



Flood Forecasting System using ML and BD

Vaishnavi V¹, P. Jahnavi Sai², Pranam B U³, Arjun Kumar A V⁴

East West Institute of Technology (ISE)¹

East West Institute of Technology (ISE)²

East West Institute of Technology (ISE)³

East West Institute of Technology (ISE)⁴

Abstract: Flood is one of the most disruptive natural hazards, responsible for loss of lives and damage to properties. A number of cities are subject to monsoons influences and hence face the disaster almost every year. Early notification of flood incident could benefit the authorities and public to devise both short and long terms preventive measures, to prepare evacuation and rescue mission, and to relieve the flood victims. Geographical locations of affected areas and respective severities, for instances, are among the key determinants in most flood administration. Thus far, an effective means of anticipating flood in advance remains lacking. Existing tools were typically based on manually input and prepared data. The processes were tedious and thus prohibitive for real-time and early forecasts. Furthermore, these tools did not fully exploit more comprehensive information available in current big data platforms. Therefore, this project proposes a novel flood forecasting system based on fusing meteorological, hydrological, geospatial, and crowd source big data in an adaptive machine learning framework. Data intelligence was driven by state of the-art learning strategies. Subjective and objective evaluations indicated that the developed system was able to forecast flood incidents, happening in specific areas and time frames. It was also later revealed by benchmarking experiments that the system configured with an MLP ANN gave the most effective prediction, with correct percentage, Kappa, MAE and RMSE of 97.93, 0.89, 0.01 and 0.10, respectively.

Keywords:

- Flood forecasting
- Natural hazards
- Monsoon influences
- Early notification
- Preventive measures
- Evacuation and rescue
- Geographical locations
- Big data
- Machine learning

I. INTRODUCTION

Overview

Natural flood is one of the most recurrent disasters. Unlike stagnant water discharge, occasionally experienced in poorly planned cities, major flood incidents always cause considerable damages to properties and, more often than not, loss of lives. Several Asian countries, particularly Thailand, are subject to both southwest and northeast monsoons and accordingly facing seasonal deluge almost every year and in most parts of the countries.

Among notable causes, sudden and enduring heavy rain is the most pertinent one in Thailand. Furthermore, overflow from main rivers along shore sides to surrounding basins can greatly spread the damages. Although being located further away from a river, an area with inappropriate land uses are unable to efficiently discharge accumulated precipitation, and hence are inevitably prone to even more frequent floods. Regardless of causes, however, a flood is generally sudden and thus almost formidable for the general public and relevant organization to be adequately prepared for the incident. This is mainly due to the lack of an effective means of anticipating the disaster well in advance.

Despite the recent extensive development of computerized flood forecasting systems, they remained based primarily on present precipitation, monitored by rain stations or rain gauges. These facilities are normally owned by a meteorology department or similar organizations.



Besides, they are scantily located in a few areas due to costly installation and maintenance. Hence, it is difficult to determine precipitation or predict flood accurately, especially in areas with no such facility. To remedy this issue, precipitation in these areas were typically estimated either by inter- or extrapolation from those with rain stations present. Due to a limited number of these stations and those readings in one area may not be a good representative to others. Therefore, estimated precipitation was insufficiently accurate to make a realistic forecast. Conventional meteorological readings, e.g., precipitation, temperature, and humidity, etc., took really long time to measure, process, record, and transfer to relevant organizations. Analyses based on past precipitation were known to be associated with several short comings.

For instance, they contribute to inaccurate and often outdated flood prediction. Limited sample size, inadequate computing capability, and inefficient prediction methods were all undermining the real potential of this scheme. Nonetheless, with the recent advances in distributed computing and especially modern machine learning (ML), resembling human intelligence, computerized flood forecasting, based on thematic factors has widely been investigated. In addition, as the number of both open and proprietary data providers escalates, Big Data has now become a central source of information in such pursuits

Thus far, according to recent surveys, most flood forecasting systems relied primarily on either monitored precipitation data or those obtained from a single source. Beside the mentioned limitations, existing systems remained lacking in other various aspects. For example, in a case where monitoring facilities or communication network of ones became malfunctioned, there would be no precipitation data available for imperative analyses. To the best of our knowledge, there was also no tool (software) that can accommodate area specific forecasting well in advance. Furthermore, existing tools were highly dependent on demanding data preparation and compilation from various sources, including Big Data. As a consequent, automated and spontaneous notification of flood incidents to the public and authorities, or realistic anticipation of ones has remained a grand challenge.

In addition, there have been recent developments in flood forecasting systems based on ML. These systems embedded both attributes and crowd sourcing data into their ML frameworks. However, most existing systems operate by analyzing these data offline on premise before presenting their prediction results on various platforms. A typical practice was proposed in, where an ML was trained with real-time rainfalls, streamflow, and other data. It was unclear, nonetheless, how a prediction result was verified against an actual event, which was obtained from crowd sourcing.

II. PROBLEM STATEMENT

- Flood is one of the most disruptive natural hazards, responsible for loss of lives and damage to properties. A number of cities are subject to monsoons influences and hence face the disaster almost every year.
 - Early notification of flood incident could benefit the authorities and public to devise
 - Short and long terms preventive measures, to prepare evacuation and rescue mission, and to relieve the flood victims.
 - Existing tools were typically based on manually input and prepared data. The processes were tedious and thus prohibitive for real-time and early forecasts.
 - Existing tools did not fully exploit more comprehensive information available in current big data platforms.
 - This project proposes a novel flood forecasting system based on fusing meteorological, hydrological, geospatial, and crowdsource big data in an adaptive machine learning framework
 - Data intelligence was driven by state-of-the-art learning strategies.
- Subjective and objective evaluations indicated that the developed system was able to forecast flood incidents, happening in specific areas and time frames.
- System will be configured with an MLP ANN give the most effective prediction.

III. SIGNIFICANCE AND RELEVANCE OF WORK

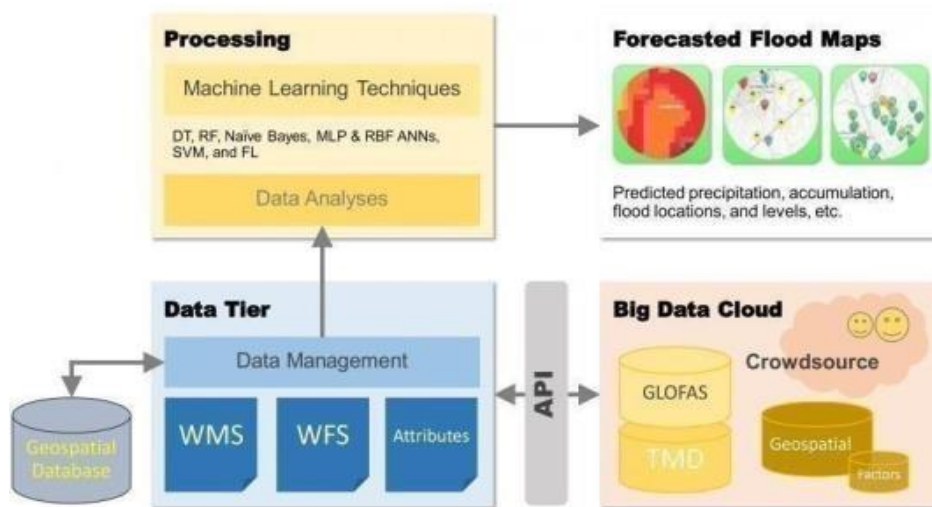
- To elucidate the merits of the proposed scheme, the experiments were carried out on Chennai coastal places. We will consider the geographical information for the preparation of dataset.
- These areas are under influences of both southeast and northwest monsoons. As a consequence, they both are prone to heavy floods.
- In this approach would show that no restriction on data nor fundamental processes was imposed regarding these specific provinces. Therefore, the proposed scheme could be generalized and applied equally well to other areas



IV. OBJECTIVES

To collect the flood forecasting data such as

1. Geospatial,
 2. Meteorological and hydrological, obtained from GLOFAS, 3. Crowdsorce (or volunteer) data.
- Normalization of data and make a common coordinate frame, a preprocessing need to be done. Data integration pre-processed data.
 - Prediction and evaluating the results using MLP ANN, RF, SVM.



V. METHODOLOGY

- This project proposed a novel distributed flood forecasting system, based on integrating meteorological, hydrological, geospatial, and crowd source data.
- Big data made available by prominent agencies were acquired by means of various cross platform APIs. Forecasting was performed based on these data learned by modern ML strategies.
- They were decision tree, RF, Naïve Bayes, MLP and RBF ANN, SVM, and fuzzy logics
- It was elucidated empirically that the developed system could be used to alert the public and authorities alike of not only a current flood but also future ones.
- This system also enhanced user experience via responsive graphical interfaces, interoperable on different computing devices including mobiles
- This advantage effectively encouraged greater contribution of crowdsorce data from the public, enriching data aggregation and hence increasing system accuracy and reliability.

VI. LITERATURE SURVEY

6.1 Flood Forecasting System Based on Integrated Big and Crowdsorce Data by Using Machine Learning Techniques.[48]

Authors : S. Puttinaovarat and P. Horkaew

Flood is one of the most disruptive natural hazards, responsible for loss of lives and damage to properties. A number of cities are subject to monsoons influences and hence face the disaster almost every year. Early notification of flood incident could benefit the authorities and public to devise both short and long terms preventive measures, to prepare evacuation and rescue mission, and to relieve the flood victims. Geographical locations of affected areas and respective severities, for instances, are among the key determinants in most flood administration. Thus far, an effective means of anticipating flood in advance remains lacking. Existing tools were typically based on manually input and prepared data.



The processes were tedious and thus prohibitive for realtime and early forecasts. Furthermore, these tools did not fully exploit more comprehensive information available in current big data platforms. Therefore, this paper proposes a novel flood forecasting system based on fusing meteorological, hydrological, geospatial, and crowdsourced big data in an adaptive machine learning framework. Data intelligence was driven by state-of-the-art learning strategies. Subjective and objective evaluations indicated that the developed system was able to forecast flood incidents, happening in specific areas and time frames.

Disadvantages:-

Data considered in this study, such as GLOFAS, were not of intrinsically high spatial resolution

6.2 Flood Prediction Using Machine

Learning Models: Literature Review

Authors : Amir Mosavi , Pinar Ozturk and Kwokwing Chau[49]

The research on the advancement of flood prediction models contributed to risk reduction, policy suggestion, minimization of the loss of human life, and reduction of the property damage associated with floods. To mimic the complex mathematical expressions of physical processes of floods, during the past two decades, machine learning (ML) methods contributed highly in the advancement of prediction systems providing better performance and cost-effective solutions. Due to the vast benefits and potential of ML, its popularity dramatically increased among hydrologists.

Researchers through introducing novel ML methods and hybridizing of the existing ones aim at discovering more accurate and efficient prediction models. The main contribution of this paper is to demonstrate the state of the art of ML models in flood prediction and to give insight into the most suitable models. In this paper, the literature where ML models were benchmarked through a qualitative analysis of robustness, accuracy, effectiveness, and speed are particularly investigated to provide an extensive overview on the various ML algorithms used in the field. As a result, this paper introduces the most promising prediction methods for both long-term and short-term floods. Furthermore, the major trends in improving the quality of the flood prediction models are investigated.

Disadvantages :-

This paper suggests that the drawbacks to major ML methods in terms of accuracy, uncertainty, performance, and robustness

Long-term prediction

6.3 Building an Intelligent Hydroinformatics Integration Platform for Regional Flood

Inundation Warning Systems.[50]

Authors:- Li-Chiu Chang and Fi-John Chang

This paper first summarizes the ML methods proposed in this special issue for flood forecasts and their significant advantages. Then, it develops an intelligent hydroinformatics integration platform (IHIP) to derive a user-friendly web interface system through the state-of-the-art machine learning, visualization and system developing techniques for improving online forecast capability and flood risk management. The holistic framework of the IHIP includes five layers (data access, data integration, services, functional subsystem, and end-user application) and one database for effectively dealing with flood disasters. The IHIP provides real-time flood-related data, such as rainfall and multistep-ahead regional flood inundation maps. The interface of Google Maps fused into the IHIP significantly removes the obstacles for users to access this system, helps communities in making better-informed decisions about the occurrence of floods, and alerts communities in advance. The IHIP has been implemented in the Tainan City of Taiwan as the study case. The modular design and adaptive structure of the IHIP could be applied with similar efforts to other cities of interest for assisting the authorities in flood risk management.

6.4 River Flooding Forecasting and Anomaly Detection Based on Deep Learning[26]

Author :- Scott Miao and Wei-Hsi Hung

Pluvial floods are rare and dangerous disasters that have a small duration but a destructive impact in most countries. In recent years, the deep learning model has played a significant role in operational flood management areas such as flood forecasting and flood warnings. This paper employed a deep learning based model to predict the water level flood phenomenon of a river in Taiwan. We combine the advantages of the CNN model and the GRU model and connect the output of the CNN model to the input of the GRU model, called Conv-GRU neural network, and our experiments showed that the ConvGRU neural network could extract complex features of the river water level.



We compared the predictions of several neural network architectures commonly used today. The experimental results indicated that the Conv-GRU model outperformed the other state-of-the-art approaches. We used the Conv-GRU model for anomaly/fault detection in a time series using open data. The efficacy of this approach was demonstrated on 27 river water level station datasets. Data from Typhoon Soudelor were investigated by our model using the anomaly detection method. The experimental results showed our proposed method could detect abnormal water levels effectively.

6.5 Rainfall–runoff modelling using Long Short-Term Memory (LSTM) networks[19]

Author :- Frederik Kratzert, Daniel Klotz, Claire Brenner, Karsten Schulz, and Mathew Herrnegger

Rainfall–runoff modelling is one of the key challenges in the field of hydrology. Various approaches exist, ranging from physically based over conceptual to fully data-driven models. In this paper, we propose a novel data-driven approach, using the Long Short-Term Memory (LSTM) network, a special type of recurrent neural network. The advantage of the LSTM is its ability to learn long-term dependencies between the provided input and output of the network, which are essential for modelling storage effects in e.g. catchments with snow influence. We use 241 catchments of the freely available CAMELS data set to test our approach and also compare the results to the well-known Sacramento Soil Moisture Accounting Model (SAC-SMA) coupled with the Snow-17 snow routine. We also show the potential of the LSTM as a regional hydrological model in which one model predicts the discharge for a variety of catchments. In our last experiment, we show the possibility to transfer process understanding, learned at regional scale, to individual catchments and thereby increasing model performance when compared to a LSTM trained only on the data of single catchments. Using this approach, we were able to achieve better model performance as the SAC-SMA + Snow-17, which underlines the potential of the LSTM for hydrological modelling applications.

VII. SYSTEM REQUIREMENTS SPECIFICATION

System Requirement Specification is used for the programming contexture that are basically being for the functionality of the system can do, and also for the work behind the organization for describing and to understand the client's needs. The purpose of SRS gives the requirement to be master planed of a system or sub-system. It generally provides for the capable clients needful conditions at a particular instance of time before the work is finalize.

7.1 System Overview

A software requirements specification is a complete that describes the behavior of system to be developed. Use case techniques can be used to find the product of functional requirements or supplementary requirements. A nonfunctional requirement performs the engineering requirements, quality of standards.

7.1.1 Purpose:

The purpose of SRS document is to list the user requirements in the organized manner. It defines all the constraints and software requirements needed to understand the application and documents. The user should be able to understand the purposed system after going through the SRS documents and should be in position to incorporate some changes required.

7.2 Functional Requirement Specification The functional requirements are designed to carry out to the clients. The requirements used by the clients should be very well defined for the operation of the system. The clients understand what the services to be provided, objectives to be defined and how the system will react with particular input. The development of functional requirement leads to the specific operations, major requirements to develop project, the types of software to be tested and what does the system need to be specified. The input and output data should be taken to the use with characteristics of clients data. The specified task of functional requirement of system should understand the brief description of clients needs. The resources those are very essential to estimate the operations, costs, analysis and the information to be carried.

7.2.1 Functional Requirements

- The analysis and design of the proposed system focused on its main functional requirements. They included gathering not only key geospatial factors from institutional data sources but also those crowdsourced from individuals.
- Users are divided into two groups. Firstly, the one labeled as “User,” are the general public, community leaders, or government officers who are anticipating the event and need access to forecasts so as to prepare appropriate measure in accordance.



- The users are able to browse map data as well as relevant information geographically, such as observed and predicted precipitations. The prediction includes accumulated amount (in mm.) and probability (in percent) of precipitation, etc.
- Based on relevant factors within an area, they may query prediction of flood event, along with its likelihood and possible severity.
- The outcomes can then be verified against the flooding data, crowdsourced and reported via the Flood Mitigation System (thaiflood.org). Secondly, the other group labeled as “Admin/ Officer” are authorized agents who are operating the services.
- In addition to generic functionalities, their administrative tasks include membership management as well as updating relevant data and forecasted results.

7.3 Non Functional Requirements Non-Functional requirements are indirectly specified for the structure of project. The requirements with the specified function may carry different constraints to perform the system. The quality of the system are presides to measure the constraints of the system capabilities. The measure of Non-Functional requirements allows the clients to indirectly concentrate on the system analysis.

Qualities of Non-Functional Requirements are

- Usability: Design very simple no pre-requisite technical knowledge is required
 - Maintainability: Maintenance is very less.
 - Scalability: it is scalable for multiple devices which does not degrade the performance.
 - Portability: it can be used in anywhere on independent of platform.
- Data used in flood forecasting could be categorized into four main groups. They were
 - 1) Geospatial,
 - 2) Meteorological and Hydrological, obtained from GLOFAS,
 - 3) Hourly rainfalls prediction from TMD Big Data platform, and Crowdsourc (or volunteer) data
 - They were stored in geo-database and then processed by one of modern ML strategies.
 - Since all data involved in this study were acquired from various sources data needs to preprocess in terms of spatial resolution, interpolation was made based on their geographic coordinates.
 - The key elements were data acquisition and interchange between the system and respective sources and their intelligence via MLs.
 - Meteorological and hydrological data were acquired from the Global Flood server, called GLOFAS. Meteorological data consisted of accumulated precipitation, and the probability of precipitation at different levels, predicted daily.

The prediction is based on ECMWF (European Center for Medium-Range Weather Forecasts) model.

REQUIREMENTS

HARDWARE NECESSITIES

The most extensively watched approach of fundamentals delineated in some running system applications is a physical PC resources , everything considered known contraption , hardware entities once over is anyway significant part of the time as could be normal joined by an apparatus resemblance list , particularly if there ought to their event of working structures. A HCL notes endeavored , flawless and once an incongruent gear devices for a certain running structure. The going with set packs separates the various bits of mechanical assemble stray pieces. All personal computers running systems are made methodologies for a particular PC structure. The hugeness of leading getting ready unit is a focal structure essential for anything. Highest programming executing on x86 dealing with delineate master minding ability as copy and time speed of CPU.

HARDWARE SYSTEM CONFIGURATION :-

- ❖ Processor : Intel dual core
- ❖ Speed : 1.1 Ghz
- ❖ RAM : 4GB (min)
- ❖ Hard Disk : 50 GB



- ❖ Key Board : Standard Window Keyboard
- ❖ Mouse : Two or Three Button Mouse

SOFTWARE REQUIREMENTS

- ❖ Operating System : Windows 10, 11
- ❖ Front End : HTML, CSS
- ❖ Back End : Python
- ❖ Database Connectivity: "https://weather.visualcrossing.com"

VIII. SYSTEM ANALYSIS

The software development is generally carried out with the System Analysis and Design. The system analysis provides various detailed information of existing performance that may lead to the configuration of the new system. The problem in the existing system will be the drawbacks and to overcome this problem, the proposed system may raise with the improvement with the solutions for existing system. The process of studying the system analysis in order to identify the goals and objectives of system

8.1 Existing System

All the Existing system statistically analyzed floods based on the historical data of water levels, and seven points of water inlets and outlets. The resultant forecast was presented as point-wise flood levels, rendered on a two-dimensional map. Its shortcoming was that the analysis did not consider any geospatial data. In addition, forecasts could only be made at specific outlets, but not at those arbitrarily queried by users. Although it was developed as a web application, it did not support responsive web technology, and as such was unable to equally well satisfy user experience on all other devices apart from personal computers (PC). According to recent surveys, most flood forecasting systems relied primarily on either monitored precipitation data or those obtained from a single source. Beside the mentioned limitations, existing systems remained lacking various aspects. Existing tools were typically based on manually input and prepared data. The processes were tedious and thus prohibitive for real-time and early forecasts. Furthermore, these tools did not fully exploit more comprehensive information available in current big data platforms.

8.1.1 Disadvantages

- Monitoring facilities or communication network of one's became malfunctioned, there would be no precipitation data available for imperative analyses.
- There was also no tool (software) that can accommodate area specific forecasting well in advance.
- Existing tools were highly dependent on demanding data preparation and compilation from various sources, including Big Data

4.1 PROPOSED SYSTEM

To elevate the limitations stated above, the proposed system thus analyzed and designed a flood forecasting system that improved over the current ones. The aspects considered herein were supports of responsive web technology, automation of key processes, and availability and usability of the system. To this end, the proposed system was developed by using both meteorological and hydrological models in forecasting accumulated precipitation from data obtained from TMD big data and GLOFAS, and ML models in forecasting flood situations in given areas.

The analyses were made based on meteorological, hydrological, geospatial and crowd sourcing data. Our novel flood forecasting system based on fusing meteorological, hydrological, geospatial, as well as crowd sourcing data, and integrating them into an ML framework. These data were compiled from various big data platforms, by using online application programming interfaces (API). The forecasting mechanism was driven by a machine learning strategy. To determine the most suitable one for the task, several state-of-the-art MLs, i.e., decision tree, random forest, naïve Bayes, artificial neural networks, support vector machine

ADVANTAGES

- Forecasting was performed based on these data learned by modern ML strategies. They were decision tree, RF, Naïve Bayes, MLP and RBF, ANN, SVM
- The proposed system also enhanced user experience via responsive graphical interfaces, interoperable on different computing devices including mobiles.
- This advantage effectively encouraged greater contribution of crowd source data from the public, enriching data aggregation and hence increasing system accuracy and reliability



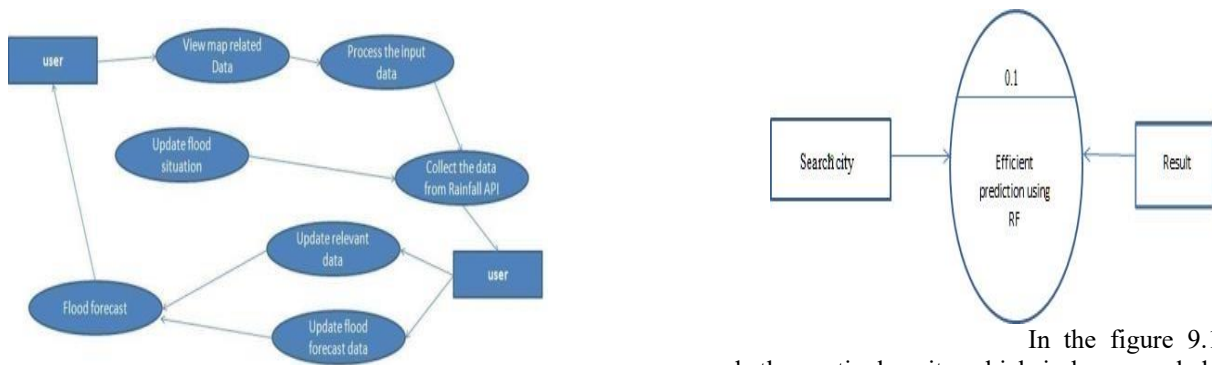
IX. SYSTEM DESIGN

This document gives the design of the overall project. Software development is the phase which is very important for the supernova of the software, which is called as design phase. The design phase should satisfy the functional and non-functional requirements for the effectiveness for satisfying all the constraints and objectives of the project. It mainly concentrates on the modules that needed for system. The design phase depends mainly on the specification of feasibility survey.

Data Flow Diagram

The information stream outline demonstrates the graphical portrayal, similar to game plans it is utilized to speak to the information through the sources of info, different sorts of information examination will be completed and the coveted yield will be produced. These parts will be utilized to demonstrate the framework and it will be displayed by to contemplate quickly regarding the info. In the framework outline the DFD will demonstrate the stream of whole parts. The stream of data will in arrangement of change utilizing this framework.

Figure 9.1,9.2: Data Flow Diagram Level 0



In the figure 9.1 , user can search the particular city which is been needed, it shows the result of the particular city through the efficient prediction by using the RF.

Figure 9.2: Data Flow Diagram Level 1 In the figure 9.2, it shows that the user types the city name and through the parameters like cloud, temperature and wind it reads the data and predicts the flood.

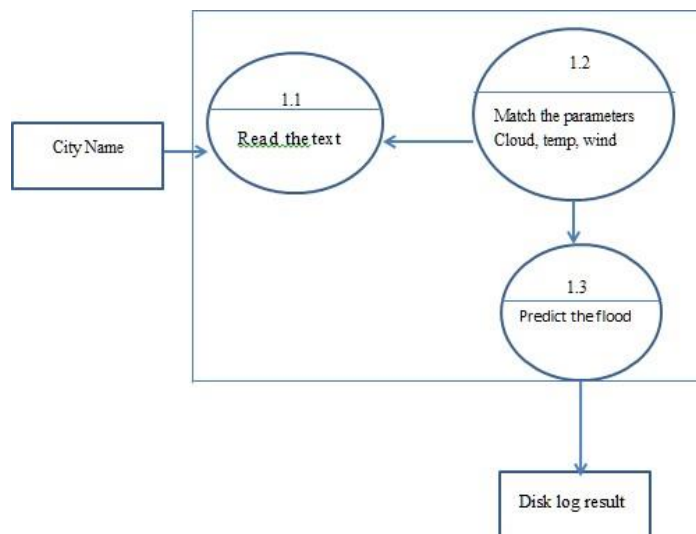
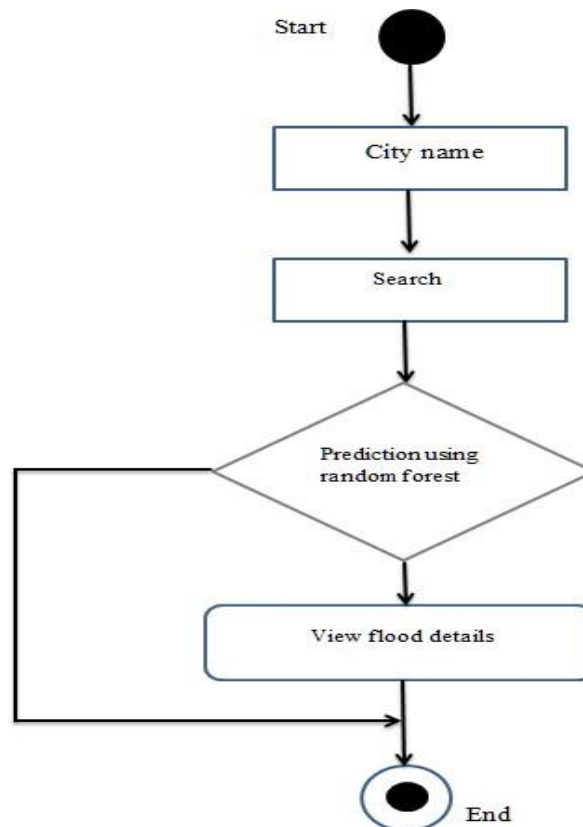


Figure 9.3: Dataflow Diagram Level



In the figure 9.3, the user can view map related data and processes the data as input, it collects the data from the rainfall API which is been updated the flood situations . The user collects the updated data and flood forecast data and predicts the flood.

9.1.1 Activity Diagram



The above figure represents the activity model of the novel flood forecasting project, it collects the data by processing of recording and transmitting the readings of an instruments and it is automated communication process from multiple data sources. The Telemetry works by using sensors at a remote source, the sensors measure physical data and electrical data which gets converted to electrical voltages and combined with timing data, they form a data system which is transmitted over wireless, wired or hybrid medium. The weather forecasting is the application of science and technology to predict the conditions of the atmosphere for a given location and time, it is a scientific estimate of future weather condition. Weather condition is the state of atmosphere at a given time expressed in terms of the most significant weather variables. The significant weather variables bring forecast differ from the place to place. Weather radar which is also called weather surveillance radar and Doppler weather radar is a type of radar used to locate precipitation, calculate its motion, and estimate its, Modern weather radars are mostly pulse-Doppler radars, capable of detecting the motion of rain droplets in addition to the intensity of the precipitation. The images are collected from satellite. These all data is been collected and modules is been transmitted.

9.1.2 Use Case Diagram A utilization case in programming designing and frameworks building is a portrayal of a framework's conduct as it reacts to a demand that starts from outside of that framework. As it were, an utilization case depicts "who" can do "what" with the framework being referred to. The utilization case method is utilized to catch a framework's behavioral necessities by specifying situation driven strings through the useful prerequisites.

Utilize cases portray the connection between at least one performing artists (an on- screen character that is the initiator of the communication might be alluded to as the 'essential performer') and the framework itself, spoken to as a succession of basic strides. Performing artists are something or somebody which exists outside the framework ('black box') under review, and that partake in a grouping of exercises in a discourse with the framework to accomplish some objective. On-screen characters might be end clients, different frameworks, or equipment gadgets. Each utilization case is a total arrangement of occasions, depicted from the perspective of the on-screen character.

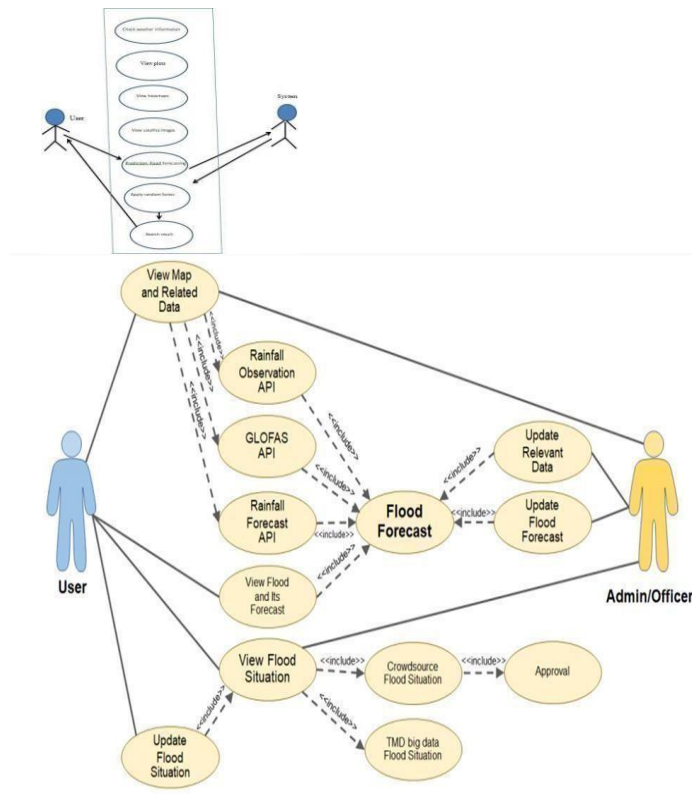


Flood forecasting was rendered based on learning of thematic data by an ML method. The resultant forecast was displayed on a web browser. With this platform, users may access this information from various devices. The design phase should satisfy the functional and non-functional requirements for the effectiveness for satisfying all the constraints and objectives of the project. It mainly concentrates on the modules that needed for system. The design phase depends mainly on the specification of feasibility survey.

As shown in the above diagram, the user can view the map and data which is related ,he can view the weather forecast and floods occurring situation. The data is been related by considering the parameters like rainfall, GLOFAS and weather forecast, through considering these parameters flood is been predicted and it updates the relevant data and flood forecasting.

X. IMPLEMENTATION

1 Algorithm Prediction Module



- Step 1: Start prediction model
- Step 2: Users should select a state required to enter the state name
- Step 3: It checks the all-weather and parameters
- Step 4: It display safe/unsafe for the flood and 2 also displays temperature, humidity, cloud, parameters
- Step 5: Stop

User Module

- Step 1: Start user model
- Step 2: view map related data

View flood statuses

- Step 4: flood forecasting information
- Step 5: end if condition
- Step 6: Stop



Output : {View map related data}

Begin

Step 1: Add city name

Step 2: connected with map data

Step 3: checks the weather data

Step 4: Shows the weather related data

End

Step 1: Read city name

Step 2: Preprocess query

Step 3: Checks the city name in map

Step 4: Apply RF

Step 5: Find flood related data

Step 6: returns result

End

Floods are one of the most dangerous and frequent natural disasters in the world. Every year, thousands of people die from the floods and millions of people lose their homes and livelihoods. In many states around India and China's Yellow River Valley, some of the world's worst floods have killed thousands, if not millions of people.

MODULES

There are three modules in the Novel Approached Flood Forecasting System

Using ML And BD. They are:-

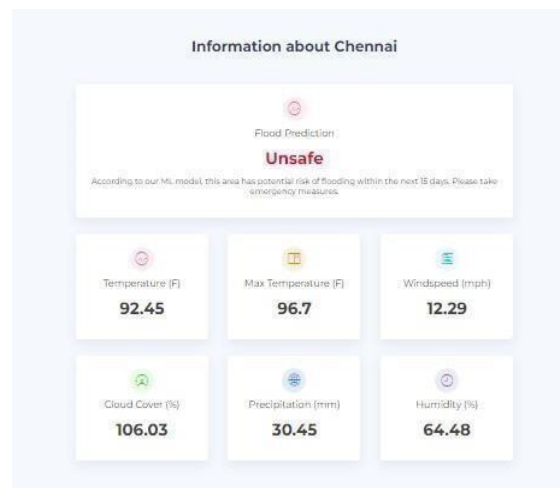
- PLOTS
- INTERACTIVE HEATMAPS

XI. TESTING

PLOTS

The visualizations show flood predictions, damage predictions, and heavy rainfall predictions across India, taking in factors such as precipitation, wind speed, humidity, temperature, cloud cover, as well as previous data history. We can view plots showing flood prediction, precipitation analysis, and damage analysis.

Flood prediction: The plot below shows our ML powered prediction of where a flood is going to occur, marked by red dots



Precipitation: The plot below shows the current precipitation data across the nation. The larger bubbles indicate more precipitation.



Interactive Heatmaps

The heatmaps show flood predictions, damage predictions, and heavy rainfall predictions across India, taking in factors such as precipitation, wind speed, humidity, temperature, cloud cover, as well as previous data history.

Damage Analysis

The plot below shows cost and damage analysis, based on the flood risk prediction. The colorscale of the heatmap indicates the extent of predicted monetary damage, measured in INR.

Flood Prediction

The plot below shows our ML powered prediction of where a flood is most likely to occur given the current environmental factors, marked primarily by the darker red spots.

The primary objective of testing is to correct the bugs, issues or blunders. To recognize mistakes the test engineers need to test every individual segments of the venture arrange module. Every module is tried for the better execution and by checking the modules the mistakes are recognized. It confirms that the frameworks achieved its prerequisites are definitely not. By examining the every modules and mistakes are pester out of framework to get particular yield

Dissimilar to every module gives the required yield, the assurance of test designer prompts the rightness of framework projects. The last module stage is intended to maintain a strategic distance from the disappointments and to expel deficiencies. So it's ideal to give the testing stage to the advancement of the venture.

11.1 Testing methods

The testing methods is one of the essential stages in the framework testing. It serves to the general population who are working outside the testing field. The correct arranging makes the item to build up to customer. The test arrange gives the documentation of the exercises performed for programming testing and its serves to approve the yield work.

11.1.1 Unit Testing

Unit testing as the name portrays that the testing procedure is completed with the testing where every individual models are tried in a steady progression. The operation to perform unit testing is to figure out where every module testing is approve or not.

The investigation of testing gives the fruitful result and to perform correct report determination. The capacity of unit test additionally upgrades the level of testing before the reconciliation procedure. By testing every module the blunders are recognized in before stages and this may prompt the yearning yield of the projects. Unit testing isolates the every individual piece of modules and redresses whether the module is executed or not. The essential execution is to give an end-clients to enhancing the application programming, business handle and the level of framework setup.

11.1.1 Integration Test

It is test where every one of the exhibitions are planned with the product testing procedure and individual set programming's are coordinated to perform in a gathering to run the one program. The fulfillment of this testing leads just when exhibitions of every necessities, programming modules and programming design. The yield execution makes when all reconciliation test modules are determined to play out the testing procedure with craving input. At long last mix testing furnishes end-client with accuracy of the yield with determined programming testing

11.1.2 System Test

Framework Testing is one of the testing procedure where the fruition of testing stage is for the most part relies on upon System. Framework testing gives the spine support to all the testing stage on the grounds that once the consummation of all the testing procedure the framework testing plays out the Hardware and Software Requirement Specifications and the Software situated examination of framework. This depends on the desire of end-client, where it ought to fulfilled to get crave result.

11.1.3 Functional testing

This is a type of black box testing that is based on the specifications of the software that is to be tested the application is tested by providing input and then the results are examined that need to conform to the functionality it was intended. Functional testing of a software is conducted on a complete, integrated system to evaluate the system's compliance with its specified requirements.



11.1.4 Validation testing

Validation testing determining the system complies with the requirements and performs functions for which it is intended and the organization's goals and user needs. Validation is done at the end of the development process and takes place after verification are completed. Performed after a work product is produced against established criteria ensuring that the product integrates correctly into the environment. Determination of correctness of the final software product by a development project with respect to the user needs and requirements.

XII. PERFORMANCE ANALYSIS

To ensure maximum versatility and most extensive coverage of crowdsourcing, the proposed methodology adopted responsive web design paradigm in developing the website. The web application once deployed was able to support various devices, ranging from personal and portable computers to mobile phones and tablet computers, with varying screen sizes and resolutions.

By using the Bootstrap framework, the rendered CSS and JavaScript automatically aligned and adjusted the layout of components and controls to maintain uniform appearance and hence satisfying user experience (UX). Furthermore, the development cost was minimized as focuses were placed on core functionalities and customizable contents, instead of variations of frontend interfaces and layouts across platforms.

Meteorological and hydrological data that were acquired from GLOFAS are illustrated in Figure 13. In this figure, accumulated precipitation in each area was color coded. Specifically, blue, light green and dark red, represents low (10 – 25 mm), moderate (25 – 100 mm), and high (> 100 mm), respectively. Areas with neutral one indicated those with no precipitation. Likewise, the data would also be used in forecasting.

Another crucial information was the probability of precipitation. In this study, the respective probability in each area was color coded in three levels, i.e., 50 mm, 150 mm, and 300 mm. Examples of these levels are illustrated in Figures 14 to 16. In each figure, pixel intensities indicate the likelihood of precipitation (i.e., 0 to 1) at that location being of the respective levels. In Figure 14, for instance, light and dark green indicate low and high probability of 50 mm precipitation in a given location. Similarly, Figures 15 and 16 display the probability of precipitation of 150 mm and 300 in blue and red, respectively.

Yellow, orange, and red pins represented flood levels of less than 20 cm, 20 – 50 cm, and greater than 50 cm, respectively. Without crowdsourcing data, there would be no better means of acquiring rainfall duration and its intensity, as well as drainage problem, which were all crucial

Flood forecasting was rendered based on learning of thematic data by an ML method. The resultant forecast was displayed on a web browser. With this platform, users may access this information from various devices. An example of forecasted flood is shown snapshots. The strength of its impacts was determined based its level, and accordingly coded in different colors. Green pin indicate that the area was not at all affected by flood. On the contrary, yellow, orange, and red pin represents flooded areas, whose levels were less than 20 cm, 20 – 49 cm, and greater than 50 cm, respectively. Figure 19 shows an example when no flood incident was anticipated.

The performance of these MLs is listed in Tables 5 and 6. It is evident that MLP ANN, SVM, and RF were placed in top three ranks, in terms of classification accuracies (i.e., 97.83%, 96.67%, and 96.67%) and Kappa coefficients (i.e., 0.89, 0.84, and 0.84). Statistically, Kappa values of greater than 0.8 indicate highly accurate forecast, while those between 0.4 – 0.8 were moderate performers. Naïve Bayes, and fuzzy logic fell in the latter category. These trends were similarly exhibited in MAE and RMSE, the closer to 0, the more accurate the forecast. Taken into account the results of, not only flooded regions but also those unaffected by flood, the balance between relevant and irrelevant predicted samples had to be considered.

Comparison of previous algorithm and updated algorithm Some data considered in this study, such as GLOFAS, were not of intrinsically high spatial resolution, However, they were accurate. The accumulated precipitations and their probability at different levels, for instance, corresponded to the actual event. Their API system was also reliable. These characteristics were updated by the proposed system. Their resolutions shortcomings were remedied by incorporating other more detailed layers as well as crowdsource factors into the ML framework. Possible improvements for this issue include involving the Internet of Things (IoT) in measuring actual meteorological data with preferred coverage actual meteorological data with preferred coverage.

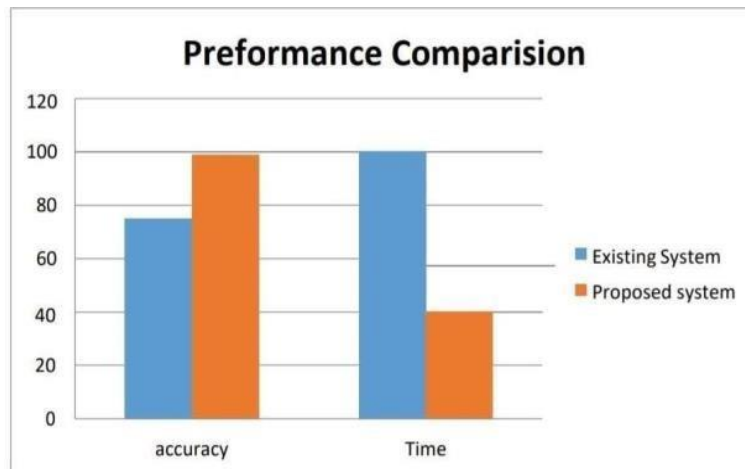


Figure 8.1: Performance Comparison

XIII. CONCLUSION

The proposed methodology is a novel distributed flood forecasting system, based on integrating meteorological, hydrological, geospatial, and crowdsource data. Big data made available by prominent agencies were acquired by means of various crossplatform APIs. Forecasting was performed based on these data learned by modern ML strategies. They were decision tree, RF, Naïve Bayes, MLP and RBF ANN, SVM, and fuzzy logics. Evaluation results on studied areas indicated that the system could forecasted flood events highly accurately.

This system also enhanced user experience via responsive graphical interfaces, interoperable on different computing devices including mobiles. This advantage effectively encouraged greater In prospects, the system can be readily employed in existing floods management schemes, e.g., those led by government agencies or non-profit organizations. Moreover, thanks to distributed architecture, the system can reach wider public, and therefore serves as an effective means of communicating with them (and especially the flood victims), regarding current status and development of the disaster. In prospects, the system can be readily employed in existing floods management schemes, e.g., those led by government agencies or non-profit organizations. Moreover, thanks to distributed architecture, the system can reach wider public, and therefore serves as an effective means of communicating with them (and especially the flood victims), regarding current status and development of the disaster.

XIV. FUTURE ENHANCEMENT

Future improvements of the system include initial flood representation and its extent being adapted to the current location of the device, so that they can be instantly made aware of by its user. Moreover, flooded location pinned by an icon may be augmented with color-coded regions, so that the conditions (e.g., levels and extents) of affected areas may be better comprehended.

REFERENCES

- [1]. Azhoni, A., Holman, I., & Jude, S. (2017a). Contextual and interdependent causes of climate change adaptation barriers: Insights from water management institutions in Himachal Pradesh, India. *Science of the Total Environment*, **576**, 817– 828.
- [2]. Baan, P., & Klijn, F. (2004). Flood risk perception and implications for flood risk management in the Netherlands. *International Journal of River Basin Management*, **2**, 113– 122.
- [3]. Alfieri, L., Salamon, P., Bianchi, A., Neal, J., Bates, P. D., and Feyen, L.: Advances in pan-European flood hazard mapping, *Hydrol. Process.*, **28**, 4928– 4937, <https://doi.org/10.1002/hyp.9947>, 2014b.
- [4]. E. Mayoraz and E. Alpaydin, “Support vector machines for multiclass classification,” in *Proc. Int. Work-Confer, Artif. Neural Netw.* Berlin, Germany: Springer, 1999, pp. 833–842. Ghosh, S. Karmakar, A. Saha, M. P. Mohanty, S. Ali, S. K. Raju, and P. L. N. Murty, “Development of India’s first integrated expert urban flood forecasting system for Chennai,” *Current Sci.*, vol. 117, no. 5, pp. 741–745, 2019.
- [5]. Q. Ran, W. Fu, Y. Liu, T. Li, K. Shi, and B. Sivakumar, “Evaluation of quantitative precipitation predictions by ECMWF, CMA, and UKMO for flood forecasting: Application to two Basins in China,” *Natural Hazards Rev.*, vol.



- 19, no. 2, May 2018, Art. no. 05018003.
- [6]. S. Ghosh, S. Karmakar, A. Saha, M. P. Mohanty, S. Ali, S. K. Raju, and P. L. N. Murty, "Development of India's first integrated expert urban flood forecasting system for Chennai," *Current Sci.*, vol. 117, no. 5, pp. 741–745, 2019.
- [7]. A. Mosavi, P. Ozturk, and K.-W. Chau, "Flood prediction using machine learning models: Literature review," *Water*, vol. 10, no. 11, p. 1536, Oct. 2018.
- [8]. L.-C. Chang, F.-J. Chang, S.-N. Yang, I.-F. Kao, Y.-Y. Ku, C.-L. Kuo, and I. Amin, "Building an intelligent hydroinformatics integration platform for regional flood inundation warning systems," *Water*, vol. 11, no. 1, p. 9, Dec. 2018. N. A. Sulaiman, N. F. A. Aziz,
- N. M. Tarmizi, A. M. Samad, and W. Z. W. Jaafar, "Integration of geographic information system (GIS) and hydraulic modelling to simulate floodplain inundation level for Bandar Segamat," in *Proc. IEEE 5th Control Syst. Graduate Res. Colloq.*, Aug. 2014, pp. 114–119.
- [9]. Sarawut Jamrusri, Yuji Toda. Simulating past severe flood events to evaluate the effectiveness of nonstructural flood countermeasures in the upper chao phraya river basin, Thailand. *Journal of Hydrology Regional Studies*, 2017, Vol.10, p.82.
- [10]. Scott Miao and Wei-Hsi Hung River Flooding Forecasting and Anomaly Detection Based on DeepLearning30 Oct 2020, 198384-198402, Vol 8
- [11]. P. Vijayalakshmi, M. Vandhana, K. Vasanthapriya, P.B. Thushikka Prediction of Flood by Rainfall using MLP Classifier of Neural Network Model 2021 ICCE51350.2021.9489161
- [12]. Thorndahl, S.; Einfalt, T.; Willems, P.; Ellerbæk Nielsen, J.; Ten Veldhuis, M.C.; Arnbjerg-Nielsen, K.; Rasmussen, M.R.; Molnar, P. Weather radar rainfall data in urban hydrology. *Hydrol. Earth Syst. Sci.* **2017**, *21*, 1359–1380
- [13]. Cecinati, F.; de Niet, A.; Sawicka, K.; Rico-Ramirez, M.A. Optimal Temporal Resolution of Rainfall for Urban Applications and Uncertainty Propagation. *Water* **2017**, *9*, 762.
- [14]. Flood inundation modelling: A review of methods, recent advances and uncertainty analysis, J. Teng, A. Jakeman, J. Vaze, B. Croke, D. Dutta, S. Kim, *Business Environ. Model. Softw.* 2017
- [15]. Weili Duan, Bin He, Daniel Nover, et al. floods and associated socio economic damages in China over the latest century. *Natural Hazards*. 2016, Vol.82, No.1, p.401.
- [16]. S.Zhang, J.Zhang, Y. Liu, and Y. Liu, "A mathematical spatial interpolation method for The estimation of convective rainfall distribution over small watersheds," *Environ. Eng. Res.*, vol. 21, no. 3, pp.226-232, Sep.2016.
- [17]. J. P. De Albuquerque, M. Eckle, B. Herfort, and A. Zipf, "Crowdsourcing geographic Information for disaster management and improving urban resilience: An overview of recent developments and lessons learned," in *European Handbook of Crowdsourced Geographic Information*. London, U.K.: Ubiquity Press, Aug. 2016, pp. 309–321.
- [18]. Ford, J. D., Tilleard, S. E., Berrang-Ford, L., Araos, M., Biesbroek, R., Lesnikowski, A. C., Bizikova, L. (2016). Opinion: Big data has big potential for applications to climate change adaptation. *Proceedings of the National Academy of Sciences*, **113**, 10729–10732.
- [19]. Al-Zahrani M, Al-Areeq A, Sharif HO. 2016. Estimating urban flooding potential near the outlet of an arid catchment in Saudi Arabia. *Geomat Nat Hazards Risk*. 1:1-17.
- [20]. Technical review of large-scale hydrological models for implementation in operational Flood forecasting schemes on continental level, Ann Kauffeldt, F. Wetterhall, F. Pappenberger, P. Salamon, J. Thielen *Environmental Science, Environ. Model. Softw.* 2016
- [21]. Xu, H.; Xu, C. Y.; Saelthun, N. R.; Xu, Y.; Zhou, B.; Chen, H. Entropy theory based multi criteria resampling of rain gauge networks for hydrological modelling—a case study of humid area in southern China. *J. Hydrol.* 2015, 525, 138–151. 2015, 525, 138–151.
- [22]. G. Schimak, D. Havlik, and J. Pielorz, "Crowdsourcing in crisis and disaster management Challenges and considerations," in *Environmental Software Systems. Infrastructures, Services and Applications*. Cham, Switzerland: Springer, 2015, pp. 5670.
- [23]. T. Simon, A. Goldberg, and B. Adini, "Socializing in emergencies— A review of the use of social media in emergency situations," *Int. J. Inf. Manage.*, vol. 35, no. 5, pp. 609–619, Oct. 2015
- [24]. L. Alfieri, J. Thielen, A European precipitation index for extreme rain-storm and flash flood early warning. *Meteorol. Appl.* **22**(1), 3–13 (2015).
- [25]. Development of a spatial decision support system for flood risk management in Brazil that combines volunteered geographic information with wireless sensor networks **April 2015**, *Computers & Geoscience* **80:84-94**
- [26]. Rico-Ramirez, M.A.; Liguori, S.; Schellart, A.N.A. Quantifying radar-rainfall uncertainties in urban drainage flow modelling. *J. Hydrol.* **2015**, *528*, 17–28
- [27]. On the Use of Global Flood Forecasts and Satellite-Derived Inundation Maps for Flood Monitoring in Data-Sparse Regions, B. Romero, F. Hirpa, +4 authors T. Groeve, *Environmental Science Remote. Sens.* 2015
- [28]. N. A. Sulaiman, N. F. A. Aziz, N. M. Tarmizi, A. M. Samad, and W. Z. W. Jaafar, "Integration of geographic



- information system (GIS) and hydraulic modelling to simulate floodplain inundation level for Bandar Segamat,” in Proc. IEEE 5th Control Syst. Graduate Res. Colloq., Aug. 2014, pp. 114–119.
- [29]. S. Puttinaovarat, P. Horkaew, and K. Khaimook, “Conguring ANN for inundation areas identification based on relevant thematic layers,” ECTI Trans. Comput. Inf. Technol.(ECT I- CIT), vol. 8, no. 1, pp. 56–66, May 2014.
- [30]. L.S. Campbell, W.J. Steenburgh, Finescale orographic precipitation variability and gap-filling radar potential in little Cottonwood Canyon, Utah. *Weather Forecast.* **29**,2014
- [31]. Al-Nammari F., Alzaghal M. Towards local disaster risk reduction in developing countries:Challenges from Jordan. *Int. J.Disaster Risk Reduct.* 2015;**12**:34–41. doi: 10.1016/j.ijdr.2014.11.005.
- [33] J.García-Pintado, J.C. Neal, D.C. Mason, S.L. Dance, P.D. Bates Scheduling satellitebased sar acquisition for sequential assimilation of water level observations into flood modelling *J. Hydrol.*, 495 (2013), pp. 252-266
- [32]. Siriluk Mongkonkerd, Siyanee Hirunsalee, Hidehiko Kanegae, et al. Comparison of Direct Monetary Flood Damages in 2011 to Pillar House and Non-pillar House in Ayutthaya, Thailand. *Procedia Environmental Sciences.* 2013, Vol.17, p.327.
- [33]. Shunji Kotsuki, Kenji Tanaka. Impacts of Mid-Rainy Season Rainfall on Runoff into Chao Phraya River, Thailand. *Journal of Disaster Research.* 2013, Vol.8, No.3, p.397.
- [34]. Abaza, M., *et al.*, 2014. A comparison of the Canadian global and regional meteorological ensemble prediction systems for short-term hydrological forecasting (vol 141, pg 3462, 2013). *Monthly Weather Review*, 142, 2561–2562. doi:10.1175/mwr-d-14-00018.1.
- [35]. Ly, S.; Charles, C.; Degré, A. Different methods for spatial interpolation of rainfall datafor operational hydrology and hydrological modeling at watershed scale: A review. *Biotechnol. Agron. Soc. Environ.* 2013, 17, 392-406. 20, 35
- [36]. Xu, H.; Xu, C.Y.; Chen, H.; Zhang, Z.; Li, L. Assessing the influence of rain gauge density distribution on hydrological model performance in a humid region of China. *J. Hydrol.* 2013, 505, 1–12.
- [37]. J.García-Pintado, J.C. Neal, D.C. Mason, S.L. Dance, P.D. Bates Scheduling satellite-basar acquisition for sequential assimilation of water level observations into floodmodelling *J. Hydrol.*, 495 (2013), pp. 252-26.
- [38]. Rogstadius, J., Vukovic, M., Teixeira, C. A., Kostakos, V., Karapanos, E., and Laredo, J. A.: CrisisTracker: crowdsourced social media curation for disaster awareness, *IBM J. Res. Dev.*, 57, 4:1– 4:13, doi:10.1147/JRD.2013.2260692, 2013.
- [39]. Pappenberger, F., *et al.*, 2013. Visualizing probabilistic flood forecast information: expert preferences and perceptions of best practice in uncertainty communication. *Hydrological Processes*, 27 (1), 132– 146. doi:10.1002/hyp.9253.
- [40]. Peleg, N.; Ben-Asher, M.; Morin, E. Radar subpixel-scale rainfall variability and uncertainty: Lessons learned from observations of a dense rain-gauge network. *Hydrol Earth Syst. Sci.* **2013**, 17, 2195–2208
- [41]. Thorndahl, S.; Einfalt, T.; Willems, P.; Ellerbæk Nielsen, J.; Ten Veldhuis, M.C.; Arnbjerg-Nielsen, K.; Rasmussen, M.R.; Molnar, P. Weather radar rainfall data in urbanhydrology. *Hydrol. Earth Syst. Sci.* **2017**, 21, 1359–1380.
- [42]. Adam, N., Shafiq, B., and Staffin, R. (2012). Spatial computing and social media in the context of disaster management. *Intelligent Systems, IEEE*, 27(6), 90 –96.
- [43]. C. Corral, D. Velasco, D. Forcadell, D. Sempere-Torres, Advances in radar-based flood Warning systems. The EHIMI system and the experience in the Besòs flash-flood pilot basin , in *Flood Risk Management: Research and Practice*, ed. by P. Samuels, S. Huntington, W. Allsop, J. Harrop (Taylor & Francis, London, 2009), pp. 1295–1303
- [44]. Dawson R.J., Speight L., Hall J.W., Djordjevic S., Savic D., Leandro J. Attribution of flood risk in urban areas. *J. Hydroinf.* 2008;**10**:275–288. doi: 10.2166/hydro.2008.054
- [45]. Chaplot, V.; Saleh, A.; Jaynes, D. Effect of the accuracy of spatial rainfall information on the modeling of water, sediment, and NO₃-N loads at the watershed level. *J. Hydrol.* 2005, 312, 223–234
- [46]. Supattra Puttinaovarat, Paramate Horkaew Flood Forecasting System Based on Integrated Big and Crowdsourced Data Using Machine Learning Techniques Jan 2020, 10.1109
- [47]. Amir Mosavi , Pinar Ozturk and Kwok-wing Chau Flood Prediction Using Machine Learning Models: Literature Review (2021) CSDE5383.2021.9718497
- [48]. Li-Chiu Chang and Fi-John Chang Building an Intelligent Hydro informatics Integration Platform for Regional Flood Inundation Warning Systems 21 Dec 2018, 9, 10.3390
- [49]. Gokmen Tayfur , Vijay P. Singh, Tommaso Moramarco and Silvia Barbeta Flood Hydrograph Prediction Using Machine Learning Methods 24 July 2018, 10(8), 968