediction (Sharifi et al. 2017).

This relationship can be represented as shown below:

$$\mathrm{FL}_{(t)} = f\left\{\mathrm{FL}_{(t-1)} + \mathrm{FL}_{(t-2)} + \dots \mathrm{FL}_{(t-7)}\right\} \tag{1}$$

where FL is the flood level, and $t$ is the day.

The "random forest" package from the $R$ statistical package, available at http://www.R-project.org in version 3.6, was used to create the model. 70% of the dataset was used to calibrate the random forest, and the remaining 30% was split into 15% validation and test sets, respectively. There is widespread speculation that key parameters for calibrating RF models include "ntree," which stands for the number of trees in the forest, and "mtry," which stands for the number of separate descriptors confirmed based on corresponding partitions (Li et al. 2016; Scornet 2017; Rakhee et al. 2020; Aiyelokun and Agebde 2021). It was determined that a ntree of 320 and a mtry of 4 were sufficient for developing the model employed in this research.

## 3.4 Evaluation of model performance

The performance of the model was assessed using Mean Absolute Error (MAE), coefficient of determination ($R^2$), Nash–Sutcliffe efficiency coefficient (NSEC) and Kling-Gupta efficiency (KGE). A MAE value close to 0, and $R^2$, NSCE and KGE values close to 1, are evidence of a good model. Performance evolution methods such as MAE, $R^2$
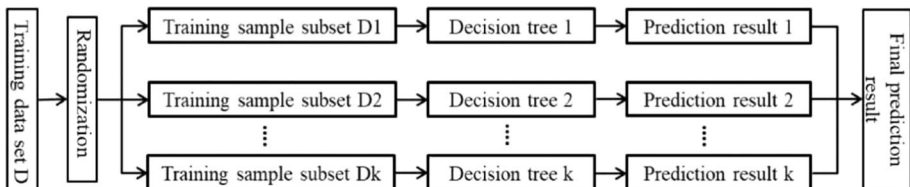


**Fig. 2** Workflow diagram of a Random Forest. (adapted from Wu et al. 2020)

and NSEC were selected because of their wildspread applications in similar studies, while KGE was selected for its robustness in the evaluation of model performance and ability to combine multiple evaluation methods (Yonaba et al. 2020). The mathematical descriptions for MEA, $R^2$, NSCE, and KGE are listed as follows:

$$\text{MAE} = \frac{\sum\limits_{i=1}^{n} |y_i - \widetilde{y}_i|}{n} \tag{2}$$

$$R^2 = \sum_{i=1}^{n} \left[ \left( y_i - \bar{y}_i \right)\left( \widetilde{y}_i - \bar{\widetilde{y}}_i \right) \right]^2 \bigg/ \left[ \sum_{i=1}^{n} \left( y_i - \bar{y} \right) \sum_{i=1}^{n} \left( \widetilde{y}_i - \bar{\widetilde{y}}_i \right) \right]^2 \tag{3}$$

$$\text{NSCE} = 1 - \sum_{i=1}^{n} (y_i - \widetilde{y}_i)^2 \bigg/ \sum_{i=1}^{n} (y_i - \bar{y})^2 \tag{4}$$

$$\text{KGE} = 1 - \sqrt{(R-1)^2 + (\beta - 1)^2 + (\gamma - 1)^2} \tag{5}$$

where for MAE, $R^2$, NSCE $y_i$ is the observed flood level, and $\widetilde{y}_i$ is the predicted flood level, while $\bar{y}$ $and$ $\bar{\widetilde{y}}_i$ indicate the average observed and predicted flood level, respectively. For KGE, $R$ is Person's product-moment correlation coefficient between $y_i$ and $\widetilde{y}_i$, $\beta = \bar{y}_i / \bar{\widetilde{y}}_i$, and $\gamma = CV_s / CV_0$ the ratio of coefficients of variation of predicted and observed flood levels. The terms 'training' 'validation' and 'testing' phases used for the data intelligent model also means calibration, validation and testing of physically based on hydrodynamic model (Li et al. 2016).

### 3.5 Important feature assessment

The ability of a random forest to handle data with outliers and noise makes it effective as a tool for feature evaluation (Jaiswal and Samikannu 2017). The ability of RF's significant feature section places it in the lead when compared to the many algorithms used to provide insights during the building of data-driven models. The research incorporated two crucial RF's feature selection techniques: minimum depth (Ishwaran et al. 2010; Zhang et al. 2018) and percentage increase in mean square error (%IncMSE).

### 3.6 Evaluation of RF uncertainty

The RF model's uncertainty was propagated using the Monte Carlo approach. This is due to the fact that the Monte Carlo technique has been shown to be effective for assessing the level of uncertainty in complicated models like RF and artificial neural networks (Khosravi et al. 2011). The Coulston et al. (2016) Monte Carlo method was modified in this work to approximate prediction uncertainty for random forest regression models. Although prediction intervals were either too large or too tight in sparse areas of the prediction distribution, this technique offered appropriate estimates of prediction uncertainty. The research used

five major phases, which are similar to Coulston et al.'s technique for measuring uncertainty in random forest regression. These phases include:

1. Fitting a random forest model on the training data;
2. Using a bootstrap resampling (sampling with replacement) to generate multiple bootstrap samples from the test data;
3. Making predictions on each bootstrap test sample based on the trained RF model (200 iterations were performed),
4. Aggregating the predictions from all bootstrap samples to obtain a distribution of the predictions, and
5. Calculating statistics such as mean, minimum, maximum and percentiles from the distribution to estimate the uncertainty.
6. Plotting the uncertainty of RF using the error plot.

## 4 Results

### 4.1 Statistical description of flood level in Lower Ogun Basin

The characterization of historical hydrologic data using statistical tools is an important step in water resources and flood management. Table 1 shows the descriptive statistics of flood levels at Mokoloki weir station for the period of 36 years (1981–2016) based on daily, seasonal and annual time scales. It can be observed from the table that, unlike the annual time scale, the flood level record is slightly negatively skewed for daily and seasonal time scales, while the mean values of 5.83 m and 5.82 m were slightly greater than the median values of 5.64 m and 5.63 m for daily and seasonal time series, respectively. This implies that flood level data at Mokoloki weir violate the textbook rule that the mean is usually right of the median under a positive skew (Von Hippel 2005) for all time scales except for annual. The low values of kurtosis of 0.16 and 0.12 for the daily and seasonal time series, respectively, are indications that the tail of the flood level distribution does not extend farther than that of a normal distribution, while a negative kurtosis of $-0.35$ is an indication that the annual series has lighter tails and is less peaked compared to the daily and seasonal

**Table 1** Statistics summary of flood levels at the Mokoloki weir

| Statistic | Daily | Seasonal | Annual |
|---|---|---|---|
| Minimum (m) | 0.2 | 0.2 | 4.71 |
| 1st Quartile (m) | 4.6 | 4.62 | 5.36 |
| Median (m) | 5.64 | 5.63 | 5.88 |
| Mean (m) | 5.829 | 5.821 | 5.83 |
| 3rd Quartile (m) | 7.04 | 7.05 | 6.14 |
| Maximum (m) | 9.5 | 9.06 | 7.21 |
| Interquartile range (m) | 2.44 | 2.43 | 0.77 |
| Standard deviation | 1.62 | 1.58 | 0.60 |
| Coefficient of variation | 0.28 | 0.27 | 0.10 |
| Skewness | $-0.067$ | $-0.092$ | 0.399 |
| Kurtosis | 0.1615 | 0.1161 | $-0.350$ |
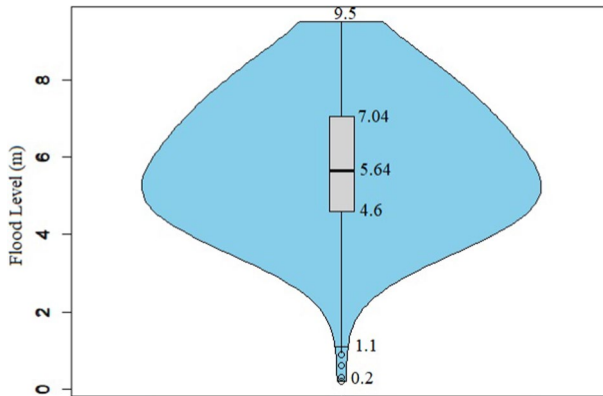
**Fig. 3** Combined violin plot and boxplot of daily flood levels. The plot provides a visual summary of the distribution of the flood level data, which includes maximum or upper whisker (9.5 m), minimum (0.2 m), interquartile range (thick gray bar), median (5.64 m), 3rd quartile (7.04 m), 1st quartile (4.6 m), lower whisker (9.5 m), and outliers represented by circles (0.2 m being the lowest outliers). The blue portion represents the kernel density estimations; the wilder blue portion represents flood levels with a higher probability of occurrence

series. The standard deviation values of 1.62, 1.58 and 0.62 and the coefficient of variation of 0.28, 0.27 and 0.10 are indications that the flood levels do not disperse largely from the mean of at least 5.82 m for all time series. The table further shows that 25% of the flood level was below 4.6 m (1st Quartile), while 75% of the flood level was below 7.04 m (3rd Quartile), implying that the majority of the flood level ranged between 4.6 and 7.04 m, with an interquartile range of 2.44 m for the daily time series.

The distribution of the flood level data is further presented in the form of a violin plot (Fig. 3). Other information that could be derived from the violin plot is that 25% of flood levels were above 7.04 m, which could have been responsible for the majority of the flooding events in downstream residential areas located in Ibafo, Magboro and Isheri North in Lagos State. The box plot further reveals that outliers in the flood level data were below 1.1 m. The outliers were retained for the modeling experimentation because RF is not sensitive to outliers during model calibrations.

## 4.2 Trend and autocorrelation of flood level

The trend analysis and the autocorrelation test for daily, seasonal, and annual time scales are presented in Table 2. The Mann–Kendall tau values of 0.19, 0.194, and 0.47 were obtained for the daily, seasonal, and annual time scales, respectively, which reveal positive

**Table 2** Summary of trends and autocorrelation of flood levels at the Mokoloki Weir

| Time Scale | Mann Kendall test | Ljung–Box test |
| --- | --- | --- |
| Daily | tau = 0.19 ($p < 0.001$) | $X$-squared = 13,059 ($p < 0.001$, df = 1) |
| Seasonal | tau = 0.194 ($p < 0.001$) | $X$-squared = 283.42 ($p < 0.001$, df = 1) |
| Annual | tau = 0.47 ($p < 0.001$) | $X$-squared = 15.821 ($p < 0.001$, df = 1) |
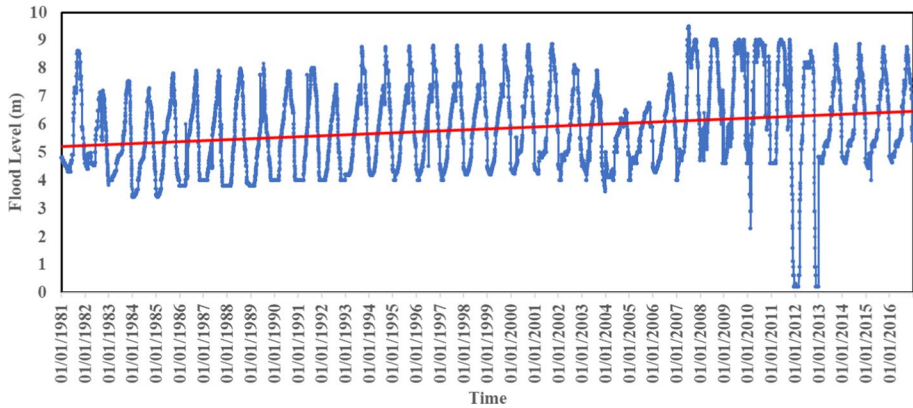
**Fig. 4** Trend of Variation of Flood Levels

trends in flood level time series based on the different time scales. The $p$ value of the trend tests was $< 0.001$, implying that flood levels are significantly increasing in the study area. While the Ljung–Box test's $X$-squared values obtained for the daily, seasonal, and annual time scales are greater than the tabulated Chi-Square of 3.841 at df of 1, the $p$ value is less than 0.05, implying that autocorrelation exists in the time series. Generally, a $p$ value of less than 0.05 is an indication of a significant trend and autocorrelation in flood level records. Figure 4 further depicts the increasing positive trend and presence of outliers in the series.

Figure 5 shows that there is a significantly higher autocorrelation of flood levels up to at least the 30th lag for different periods of investigation ((a) 2007–2016, (b) 2002–2016, (c) 1997–2016, (d) 1992–2016, (e) 1987–2016, (f) 1981–2016), which is an indication that the autocorrelation of the flood level is not dependent on the period of investigation.

From Fig. 5, ACF dropped below 0.97 for the observed flood levels after the seventh day (Lag 7). As a result, antecedent flood levels of one to 7 days were considered appropriate and adopted as the predictors, while the present day's flood level served as the response variable.

## 4.3 Performance of RF in flood level prediction

Table 3 shows that RF was very efficient in generalizing flood levels. Since it was able to achieve a very low prediction error and more than 99% prediction accuracy based on $R^2$, NSCE, and KGE for the three data sets.

Furthermore, using a mix of scatter plots and time series plots (Fig. 6), the degree of agreement between the observed and anticipated flood level values was examined. Figure 6 shows that during the training, validation, and testing phases, RF was not only able to achieve a high level of closeness between the observed and anticipated flood level values but also performed well in predicting the outliers.
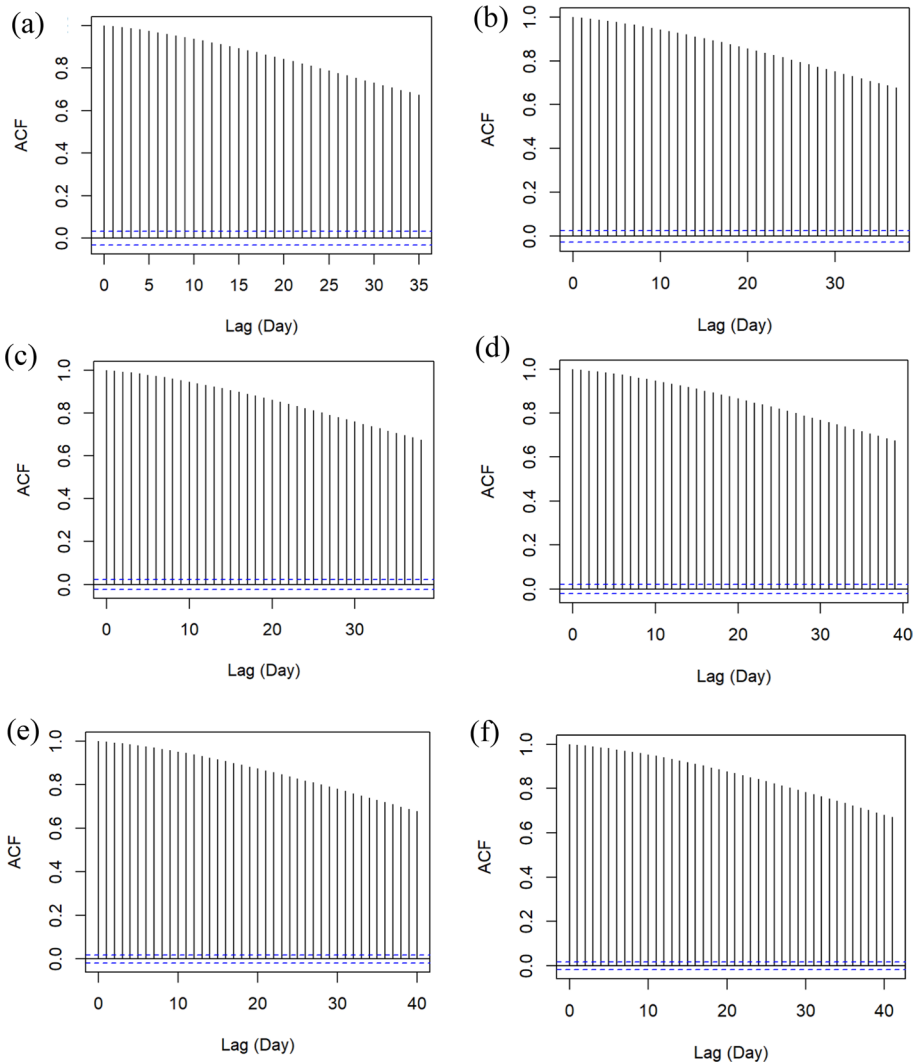
**Fig. 5** Autocorrelation plot of flood levels at Mokoloki weir for (**a**) 2007–2016, (**b**) 2002–2016, **c**) 1997–2016, (**d**) 1992–2016, (**e**) 1987–2016, and (**f**) 1981–2016: AFC represents the strength of the correlation between flood levels on a particular day and their lagged days. The dashed blue lines on the plot represent the 95% confidence interval. Autocorrelation values outside the blue lines are considered statistically significant

| Table 3 Performance of RF in predicting flood level | Training | Validation | Testing |
|---|---|---|---|
| MAE | 0.0292 | 0.0484 | 0.0519 |
| $R^2$ | 0.9966 | 0.9943 | 0.9924 |
| NSCE | 0.996 | 0.9943 | 0.9924 |
| KGE | 0.9950 | 0.9948 | 0.9930 |
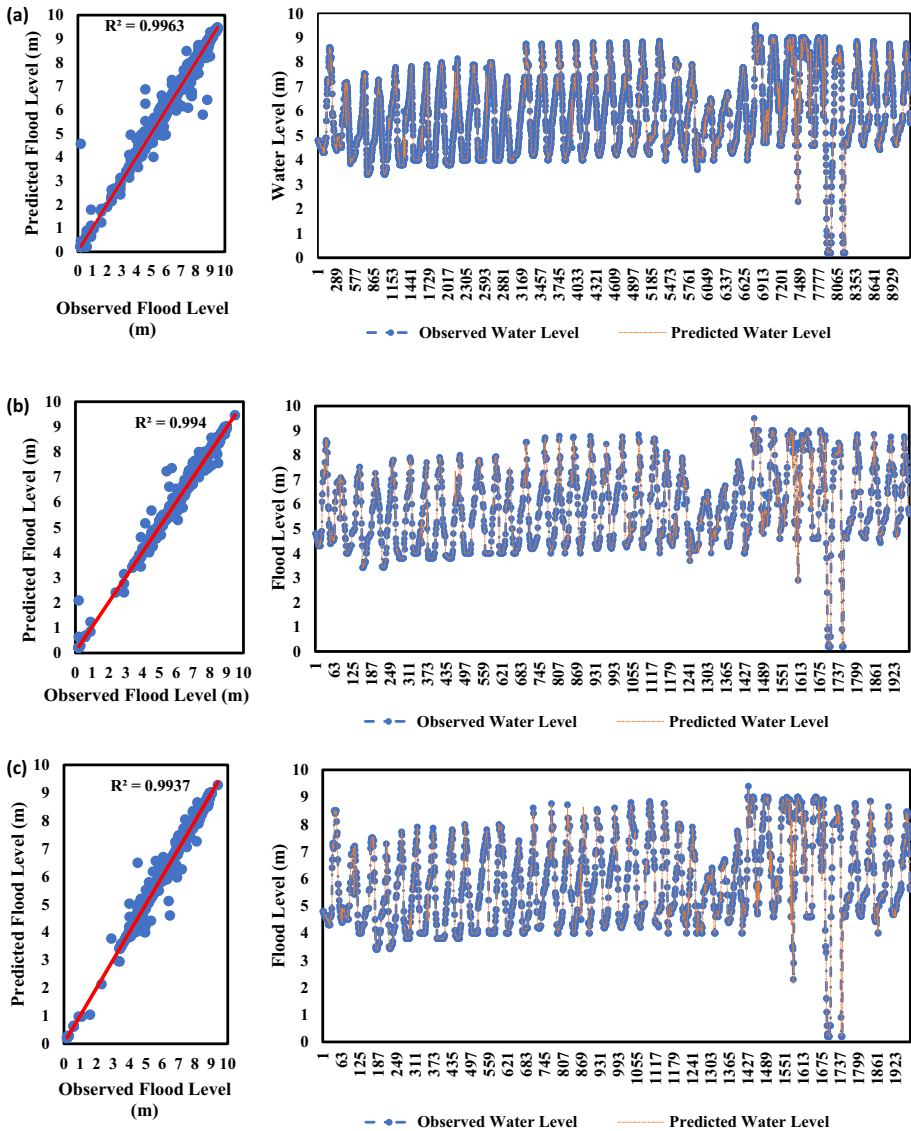
**Fig. 6** Scatter and time series plot between the predicted and observed values of flood levels for (**a**) training (**b**) validation (**c**) testing

## 4.4 Importance of input features for predicting flood level

The important input features for predicting flood levels were assessed using the minimum depth plot (Fig. 7) and the %IncMSE with their significance at 95% confidence level (Table 4). Based on Fig. 7, $FL_{(t-1)}$, $FL_{(t-2)}$ and $FL_{(t-3)}$ have the lowest mean minimum depth of 1.16, 1.42 and 1.55, respectively, with less than 100 trees. The $FL_{(t-4)}$ and
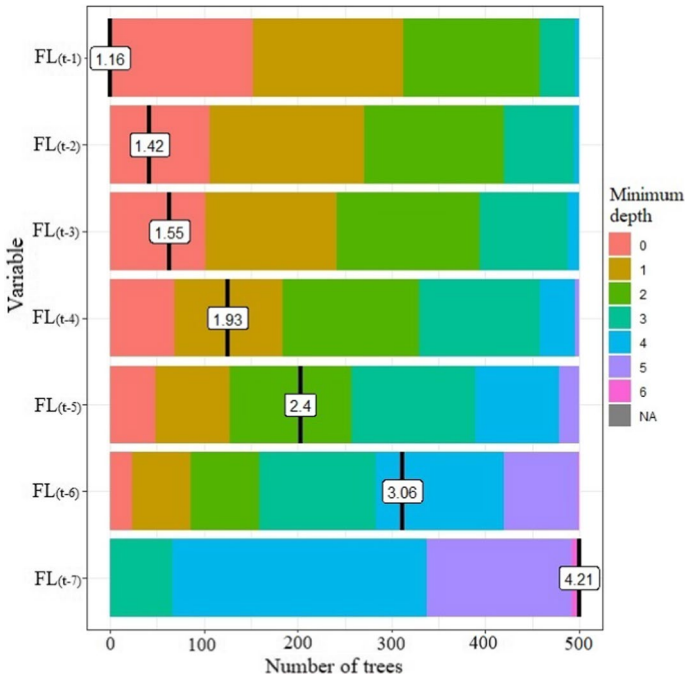
**Fig. 7** Distribution of minimum depth and mean of daily antecedent flood levels

**Table 4** Statistical evaluation of important climate parameters in predicting flood levels

| Features | %IncMSE | $p$ value |
|----------|---------|-----------|
| FL$(t-1)$ | 0.03 | 0.01 |
| FL$(t-2)$ | 0.94 | 0.01 |
| FL$(t-3)$ | 0.99 | 1.00 |
| FL$(t-4)$ | 0.97 | 1.00 |
| FL$(t-5)$ | 0.95 | 1.00 |
| FL$(t-6)$ | 0.99 | 1.00 |
| FL$(t-7)$ | 0.98 | 1.00 |

FL$_{(t-5)}$ have a mean minimum depth of tress of 1.93 and 2.4, respectively, while FL$_{(t-6)}$ and FL$_{(t-7)}$ had the highest mean minimum depth of trees of 3.06 and 4.21, respectively. This is further established in Table 4, which shows that the selected input features were important for the prediction of flood levels in the order of FL$_{(t-1)}$ > FL$_{(t-2)}$ > FL$_{(t-3)}$ FL$_{(t-4)}$ > FL$_{(t-5)}$ > FL$_{(t-6)}$ > FL$_{(t-7)}$ based on their IncMSE%. Implying that FL$_{(t-1)}$ and FL$_{(t-2)}$ are the most important independent variables for predicting the daily flood levels in the study area. Since the $p$ value of FL$_{(t-1)}$ and FL$_{(t-2)}$ are lesser than 0.05, then, it can be concluded that antecedent flood levels of up to 2 days are significant input variables for predicting daily flood levels (Table 4).
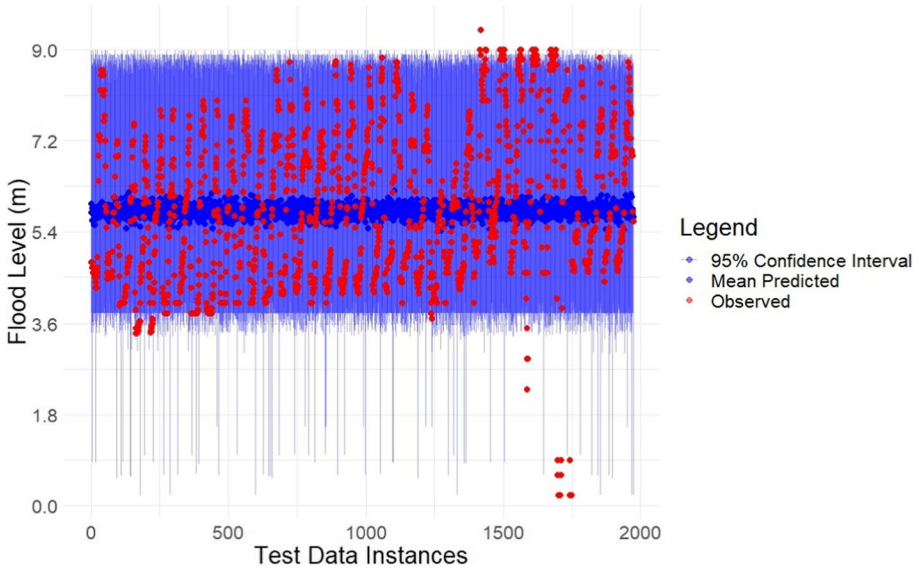
**Fig. 8** Error bar plot of uncertainty associated with model prediction

**Fig. 9** Line plot of daily flood level prediction with uncertainty indications based on Monte Carlo simulations
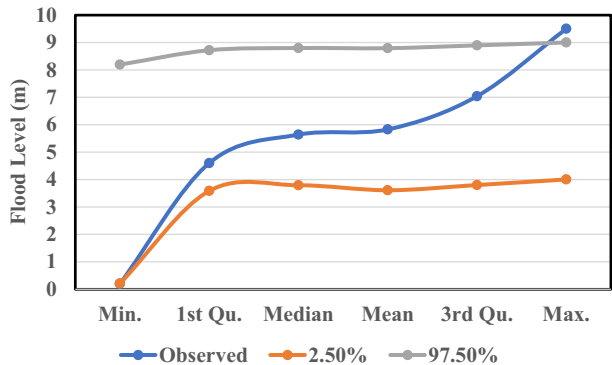


## 4.5 Uncertainty of RF model

Figure 8 shows the uncertainty around the mean prediction of the test dataset by RF based on the Monte Carlo simulation. It could be observed that there are cases of wide and narrow error bars, which, respectively, represent uncertain and more confident predictions. The figure further shows that the RF model is uncertain with the prediction of flood levels higher than 9.0 m and those that are lower than 3.15 m at a 95% confidence level, since at these records, most of the flood levels were outside the lower and upper bounds. Figure 9 further emphasized that uncertainty in flood level prediction increased for lower flood levels and higher flood levels greater than 9 m, since the blue line crossed the gray line (97.50% confidence level) at 9.0 m.

## 4.6 Discussion of findings

Researchers have recently campaigned for the implementation of integrated flood risk management strategies, which combine the wide deployment of structural and non-structural interventions (Pitt 2008; Kazmierczak and Carter 2010; Sayers et al. 2013; Nkwunonwo et al. 2016). The structural measures work to reduce flood hazards by ensuring that flood water flow is regulated in the urban built-up region. They employ technically sound processes to reduce the likelihood of flood threats and entail channelization as well as the use of both natural and artificial structures to hold back water in rivers and oceans. On the other hand, non-structural interventions use interdisciplinary techniques to reduce flood risks and increase the ability of environmental systems to withstand floods (Sayers et al. 2013). Examples of non-structural measures include low-impact development (LID), land use zoning and planning, flood risk mapping, flood proofing, flood modeling, institutionalization of policies, flood awareness campaigns, resettlement of the human population, flood insurance, flood vulnerability assessment, flood forecasting, relocation of properties, and green infrastructure planning (Jha et al. 2012; Smith 2013; Nkwunonwo et al. 2016). The implementation of the ML technique is also an example of non-structural interventions for flood risk management.

This study assessed the performance of RF in modeling and predicting flood levels in the Lower Ogun Basin, Southwest Nigeria. The statistical description of the flood levels revealed that the annual series of the flood level has a different characteristic from the daily and seasonal series. For instance, a negative kurtosis of $-0.35$ for the annual series indicated that the annual series had lighter tails and was less peaked, implying that compared to the daily and seasonal series, the annual series significantly disregards the extreme values of the flood level, which are important for the development of efficient early warning systems. The scope of the study is limited to the prediction of daily flood levels, even though a significant ($p < 0.001$) trend and autocorrelation were found on other time scales.

This study further investigates the performance of the RF model as a complementary tool in flood early warning systems for the Lower Ogun Basin. Based on three evaluation criteria, which include MAE, $R^2$, NSCE, and KGE, the performance of RF in predicting flood levels is satisfactory and can be used for developing flood early warning systems. For instance, in a similar study by Albawi et al. (2017), lesser accuracy was recorded for the linear regression algorithms SVM and ANN, with MAEs of 40.25%, 90.61%, and 21.81%, respectively, compared to the MAEs of 3% to 5% obtained by RF in the present study. In addition, Kunverji et al. (2021) recently established that the Decision Tree Algorithm, with an accuracy of 94.4%, outperformed the Gradient Boost Algorithm with an accuracy of 87.9% and the RF with an accuracy of 92.4% for the development of flood early warning systems. This implies the RF model is adequate for serving as a complementary tool in flood early warning systems, which could be adopted as a non-structural measure for flood management in the Lower Ogun Basin.

Furthermore, the assessment of important input features for the prediction of flood levels revealed that $FL_{(t-1)}$ and $FL_{(t-2)}$ are the most important input features for daily flood level prediction, while the model uncertainty assessment based on Monte Carlo simulation emphasized that uncertainty in flood level prediction by RF increased for lower flood levels and higher flood levels greater than 9.0 m, since the blue line crossed the gray line (97.50% confidence level) at 9.0 m (Fig. 9). Implying that the developed RF model is uncertain at 95% confidence interval in predicting extreme floods of greater than 9.0 and lesser than 3.15 m.

## 5 Conclusion

This study has expounded on the efficiency and usefulness of the RF model as a predictive tool for the development of flood early warning systems in the Lower Ogun Basin. The study established that the majority of the flood level ranged between 4.6 and 7.04 m, with an interquartile range of 2.44 m. Flood levels have a significant increasing trend, and 25% of flood levels were above 7.04 m, which could have been responsible for the majority of the flooding events in downstream residential areas located in Ibafo, Magboro, and Isheri North in Lagos State. A RF model was developed to model and predict flood levels based on antecedent records. Based on the evaluation criteria, RF is reliable in predicting daily flood levels, having performed well at both the validation (MAE=0.0484, $R^2$=0.9924, NSCE=0.9924, KGE=0.9930) and test (MAE=0.0519, $R^2$=0.9943, NSCE=0.99943, KGE=0.99948) phases. Furthermore, antecedent flood levels of up to 2 days are the most important input features for the RF model, while based on Monte Carlo simulation, the RF model is limited in that it is less certain in the prediction of extreme low and high flood levels. The present study is limited to the use of 7-day antecedent flood levels as input features of the RF model; however, future studies should focus on the application of different input combinations. The high predictive functioning of RF makes it an efficient complementary tool for the development of early warning systems for vulnerable communities in the Lower Ogun Basin.

## Declarations

## References

Adetunji OJ, Adeyanju IA, Esan AO (2023) Flood Areas prediction in Nigeria using artificial neural network. In: 2023 International conference on science, engineering and business for sustainable development goals (SEB-SDG), vol 1, pp 1–6. https://doi.org/10.1109/SEB-SDG57117.2023.10124629

Agbede OA, Aiyelokun OO (2016) Establishment of a stochastic model for sustainable economic flood management in Yewa Sub-Basin, southwest Nigeria. Civ Eng J 2(12):646–655. https://doi.org/10.28991/cej-2016-00000065

Agbede OA, Aiyelokun O, Ojelabi A, Oyelakin J (2019) Influence of low impact development on peak floods using system dynamics. UI J Civ Eng Technol 1(1):50–62

Aiyelokun OO, Agbede OA (2021) Development of random forest model as decision support tool in water resources management of Ogun headwater catchments. Appl Water Sci 11(7):1–9. https://doi.org/10.1007/s13201-021-01461-x

Aiyelokun O, Pham QB, Aiyelokun O, Malik A, Adarsh S, Mohammadi B et al (2021a) Credibility of design rainfall estimates for drainage infrastructures: extent of disregard in Nigeria and proposed framework for practice. Nat Hazards 109(2):1557–1588. https://doi.org/10.1007/s11069-021-04889-1

Aiyelokun O, Ogunsanwo G, Ojelabi A, Agbede O (2021b) Gaussian Naïve Bayes classification algorithm for drought and flood risk reduction. In: Deo R, Samui P, Kisi O, Yaseen Z (eds) Intelligent data analytics for decision-support systems in hazard mitigation. Springer Transactions in Civil and Environmental Engineering. Springer, Singapore. https://doi.org/10.1007/978-981-15-5772-9_3

Albawi S, Mohammed TA, Al-Zawi S (2017) Understanding of a convolutional neural network. In: 2017 international conference on engineering and technology (ICET), Antalya, Turkey, pp 1–6. IEEE. https://doi.org/10.1109/ICEngTechnol.2017.8308186

Atashi V, Gorji HT, Shahabi SM, Kardan R, Lim YH (2022) Water level forecasting using deep learning time-series analysis: A case study of red river of the north. Water 14(12):1–18. https://doi.org/10.3390/w14121971

Blain GC (2013) The Mann-Kendall test: the need to consider the interaction between serial correlation and trend. Acta Sci Agron 35:393–402. https://doi.org/10.4025/actasciagron.v35i4.16006

Boulesteix A, Janitza S, Kruppa J, König I (2012) Overview of random forest methodology and practical guidance with emphasis on computational biology and bioinformatics. Wiley Interdiscip Rev Data Min Knowl Discov 2(6):493–507. https://doi.org/10.1002/widm.1072

Coulston JW, Blinn CE, Thomas VA, Wynne RH (2016) Approximating prediction uncertainty for random forest regression models. Photogramm Eng Remote Sens 82(3):189–197. https://doi.org/10.14358/PERS.82.3.189

Ding Y, Zhu Y, Wu Y, Jun F, Cheng Z (2019) Spatio-temporal attention LSTM model for flood forecasting. In: 2019 International conference on Internet of Things (IThings) and IEEE green computing and communications (GreenCom) and IEEE cyber, physical and social computing (CPSCom) and IEEE smart data (SmartData). IEEE. pp 458–465. https://doi.org/10.1109/iThings/GreenCom/CPSCom/SmartData.2019.00095

Felix AY, Sasipraba T (2019) Flood detection using gradient boost machine learning approach. In: 2019 International conference on computational intelligence and knowledge economy (ICCIKE) pp 779–783. IEEE. Doi:https://doi.org/10.1109/ICCIKE47802.2019.9004419

Ishwaran H, Kogalur UB, Gorodeski EZ, Minn AJ, Lauer MS (2010) High-dimensional variable selection for survival data. J Amer Stat Assoc 105(489):205–217. https://doi.org/10.1198/jasa.2009.tm08622

Jaiswal JK, Samikannu R (2017) Application of random forest algorithm on feature subset selection and classification and regression. In: 2017 world congress on computing and communication technologies (WCCCT), Tiruchirappalli, India, pp 65–68. https://doi.org/10.1109/WCCCT.2016

Jha AK, Bloch R, Lamond J (2012) Cities and flooding: a guide to integrated urban flood risk management for the 21st century. The World Bank. GFDRR. http://hdl.handle.net/10986/2241

Karyotis C, Maniak T, Doctor F, Iqbal R, Palade V, Tang R (2019) Deep learning for flood forecasting and monitoring in urban environments. In: 2019 18th IEEE international conference on machine learning and applications (ICMLA) pp 1392–1397. IEEE. https://doi.org/10.1109/ICMLA.2019.00227

Kazmierczak A, Carter J (2010) Adaptation to climate change using green and blue infrastructure. A database of case studies. University of Manchester, School of Environment, Education, and Development, Manchester

Khosravi A, Nahavandi S, Creighton D, Atiya AF (2011) Comprehensive review of neural network-based prediction intervals and new advances. IEEE Trans Neural Netw 22(9):1341–1356. https://doi.org/10.1109/TNN.2011.2162110

Khosravi K, Golkarian A, Booij MJ, Barzegar R, Sun W, Yaseen ZM, Mosavi A (2021) Improving daily stochastic streamflow prediction: comparison of novel hybrid data-mining algorithms. Hydrol Sci J 66(9):1457–1474. https://doi.org/10.1080/02626667.2021.1928673

Kunverji K, Shah K, Shah, N (2021) A flood prediction system developed using various machine learning algorithms (May 7, 2021). In: Proceedings of the 4th international conference on advances in science & technology (ICAST2021). https://doi.org/10.2139/ssrn.3866524

LeCun Y, Bengio Y, Hinton G (2015) Deep Learning. Nature 521(7553):436–444. https://doi.org/10.1038/nature14539

Li B, Yang G, Wan R, Dai X, Zhang Y (2016) Comparison of random forests and other statistical methods for the prediction of lake water level: a case study of the Poyang Lake in China. Hydrol Res 47(S1):69–83. https://doi.org/10.2166/nh.2016.264

Li J, Wang Z, Lai C, Zhang Z (2019) Tree-ring-width based streamflow reconstruction based on the random forest algorithm for the source region of the Yangtze River, China. CATENA 183:104216. https://doi.org/10.1016/j.catena.2019.104216

Li Y, Liang Z, Hu Y, Li B, Xu B, Wang D (2020) A multi-model integration method for monthly streamflow prediction: modified stacking ensemble strategy. J Hydroinf 22(2):310–326. https://doi.org/10.2166/hydro.2019.066

Ljung GM, Box GE (1978) On a measure of lack of fit in time series models. Biometrika 65(2):297–303. https://doi.org/10.1093/biomet/65.2.297

Mosavi A, Ozturk P, Chau K (2018) Flood prediction using machine learning models: literature review. Water 10:1–41. https://doi.org/10.3390/w10111536

Nkwunonwo U, Whitworth M, Baily B (2016) Review article: a review and critical analysis of the efforts towards urban flood risk management in the Lagos region of Nigeria. Nat Hazards Earth Syst Sci 16:349–369. https://doi.org/10.5194/nhess-16-349-2016

Odunuga S, Raji SA (2014) Flood frequency analysis and inundation mapping of lower Ogun River basin. J Water Resour Hydraul Eng 3(3):48–59

Pitt M (2008) Lessons from the 2007 Floods. Pitt Review, London, p 2008

Rakhee R, Singh A, Mittal M, Kumar A (2020) Qualitative analysis of random forests for evaporation prediction in Indian Regions. Indian J Agric Sci 90(6):1140–1144

Rani DS, Jayalakshmi GN, Baligar VP (2020). Low cost IoT based flood monitoring system using machine learning and neural networks: flood alerting and rainfall prediction. In: 2020 2nd international conference on innovative mechanisms for industry applications (ICIMIA) (pp 261–267). IEEE

Sayers P, Li Y, Galloway G, Penning-Rowsell E, Shen F, Wen K, Chen Y, Le Quesne T (2013) Flood risk management: a strategic approach. UNESCO, Paris, p 2013

Schoppa L, Disse M, Bachmair S (2020) Evaluating the performance of random forest for large-scale flood discharge simulation. J Hydrol 590:125531

Scornet E (2017) Tuning parameters in random forests. ESAIM Proc Surv 60:144–162. https://doi.org/10.1051/proc/201760144

Sharifi A, Dinpashoh Y, Mirabbisi R (2017) Daily runoff prediction using linear and non-linear models. Water Sci Technol 76(3–4):793–805. https://doi.org/10.2166/wst.2017.234

Smith K (2013) Environmental hazards: assessing risk and reducing disaster. Routledge, London

Talukdar S, Pal S, Fahad S, Naikoo MW, Parvez A, Rahman A (2022) Trend analysis and forecasting of streamflow using random forest in the Punarbhaba River basin. Environ Monit Assess 195(1):1–27. https://doi.org/10.1007/s10661-022-10696-3

Turner G, Said F, Afzal U, Campbell K (2014) The effect of early flood warnings on mitigation and recovery during the 2010 Pakistan floods (August 4, 2014). Turner G, Said F, Afzal U, Campbell K, Preventing disaster: early warning systems for climate change. United Nations Environmental Programme, Singh A, Zommers Z (Ed), Springer Netherlands. Available at SSRN: https://ssrn.com/abstract=2476039

United Nations (2014) World urbanization prospects: the 2014 revision, highlights (ST/ESA/SER.A/352) Department of Economic and Social Affairs, Population Division

Von Hippel PT (2005) Mean, median, and skew: correcting a textbook rule. J Stat Educ 13(2):1–13. https://doi.org/10.1080/10691898.2005.11910556

Wu Y, Yukai D, Feng J (2019) Sparse bayesian flood forecasting model based on smoteboost. In: 2019 International Conference on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData), pp 279–284. https://doi.org/10.1109/iThings/GreenCom/CPSCom/SmartData.2019.00067

Wu M, Feng Q, Wen X, Deo R, Yin Z, Yang L, Sheng D (2020) Random forest predictive model development with uncertainty analysis capability for the estimation of evapotranspiration in an arid oasis region. Hydrol Res 51(4):648–665. https://doi.org/10.2166/nh.2020.012

Yonaba R, Biaou AC, Koïta M, Tazen F, Mounirou LA, Zouré CO, Karambiri H, Yacouba H (2020) A dynamic land use/land cover input helps in picturing the Sahelian paradox: assessing variability and attribution of changes in surface runoff in a Sahelian watershed. Sci Total Environ 103:1–18. https://doi.org/10.1016/j.scitotenv.2020.1437

Yonaba R, Koïta M, Mounirou LA, Tazen F, Queloz P, Biaou AC, Yacouba H (2021) Spatial and transient modelling of land use/land cover (LULC) dynamics in a Sahelian landscape under semi-arid climate in northern Burkina Faso. Land Use Policy 103:1–18. https://doi.org/10.1016/j.landusepol.2021.105305

Zhang J, Zhu Y, Zhang X, Ye M, Yang J (2018) Developing a Long Short-Term Memory (LSTM) based model for predicting water table depth in agricultural areas. J Hydrol 561:918–929. https://doi.org/10.1016/j.jhydrol.2018.04.065