Archives of Hydro-Engineering and Environmental Mechanics 69 (2022), pp. 59–77, doi: 10.2478/heem-2022-0005

sciendo

ARCHIVES OF HYDRO-ENGINEERING
AND ENVIRONMENTAL MECHANICS

Published since 1954 ISSN 1231–3726

Artificial Intelligence Based Flood Forecasting for River Hunza at Danyor Station in Pakistan

Muhammad Waseem Yaseen[1,2], **Muhammad Awais**[1], **Khuram Riaz**[1,2*], **Muhammad Babar Rasheed**[1], **Muhammad Waqar**[2], **Sajid Rasheed**[2]

[1]The University of Lahore, 1-km off Defence Road, Lahore, Pakistan, 54000,
[2]The University of Punjab, Quaid-e-Azam campus Canal Bank Road, Lahore, Pakistan, 54000
[*] Corresponding Author: khurrammalghani@gmail.com, +92-320-5705114

(Received 06 October 2022; revised 19 December 2022)

Abstract. Floods can cause significant problems for humans and can damage the economy. Implementing a reliable flood monitoring warning system in risk areas can help to reduce the negative impacts of these natural disasters. Artificial intelligence algorithms and statistical approaches are employed by researchers to enhance flood forecasting. In this study, a dataset was created using unique features measured by sensors along the Hunza River in Pakistan over the past 31 years. The dataset was used for classification and regression problems. Two types of machine learning algorithms were tested for classification: classical algorithms (Random Forest, RF and Support Vector Classifier, SVC) and deep learning algorithms (Multi-Layer Perceptron, MLP). For the regression problem, the result of MLP and Support Vector Regression (SVR) algorithms were compared based on their mean square, root mean square and mean absolute errors. The results obtained show that the accuracy of the RF classifier is 0.99, while the accuracies of the SVC and MLP methods are 0.98; moreover, in the case of flood prediction, the SVR algorithm outperforms the MLP approach.

Key words: Hydrometeorology, Random Forest, Support Vector, Multilayer Perceptron, Machine Learning, Flood Forecasting

List of Acronyms:

AI	Artificial Intelligence
ANFIS	Adaptive Neuro-Fuzzy Inference System
ANN	Artificial Neural Network
BPNN	Back Propagation Neural Networks
ELM	Extreme Learning Machine
FMLP	Feedforward Multilayer Perceptron

HKH	Hindukush-Karakuram-Himalayan
MAE	Mean Absolute Error
MCC	Most Common Category
MLP	Multilayer Perceptron
MNLR	Multinomial Logistics Regression
MSE	Mean Square Error
RBF	Radial Basis Function
RF	Random Forest
RFE	Recursive Feature Elimination
RMSE	Root Mean Square Error
SVC	Support Vector Classifier
SVR	Support Vector Regression
WAPDA	Water and Power Development Authority
WBANN	Wavelet–bootstrap Artificial Neural Networks

1. Introduction

Floods in Khyber Pakhtunkhwa and South Punjab have led to a desire for flood forecasting systems. In urban areas, floods can cause significant damage to property, infrastructure, and lives. These types of disasters result in significant economic and human losses globally, with over 15 billion in property damage and around 7,500 deaths occurring between 1985 and 2004. In areas with limited data on factors such as soil retention and permeation rates, snow exposure, and moisture levels, it can be difficult to accurately predict and mitigate the risk of flooding. This is particularly relevant in developing countries like Pakistan. Accurate river forecasting is necessary for water management, planning, and risk assessment. It can support economic activities, such as hydroelectric power and irrigation, and reduce the risk of flooding. In the HKH region, where many people depend on rivers for agriculture and economic activities, reliable river forecasting is especially crucial. There have been proposals for flood forecasting methods that range from traditional, physical-based techniques to more recent approaches using artificial intelligence algorithms. These AI algorithms can make accurate predictions by learning from past conditions and responses.

Tiwari et al (2010) pose an hourly flood prediction for Mahanadi River Basin India, they built the hybrid wavelet – bootstrap ANN (WBANN). The results illustrate that traditional ANN and WANN generate less accuracy than hybrid models like WBANN and BANN. Rezaeianzadeh et al (2014) note that the dataset of precipitation as input to find the behaviour of ANN and ANFIS techniques. The outcomes demonstrated that the area-weighted precipitation is better when applied as a contribution than ANNs and MNLR, while spatially differed rainfall contribution to the ANFIS and MLR models shows increasingly precise conjectures.

Furquim et al (2016) have worked to find the temporal correlation between unlike observations of the level of the channel and also to predict water level in a river. The results demonstrate that to our knowledge the MLP is stronger than the E-RNN. Latt (2015) provide an extensive overview of Muskingum flood routing with the addition of a feedforward artificial neural network called a black-box forecasting approach for making a model for flood forecasting and to decrease the disparity between actual and routed flows in the season of monsoon in Myanmar. The results showed that the FMLP model gives better results as compared to others.

Patel and Ramachandran (2015) used the ANN and SVR techniques of machine learning also with the ARIMA for the prediction of river flow in the river Cauvery in India. The result showed that the SVR gave a better performance in correlation and RMSE while the ANN gave a better performance in NRMSE and NSE. Hong and Hong (2016) use neural device models for the MLP water prevision at a control station located in downtown Kuala Lumpur, Malaysia, using upstream station records. The best implementation has been achieved with 15 upstream current and previous water level vectors, 7 hidden nodes, and a Kuala Lumpur core performance vector. R^2 data sets are 0.81, 0.85, and 0.85 for preparation, checking, and validation.

Tahmasebi et al (2016) used the ANN and data fusion technique for the mathematical modelling of forecasting. Regardless of the variety of information, the resulting model indicated more prominent exactness in predicting floods contrasted with models with less variety of information factors. Ghorbani et al (2016) worked to expect the monthly river flow by comparing SVM with the MLP and circular basis neural network (RBF) in the Zarrinehrud River in Iran. The findings suggest that in the SVM model, the weakness in the month-to-month waterway stream was not precisely the same in the RBF and MLP models.

In Liu et al (2017) a deep search algorithm has been suggested with a neural back package network (BPNN) and auto-encoders (SAE) stacked by the authors. Comparison of the results of proposed algorithms with the results of the BP model, the support vector machine (SVM), the RBF model, and the ELM model shows that SAE-BP is far better than others. Widiasari and Nugroho (2017) they have used the MLP model because ANN is useful in time series forecasting. The result showed that MLP has better outcomes in forecasting water rise levels on the downstream channel.

Jabbari and Bae (2018) have to evaluate and betterment of rainfall data and to the improvement of real-time flood forecasting models by Artificial Neural Network (ANN). They concluded that the implementation of ANN for inclination improvement increased the results in the 2002, 2007 and 2011 years. Ghumman et al (2018) used the past 30 year's data of rainfall, temperature, and discharge of the upper Indus river basin for the comparison of discharge results of three involvement types concerning original input variables. The results showed that Broyden Fletcher Goldfarb Shannon and radial basis function give better results to others obtained from ANN and SVR respectively.

Hussain et al (2020) used different machine learning techniques including SVR, MLP, and RF for the monthly prediction of the Hunza river. The results indicated that the RF performs far better than MLP and SVR. Campolo et al (1999) made a neural network model during heavy rainfall in Italy to analyze and predict the behaviour of the Tagliamento river. The developed model gives the Root Mean Square Error less than 4% when the model is used with a 1 to 4 hours time horizon. If the time is increased by this time limit, the accuracy of the model decreased and the model goes in the favor of flood forecasting.

Mitra et al (2016) proposed an integrated framework based on IoT and machine learning to forecast the risk of river basin flooding. The model uses ZigBee modifications to the mesh network for WSN data collection and a GPRS module for data transmittal across the internet. Wang et al (2017) updated the previous calculation of the Muskingum method by applying the Back Propagation Correction (BPC) to the semi-divided Xinanjiang model. Results showed that the accuracy of flood forecasting is comprehensively increased by this model. Phitakwinai et al (2016) have been working to forecast seven hours in advance the water level of the Ping waterway in the midtown territory of Chiang Mai, Thailand by using MLP with the cuckoo search (CS) algorithm. The results of the CS MLP model are far better than the regular multilayer perceptron (MLP) model. Ruslan et al (2015) projected a 1 hour in advance flood forecast model utilizing a Better MLPNN arrangement. For the forecasting of the level of water at Kelang river bowl 1 hour early using [4,10,1] structure current water level entered in the MLPNN structure. After this Improved MLPNN structure is used to improve the predicted results. Noteworthy improvement can be seen utilizing the Improved MLPNN structure from the first MLPNN structure. Puttinaovarat and Horkaew (2020) worked on the hydrological, meteorological, geospatial, as well as crowdsourcing data for the making of a flood forecasting model by integrating these data with the ML model. The results have shown that the MLP ANN gives more accurate results as a comparison of other techniques with correct percentages, MAE, Kappa, and RMSE of 0.89, 97.93, 0.01, and 0.10, respectively.

Hussain et al (2020) checked the suitability and capacity of the CNN 1D and ELM model for forecasting day by day, week after week, and month-to-month streams in the Gilgit River basin of Pakistan. The outcomes showed that the ELM model performed generally better than the CNN-1D model dependent on factual actions on a three-time scale. In Elsafi (2014) the Nile Flow at Dongola Station in Sudan was predicted by using ANN. The ANN model was developed to simulate flow in an upstream position at a defined stage in the river range. The research reveals that the ANN is an authentic way to detect the flood danger in the Nile. Darbandi and Pourhosseini (2018) worked on a Hybrid multi-layer perceptron to estimate monthly river flow, and multilayer perceptron ANN verified the subsequent results. A hybrid multilayer perceptron gives satisfactory results for the water flow forecast.

Berkhahn et al (2019) presented an artificial network-based model to forecast maximum volumes of water during a flash flood. After successfully testing this model in two different real catchments of different slopes this model may be suggested for real-time forecasting. Schoppa et al (2020) aimed to test the ability, on a daily scale of 95 river basins with heterogeneous characteristics, of the machine learning algorithm random forest to predict flood returns. The results of the random forest model illustrated that this model performs better compared to other conventional rainfall-runoff models. Dtissibe et al (2020) used discharge as input-output variables to design a flood forecasting model by using a multilayer perceptron. The designed model was tested in intensive trials and the results revealed that the proposed model was successful and has strong predictability. Ali and Shahbaz (2020) proposed an efficient ANN-based method for determining the impact on the streamflow of precipitation. The results verified that the ANN-based model can be an effective alternative for solving hydrological problems. Kumar and Yadav (2021) used ANN for simulating real-time flooding in the lower Tapi Basin. The results of the Levenberg Marquardt learning rule, the Feed–Forward network, and the Sigmoidal Axon transfer function are satisfactory. The observed values of discharge are in line with the predicted flood discharge of the ANN. Ali et al (2021) estimated the local scour depth around bridge piers by using ANN. They used the MATLAB software to train the ANN models with dataset of various parameters. The result showed that the ANN model with the Levenberg-Marquardt algorithm performed better than other ANN models. Sayari et al (2022) used the meta-learner mehods in forecasting natural and regulated river flow. They investigated the multiple EML and individual models but results showed that the EMLs outperformed the individual models in predicting more accurate and reliable results.

The problem of flood forecasting has been mainly explored by using the deterministic methods and very few of them addressed the real time flood scenario. Futhermore, most of them deals the flood forecating problem as a regression machine learning problem with few flooding attrebutes. Whereas, Pakistan is one of the region where climate change drastically impact the flooding in the region. Therefore, it is difficult to analyze the most dominating features in which the flooding depends. This study intends to fill this gap by considering 9 distinct features of flooding and creating a large benchmark real-time corpus for the flood forecasting. In addition, the previous studies only focused on the regression study of the flood forecasting. Whereas, we deal the flood forecasting problem both as regression and classification problem.

In this work, the main objective is to correctly classify the flood situation in the Hunza River by using 9 different types of features with the help of three classifiers: Linear SVC, RF, and MLP. For this purpose, the objective of this work is to transform the flood forecasting problem into a machine learning problem. After that, three different classical and deep learning algorithms are trained on the available Hunza River dataset constructed based on Water and Power Development Authority (WAPDA) provided data which includes 9 different features to predict the flood situation. The main contribution of this paper is as follows,

1. First, with the help of Pakistan WAPDA compiling the sensors data into a machine--understandable format to construct the gold standard dataset of the unique features for the last 31 years for the site.

2. Second, to validate the dataset, we first deal with the flood forecasting problem as a classification problem just to predict the flood condition, i.e., flood or no-flood. For this purpose, we used the feature selection method to obtain the best learning features (on which flooding depends) from the obtained dataset.
3. Third, to predict the exact discharge value by which the actual severity of the flood is dependent we use ANN. Because, by measuring the discharge value we can calculate the severity of the flood priorly, which helps to avoid any uncertain disaster by generating the flood warning to the adjacent areas.
4. Finally, for simulation purposes, we used the k-fold cross-validation approach to check the validity and diversity of the dataset and to obtain the realistic model which best maps to the practical scenario.

The rest of this paper is organized as follows. Section 2 discussed the study area. Section 3 discussed the methodology used in this work. Sections 4 and 5 discuss experimental setup and result and analysis, respectively. Section 6 concludes the proposed work.

2. Study Area

The research area is located at 35.562°N and longitudinally 74.23°E in the district of Hunza, Gilgit-Baltistan province of Pakistan. The research area ranges from 2438 m below the bottom to 4693 m above the maximum stage. The meteorological conditions of Hunza are normal in summers and extreme cold in winters. Snow and glaciers cover much of the area, which mostly holds temperatures below freezing throughout the year. The range of temperature is between −10°C and 35°C. The Hunza river starts with the water stream from the high northern heaps bordering Chinese and finishes at Gilgit by the river Gilgit. The average water outlet is 325 Q (m^3/s). The location map of the study area is shown in Figure 1. The Hunza River bay is one of eight sub-waterway bowls within the Upper Indus Basin, covering an area of 13734 km^2 with 1384 glaciers. The key cause of river rushing in this area of the Hunza is the snow and glaciers. The biggest ice sheets are arranged in the Upper Indus Basin, whose dissolving water stream into Hunza River Bowl. 80% of waterway spillover is added by icy masses and heavy snowfall dissolves at a height of over 3500 m, which is 20% of its catchment zone. Characteristics of the Study Area are given in Table 1.

3. Methodology

The flood forecasting task is treated as the supervised classification task. We deal with the problem as a binary classification problem because our goal is to distinguish between two classes: flood and no-flood. Furthermore, we have used the K-fold cross-validation approach for better estimation of performance. Three different machine learning algorithms were used for this classification task including RF, linear SVC of the support vector machine classifier and MLP of the Deep learning classifier. The numeric values obtained from the sensors after feature engineering methods are

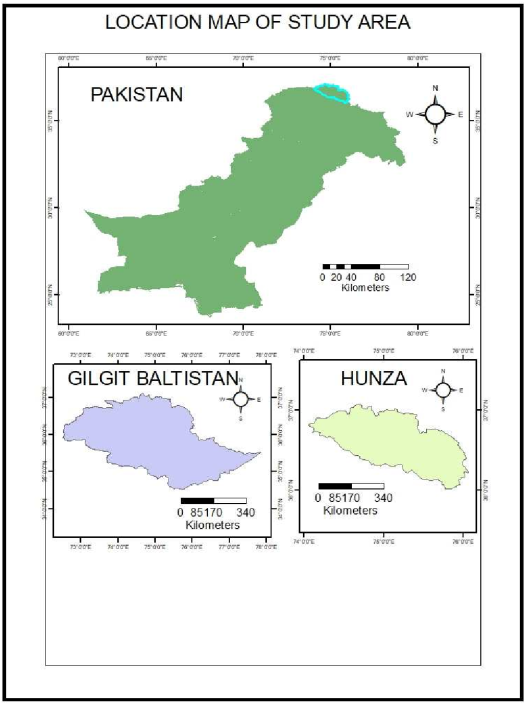

Fig. 1. Location Map of Study Area

Table 1. Characteristics of Study Area

Gauging Station	Danyor
Latitude	35°56′ N
Longitude	74°23′ E
Elevation of gauging station	1450 m
Drainage Area	13733 km^2
Glacier-covered area	4688 km^2
Glacier cover Percentage	34%
Mean Elevation (Computed from ASTER GDEM)	4631 m
Area above 5000 m	32.5%
No. of meteorological stations (Installed by WAPDA)	3

the input of these classifiers. In this work, the RFE method is used to select the most distinct features from the extracted features. Many researchers have revealed that the flood depends on some unique features. These flood features develop a pattern for identifying flood forecasting by extracting the unique features from the environment.

In this work, many features, i.e., precipitation, airspeed, dew factor, humidity, wind, and evaporation for flood forecasting have been used. A dataset is constructed of those features for the last 31 Years of Hunza River, in Pakistan. There are many other features as well but the performance is calculated based on 9 features. After that, the model is trained and tested by using two different categories of machine learning algorithms, classical and deep learning. The steps which have been done for the flood forecasting are given in Figure 2.

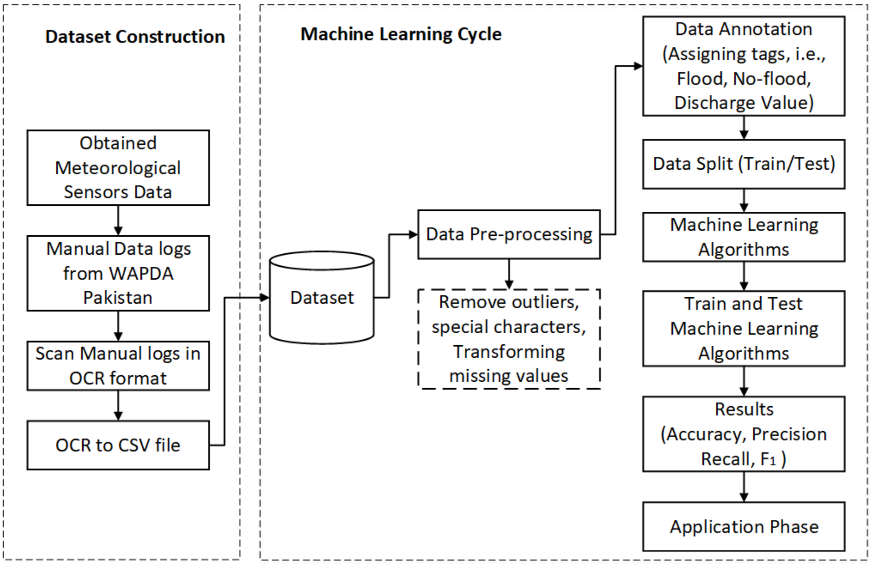

Fig. 2. Methodology adopted for the Flood Forecasting Task

We have constructed the dataset from the data provided by the WAPDA Pakistan for the last 31 years. The data is monitored by the meteorological stations via on-site sensors. After that, to maintain the record the data is maintained in manual documents for record purposes. We obtained this data and after scanning converts it into the Optical Character Recognition (OCR) format. After that, we convert it into Comma Separated Value (CSV) file to shape up the data in a structured format. The data is in numeric forms with some of the missing values due to some hardware limitations. The dataset construction steps can be illustrated in Figure 2 and the characteristics of a dataset are given in Table 2.

Table 2. Characteristics of dataset

Dataset Characteristics	Count
Total Instances	11323
Flood	6248
No Flood	5075

A simple Python script is written that reads the dataset, preprocessed it (removing URLs, unwanted spaces, and special characters, and completes the missing data), and extracts the features by transforming the environmental data into attribute-value pairs of airspeed, wind speed, precipitation, humidity and dew factor. A machine cannot understand the raw data; it can only understand the numeric values. After this step is completed, we now move towards the actual classification part by first splitting all our data into the training and test sets.

We used a 10-fold cross-validation approach to split the data set with multiple subsets and select any pair for the estimation of the performance using features. We also used the train test split module from the sklearn. model selection library. We set the ratios of the division so that our data is divided into the corresponding ratio automatically. This split approach is used after the feature selection method. Before that, the algorithms are trained by using a 10-fold cross-validation approach which returns the classifier with the highest accuracy score. We divide the data into 20% test set and 80% training set. This utility also shuffles or randomizes the data automatically. To improve the accuracy of the classification and reduce the time needed by the algorithms to produce the results, we need to select the most distinguishing features from our data. The classification report consists of *accuracy, precision, recall and* F_1 is obtained by using the following equations.

$$Accuracy = \frac{\mathrm{TP} + \mathrm{TN}}{\mathrm{TP} + \mathrm{TN} + \mathrm{FP} + \mathrm{FN}}, \tag{1}$$

$$Precision = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FP}}, \tag{2}$$

$$Recall = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}}, \tag{3}$$

$$F_1 = \frac{2 \cdot Precision \cdot Recall}{Precision + Recall}. \tag{4}$$

where TP, TN, FP, and FN represent True Positive, True Negative, False Positive and False Negative, respectively. The environmental data consists of a large number of features and the most distinct features are extracted by applying feature extraction models to construct a large feature space. It is exhaustive for most of the classifiers to deal with such a large feature space. To mitigate this issue, multiple feature selection methods are used to extract the most discriminating features from the large feature space and remove the redundant features. In this work, we used multiple sensors for feature extraction. Whereas, for feature selection, we have used the Recursive Feature Elimination (RFE) approach. It gives an external estimator that assigns weights to the features to select a smaller to a smaller set of features. We have used linear SVC as an external estimator in feature selection. We also set the step parameter of RFE to 1 so that it eliminates a single feature in each iteration. After feature selection, we are now ready for the final and real task of flood forecasting.

Furthermore, to predict the accurate value to discharge before avoiding any uncertain situation of the flooding, we can also deal with this machine learning problem as a regression problem. The actual value of discharge is considered as an output. The same features are used in this regression problem as we discussed in the classification problem. While dealing with a flood forecasting problem as a classification problem we annotate the dataset, but in the case of regression analysis, we used unannotated data as obtained from the WAPDA Pakistan. To solve this regression problem, we used two different machine learning algorithms, i.e., MLP and SVR. The results are obtained by using similar parameters of both algorithms, the only difference is the activation function of the models on the output layer. The results are evaluated based on Mean Absolute Error (MAE), Mean Squared Error (MSE), and Root Mean Squared Error (RMSE). MAE calculates the difference between the actual and forecasted values derived from the overall cumulative difference in the data collection. The main purpose is to using MAE as an evaluation measure is that, we are doing classification as well, so, in regression, the MAE maps the accuracy of the trained model. In our proposed dataset, there are some variations in the parameters due to uncertain climate conditions. Therefore, we also report the MSE and RMSE along with the MAE. The calculation of MAE, MSE, and RMSE is represented in the following equations.

$$\mathrm{MAE} = \frac{1}{N}\sum_{I=1}^{N}\left(y_p - y_a\right) \tag{5}$$

where N is number of total instances in test data, y_p is the predicted output and y_a is the actual output.

MSE is the variance between the initial and expected values, which are derived by squared data from the average difference.

$$\mathrm{MSE} = \frac{1}{N}\sum_{I=1}^{N}\left(y_p - y_a\right)^2. \tag{6}$$

RMSE is the error rate by the square root of MSE.

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{I=1}^{N}\left(y_p - y_a\right)^2}. \tag{7}$$

The brief description and basic working of the machine learning algorithms we adopt are as follows.

3.1. Artificial Neural Network Algorithm

Artificial Neural Networks (ANN) are a type of Machine Learning system modeled on the biological neurons and synapses within the human brain (Gholami et al 2015).

ANNs approximate complex non-linear functions, can use large datasets for training, and provide extremely accurate results for classification, prediction, and pattern recognition tasks. A Multi-Layer Perceptron (MLP) ([Delashmit et al 2005) is an ANN architecture consisting of three or more layers: an input layer, an output layer, and one or more hidden layers that process given inputs to produce desired outputs with the help of weights attached to each connection between nodes. The MLP can be interpreted mathematically in Eq. 8

$$y = f\left(\sum_{i=1}^{n} j_i r_i + k\right). \tag{8}$$

where j_i is the weight vector, r_i is input vector, k is bias, i is $1 \dots n$, f is the transfer function and y refers to output. To calculate the output at each layer the activation function is used. We are using Rectified Linear Unit (ReLU) function at hidden and input layers and sigmoid the output for the classification problem. The ReLU function is a piece-by-piece function that outputs the entry directly, otherwise, it outputs zero. For certain varieties of neural networks, it is becoming the default activation mechanism when a model is more educated and therefore performs better. The architecture of ANN used in this work with 9 features is given in Figure 3.

3.2. Support Vector Machine Algorithm

The mechanism of this algorithm is slightly different from the previously discussed algorithm. It is based on a hyper-line which separates the false and true labels. The placement of class near or close to this line tells us how confident we are for this particular class of the given example (Vishwanathan and Narasimha Murty 2002). For instance, if the cline falls close to the boundary then we can say that we have low confidence. In it, we plot the data in n-dimensions and where each value is represented by its coordinates. If we classify these coordinated with hyper-line the classes may also be differentiated.

3.3. Random Forest Algorithm

RF algorithm is a supervised machine learning algorithm (Biau and Scornet 2016). This algorithm, as the name implies, generates several trees in the forest. In general, the stronger the forest looks the more trees in the forest. Likewise in the random forest classifier, the more trees in the forest, the more accurate results are obtained. This algorithm is designed especially for the analysis of high-dimensional multi-class results. In this work, we are dealing with only two classes, i.e., flood and no-flood. RF groups have high characterization precision, endure anomalies and commotion well and have a narrow chance of overfitting. RF is a set of separate classification and regression, proposed by Breiman in 2001, which can then be described in equation 9:

$$a(t, \theta_j), \quad j = 1, 2, \dots, n, \tag{9}$$

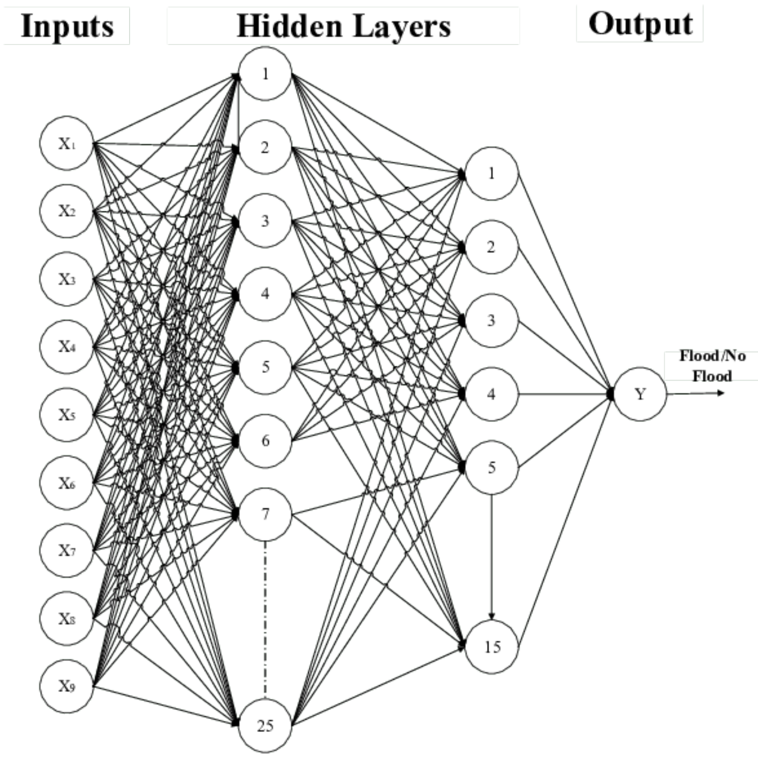

Fig. 3. Architecture of ANN

where a is an RF classifier, t is an input variable and θ_j refers to the independent and identical distribution of the input used for producing each tree. RF's final answer is based on the performance of all the concerned decision-making trees. RF has a comparatively smaller statistical burden compared to other machine learning approaches and is resistant to cross-linear variables or outliers. Furthermore, RF exhibits excellent performance in large-scale features that have great potential for the classification of composite and texture-abundant UAV images. A further valuable benefit of RF is that the value of input variables can be calculated so that scientists can better consider how each variable relates to the precision of class. The architecture of the RF classifier is given in Figure 4.

4. Results and Analysis

As we discussed earlier that we have to deal with this flood forecasting problem as both classification and regression problem. The characteristics of the dataset we used are discussed in Table 2. We annotate the dataset based on discharge value and the historical data of flooding. Therefore, based on discharge value the dataset is classified into two categories and for flood its 6284 instances and 5075 for the no-flood category. The dataset is large, balanced, and of high quality. The performance of the algorithms

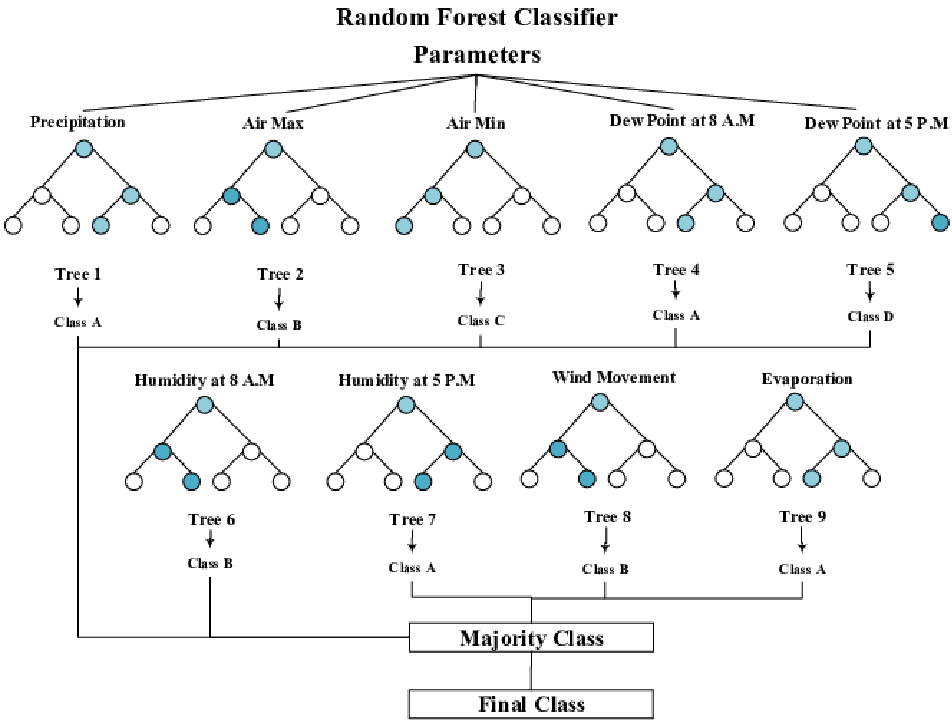

Fig. 4. Tree Structure of Random Forest Classifier

is evaluated based on the accuracy and confusion matrix. The data is nearly balanced therefore the accuracy is the better evaluation measure and to look for deep insight into the performance of an algorithm, we used a confusion matrix as well.

In this task, the performance of the classifiers is compared with the baseline approach called Most Common Category (MCC). The accuracy of MCC is calculated by assigning the most common category and according to this, the MCC accuracy is 0.55. Moreover, to justify the results as our dataset slightly has a large number of instances from the no-flood category, therefore, we also report *precision*, *recall* and F_1 score. For training and prediction, we have used RF, Linear SVC, and MLP machine learning algorithms. For the RF classifier, we set the parameters n_estimators = 1000 for the number of trees in the forest, and random state = 0 for the selections of instances to remain the same during each iteration, so that our classification results remain consistent. For the MLP classifier, we used 'relu' as an activation function to calculate process the input and calculate output for the next hidden layers and 'sigmoid' at the output layer to predict output as '0' and '1'; '0' for No-flood and '1' for flood. Furthermore, we used two hidden layers, the first consists of 25 units and the second is 15 units, and the learning rate is 0.01. Detail Parameters of MLP Classifier and Regression are given in Table 3. For the Linear SVC algorithm, we used default parameters.

Table 3. Parameters for MLP classifier and regression

Parameter	For Classifier	For Regression
Activation Function	Relu	Relu
Batch Size	Auto	Auto
Hidden Layer Size	25, 15	25, 15
Learning Rate	Constant	Invscaling
Learning Rate Init	0.001	0.001
Max Iter	200	1000
Random State	1	1
Solver	Lbfgs	Adam
Tolerance	0.0001	0.0001
N iter no change	5	5
Alpha	–	1e-05

After applying the algorithms, we calculate the prediction basis on the data-driven Hydrological method. The detailed performance of the algorithms after applying the feature selection method is illustrated in Table 4. It is clear from Table 4, that the feature selection has an impact on the results, and RF outperforms the Linear SVC and MLP classifier in terms of accuracy. The best result is obtained when we considered the feature selection method. This shows that the combination of distinct features improves the performance of the algorithms. The deep insight of the classifier is clear from the confusion matrices given in Figure 5. It is clear from Figure 5(a) that how many instances are truly classified by the machine learning classifier, i.e. the 1054 instances are truly classified as No-Flood because "0" is for No-Flood and "1" is for Flood, and 0 instances are misclassified as No-Flood. In contrast, in Figure 5(b) you can see that some of the instances are misclassified. Although, they are small in number. We can say that in this dataset all the classifiers perform well and are ready to go in the application phase, where they can be deployed in the real world to solve real-world problems. We have also developed its application phase as well and the results are quite astonishing. The k-fold cross validation result after training is depicted in Figure 6. It is clear from Figure 6 that the RF classifier outperforms in every fold. For each fold, the accuracy of the classifier is approaching 1. However, the mean accuracy of the RF classifier is 0.99.

Table 4. Details on performance of algorithms

Classifier	Accuracy	Precision	Recall	F1-Score
MCC	0.55	–	–	–
Linear SVC	0.98	0.97	0.97	0.97
RF	0.99	1.00	1.00	1.00
MLP	0.98	0.99	0.99	0.99

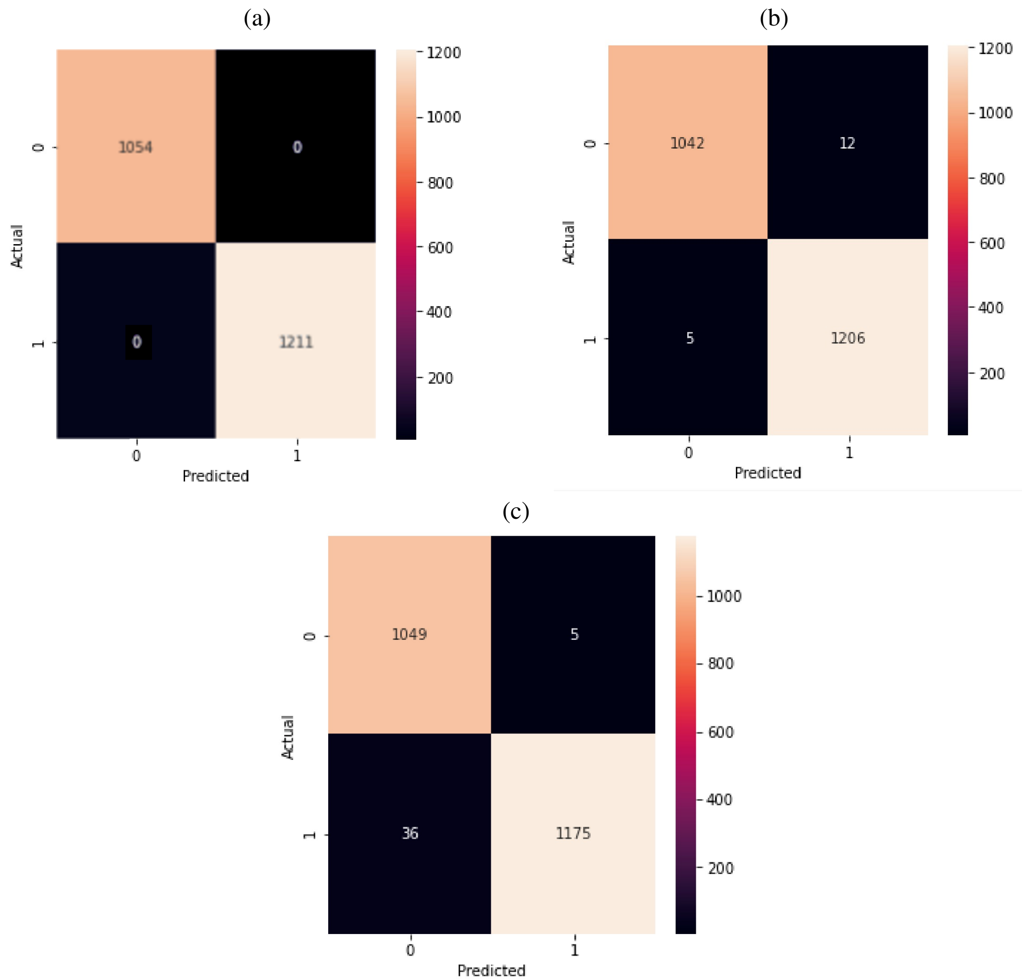

Fig. 5. Confusion matrix

After measuring the accuracy of flood in the classification process a comparison of two machine learning algorithms (MLP and SVR) has been carried out to appraise their performance. The model accuracy assessments are usually carried out according to the error variation between actual and forecast values, however, several studies have used different performance assessment measures. Legates and McCabe (1999) has shown that it is useful to look at the degree of accuracy for "good-fit" measures which demonstrated the regression and absolute error statistical measurements. Therefore, we also carried out the regression analysis of our proposed dataset. As we discussed in Section 3, the dataset is constructed on basis of real-time parameters measured by the meteorological department of WAPDA. The main issue of this type of data collection is the high diversity and variance in data. As the dataset is of large amount, diverse, and high variance. Therefore, we adopt Tukey's method to

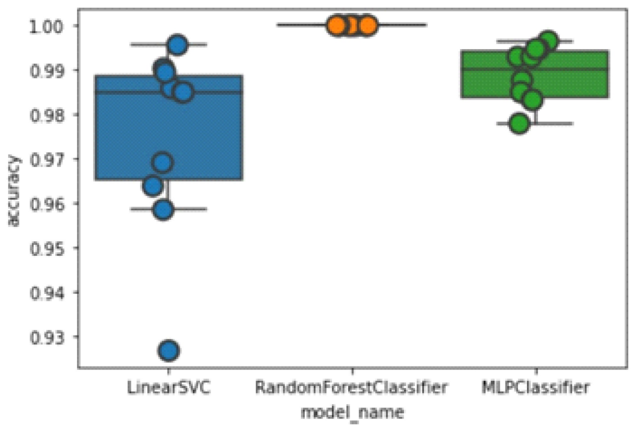

Fig. 6. Accuracy results after 10 fold cross-validation

remove outliers from such data (Kolbaşi and Ayidin (2019). This technique is adapted where the variance is right or left skewed from its mean value. These outliers have no/little impact on the overall performance if removed. Hence, we used this technique to filter the dataset and it improves the performance of the deployed model in case of regression to predict the water discharge value. The performance of the regression models is evaluated by using MAE, MSE, and RMSE as performance measurements. The results in Table 5 show that the MAE score of the MLP and SVR is 0.16 and 0.06, respectively. This shows that the model predicted values and actual values have a mean deviation of 0.16 and 0.06 in the case of MLP and SVR models, respectively. The MSE and RMSE further elaborate on the results in the case of variance in the dataset. Hence, the remaining variation in the predicted values is penalized by these evaluation measures. Our results demonstrate that the SVR model outperformed the MLP model based on the performance measures. This is because of the nature of the dataset because we have removed noise from the data so SVR took advantage of that and outperforms MLP. In the case of noisy data, the performance of the MLP is quite better than other classical machine learning algorithms. The complete numeric results of regression analysis are given in Table 5. Furthermore, Figure 7 shows the yearly hydrograph of the peak values of runoff from each year. The graph shows the time on X-axis and the discharge rate on Y-axis. It is clearly seen from the graph that there is uncertainty present in the peak discharge which is solved in this study by classification and regression analysis.

5. Conclusions

In this study, a dataset was developed for flood forecasting in the Hunza River in Pakistan over the last 31 years using data collected from various sensors. The dataset was then used to predict floods using machine learning algorithms for classification and

Table 5. Regression results of MLP and SVR

Models	MSE	RMSE	MAE	R^2
Proposed Linear SVC	0.004	0.07	0.06	0.85
Proposed MLP	0.05	0.24	0.16	0.95
MLP (Rezaeianzadeh et al (2014)	–	–	–	0.91
MLP (Furquim et al (2016)	–	–	–	0.96
MLP & Linear SVC (Patel et al (2015)	–	–	–	0.78 & 0.79
MLP (Hong and Hong (2016)	–	–	–	0.85
MLP (Tahmasebi et al (2016)	–	–	–	0.91
Linear SVC (Ghumman et al (2018)	–	–	–	0.811
MLP & Linear SVC (Hussain et al (2020)	–	–	–	0.91 & 0.83

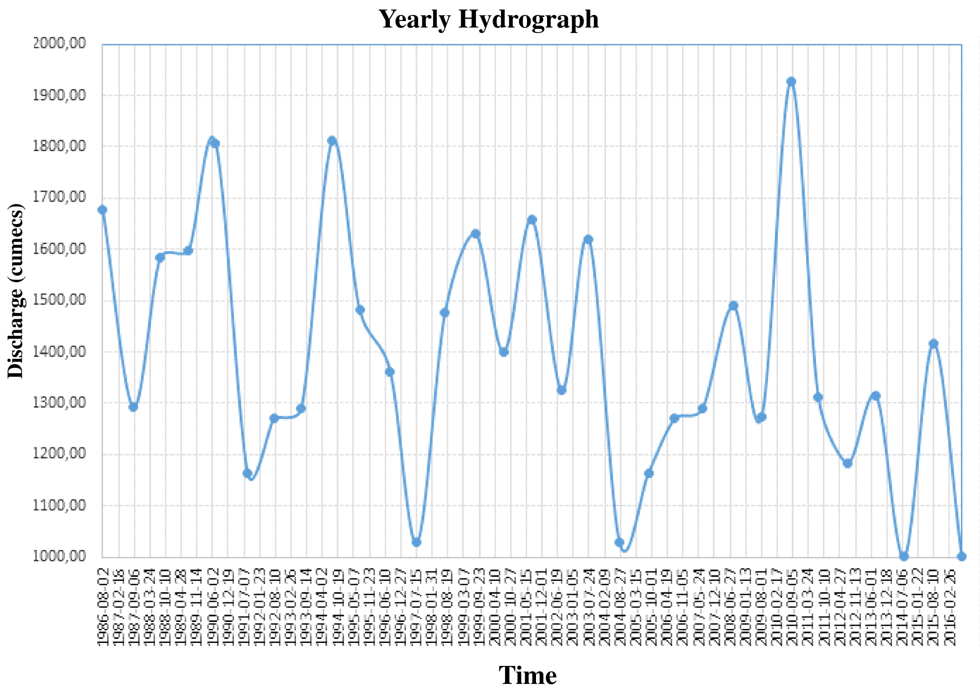

Fig. 7. Runoff Hydrograph

regression. Two families of machine learning algorithms, classical and deep learning, were tested using three different classifiers (Random Forest, Linear Support Vector Classifier, and Multi-Layer Perceptron) and two regression algorithms (Multi-Layer Perceptron and Support Vector Regression). The results of the k-fold cross-validation showed that the Random Forest classifier was the most accurate in predicting floods, while the Support Vector Regression algorithm was the most accurate in predicting discharge values. The Random Forest classifier was found to be more accurate than the Linear Support Vector Classifier and the Multi-Layer Perceptron, with an accuracy of 0.99 in the classification study. In the regression analysis, the Support Vector Regression algorithm outperformed the Multi-Layer Perceptron, with better results in terms of mean absolute error, mean squared error, and root mean squared error. Overall, the results of this study show the potential for using machine learning algorithms to accurately predict floods and discharge values in rivers, as has been shown on the example of the Hunza River.

References

Ali S., Shahbaz M. (2020) Streamflow forecasting by modeling the rainfall–streamflow relationship using artificial neural networks, *Modeling Earth Systems and Environment*, **6** (3), 1645–1656.

Ali A. S. A., Günal M. (2021) Artificial Neural network for estimation of local scour depth around bridge piers, *Archives of Hydro-Engineering and Environmental Mechanics*, **68** (2), 87–101.

Berkhahn S., Fuchs L., Neuweiler I. (2019) An ensemble neural network model for real-time prediction of urban floods, *Journal of Hydrology*, **575**, 743–754.

Biau G., Erwan S. (2016) A random forest guided tour, *Test*, **25** (2), 197–227.

Breiman L. (2001) Random forests, *Machine Learning*, **45** (1), 5–32.

Campolo M., Andreussi P., Soldati A. (1999) River flood forecasting with a neural network model, *Water Resources Research*, **35** (4), 1191–1197.

Darbandi S., Pourhosseini F. A. (2018) River flow simulation using a multilayer perceptron-firefly algorithm model, *Applied Water Science*, **8** (3), 1–9.

Delashmit W. H., Manry M. T. (2005) Recent developments in multilayer perceptron neural networks, *Proceedings of the seventh Annual Memphis Area Engineering and Science Conference*, MAESC.

Dtissibe F. Y., Ari A. A. A., Titouna C., Thiare O., Gueroui A. M. (2020) Flood forecasting based on an artificial neural network scheme., *Natural Hazards*, **104** (2), 1211–1237.

Elsafi S. H. (2014) Artificial neural networks (ANNs) for flood forecasting at Dongola Station in the River Nile, Sudan, *Alexandria Engineering Journal*, **53** (3), 655–662.

Furquim G., Pessin G., Faiçal B. S., Mendiondo E. M., Ueyama J. (2016) Improving the accuracy of a flood forecasting model by means of machine learning and chaos theory, *Neural Computing and Applications*, **27** (5), 1129–1141.

Gholami V., Darvari Z., Mohseni Saravi M. (2015) Artificial neural network technique for rainfall temporal distribution simulation (Case study: Kechik region), *Caspian Journal of Environmental Sciences*, **13** (1), 53–60.

Ghorbani M. A., Zadeh H. A., Isazadeh M., Terzi O. (2016) A comparative study of artificial neural network (MLP, RBF) and support vector machine models for river flow prediction, *Environmental Earth Sciences*, **75** (6), 476.

Ghumman A. R., Ahmad S., Hashmi H. N. (2018) Performance assessment of artificial neural networks and support vector regression models for stream flow predictions, *Environmental Monitoring and Assessment*,**190** (12), 704.

Hussain D., Hussain T., Khan A. A., Naqvi S. A. A., Jamil A. (2020) A deep learning approach for hydrological time-series prediction: A case study of Gilgit river basin, *Earth Science Informatics*, **13** (3), 915–927.

Hussain D., Khan A. A. (2020) Machine learning techniques for monthly river flow forecasting of Hunza River, Pakistan, *Earth Science Informatics*, 1–11.

Hong J. L., Hong K. (2016) Flood forecasting for Klang river at Kuala Lumpur using artificial neural networks, *International Journal of Hybrid Information Technology*, **9** (3), 39–60.

Jabbari A., Bae D. H. (2018) Application of Artificial Neural Networks for Accuracy Enhancements of Real-Time Flood Forecasting in the Imjin Basin, *Water*, **10** (11), 1626.

Kolbaşi A., Aydın Ü. (2019) A comparison of the outlier detecting methods: an application on Turkish foreign trade data, *J. Math. Stat. Sci.*, **5**, 213–234.

Kumar V., Yadav S. M. (2021) Real-Time Flood Analysis Using Artificial Neural Network, [In:] *Recent Trends in Civil Engineering*, Springer, Singapore, 973–986.

Latt Z. Z. (2015) Application of feedforward artificial neural network in Muskingum flood routing: a black-box forecasting approach for a natural river system, *Water Resources Management*, **29** (14), 4995–5014.

Legates D. R., McCabe Jr G. J. (1999) Evaluating the use of "goodness-of-fit" measures in hydrologic and hydroclimatic model validation, *Water Resources Research*, **35** (1), 233–241.

Liu F., Xu F., Yang S. (2017) A flood forecasting model based on deep learning algorithm via integrating stacked auto encoders with BP neural network, *2017 IEEE third International conference on multimedia big data (BigMM)*, IEEE, 58–61.

Mitra P., Ray R., Chatterjee,R., Basu R., Saha P., Raha S., Saha S. (2016) Flood forecasting using Internet of things and artificial neural networks, *2016 IEEE 7th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON)*, IEEE, 1–5.

Patel S. S., Ramachandran P. (2015) A comparison of machine learning techniques for modeling river flow time series: the case of upper Cauvery river basin, *Water Resources Management*, *29* (2), 589–602.

Phitakwinai S., Auephanwiriyakul S., Theera-Umpon N. (2016) Multilayer perceptron with Cuckoo search in water level prediction for flood forecasting, *2016 International Joint Conference on Neural Networks (IJCNN)*, 519–524.

Puttinaovarat S., Horkaew P. (2020) Flood Forecasting System Based on Integrated Big and Crowdsource Data by Using Machine Learning Techniques, *IEEE Access*, **8**, 5885–5905.

Rezaeianzadeh M., Tabari H., Yazdi A. A., Isik S., Kalin, L. (2014) Flood flow forecasting using ANN, ANFIS and regression models, *Neural Computing and Applications*, **25** (1), 25–37.

Ruslan F. A., Tajuddin M., Adnan R. (2015) Flood prediction modeling using improved MLPNN structure: Case study Kuala Lumpur, *2015 IEEE Conference on Systems, Process and Control (ICSPC)*, IEEE, 101–105.

Sayari S., Meymand A. M., Aldallal A., Zounemat-Kermani M. (2022) Meta-learner methods in forecasting regulated and natural river flow, *Arabian Journal of Geosciences*, **15** (11), 1–12.

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

Tahmasebi Biragani Y., Yazdandoost F., Ghalkhani H. (2016) Flood Forecasting Using Artificial Neural Networks: an Application of Multi-Model Data Fusion Technique, *Journal of Hydraulic Structures*, **2** (2), 62–73.

Tiwari M. K., Chatterjee C. (2010) Development of an accurate and reliable hourly flood forecasting model using wavelet–bootstrap–ANN (WBANN) hybrid approach, *Journal of Hydrology*, **394** (3–4), 458–470. Vishwanathan S. V. M., M. Narasimha Murty (2002) SSVM: a simple SVM algorithm, *Proceedings of the 2002 International Joint Conference on Neural Networks. IJCNN'02 (Cat. No. 02CH37290)*, IEEE, Vol. 3, 2393–2398.

Wang J., Shi P., Jiang P., Hu J., Qu S., Chen X., Xiao Z. (2017) Application of BP neural network algorithm in traditional hydrological model for flood forecasting, *Water*, **9** (1), 48.

Widiasari I. R., Nugroho L. E. (2017), Deep learning multilayer perceptron (MLP) for flood prediction model using wireless sensor network based hydrology time series data mining, *2017 International Conference on Innovative and Creative Information Technology (ICITech)*, IEEE, 1–5.