Journal of Computation Physics and Earth Science
ISSN: 2776-2521 (online)
Volume 5, Number 1, April 2025, Page 80 - 88
https://journal.physan.org/index.php/jocpes/index

Flood Early Warning Monitoring System Using KNN Methods In Bangkalan District

Iqbal Fariansyah Ridwan[1]
[1]Undergraduate Program Aplied Instrumentation Meteorology Climatology Geophysics (STMKG)

Article Info

Article history:

Received March 12, 2025
Revised March 17, 2025
Accepted March 18, 2025

Keywords:

Flood
Early warning system
Prediction
KNN
Algorithm.

A B S T R A C T

Bangkalan Regency faces serious challenges due to flood disasters that periodically threaten the safety and welfare of the community. The flood phenomenon in this area is caused by a complex combination of geographic and climate factors, including increased extreme rainfall and the dynamics of sea level rise. Annual floods not only cause infrastructure damage but also threaten the livelihoods and lives of residents living around the river flow. This study aims to develop an innovative flood early warning system using the K-Nearest Neighbors (KNN) method to predict potential disasters before floods occur. By using water flow data analysis and machine learning algorithms, this system is designed to provide accurate and timely early estimates. The main advantage of this study is its ability to proactively mitigate disaster risks using modern computer technology. The study produced a prototype of a flood detection and simulation system that can help local governments, related agencies, and the Bangkalan community in taking preventive and mitigation actions at an earlier stage. Therefore, this system is expected to make a significant contribution to reducing the impact of disasters and protecting the lives of Bangkalan Regency residents.

Corresponden Author:

Iqbal Fariansyah Ridwan,
Undergraduate Program Aplied Instrumentation Meteorology Climatology Geophysics (STMKG)
Tangerang, Indonesia
Email: iqbalfariansyahridwan@gmail.com

1. INTRODUCTION

One of the natural disaster problems that often occurs in several regions in Indonesia is flooding, according to data from the Indonesian Disaster Information Data (DIBI) the number of flood disasters that occurred in Indonesia was 751 out of 2,313 natural disasters. Then data from the National Flood Management Agency (BNPB) stated that there were 7,812 natural disasters from 21,368 total natural disasters, this places flooding as number one in the number of natural disasters [1]. Floods cause many losses including material and immaterial, the impacts caused after the flood also make many flood victims remain vigilant against subsequent floods. Material losses experienced by victims can be in the form of damaged goods, disease outbreaks, damaged plants, and difficulty in getting clean water. One of the causes of flooding in Indonesia that most often results in flooding is high rainfall, if an area has high rainfall and continues for a long period of time, then there is a high risk of flooding. Bangkalan is one of the regencies in East Java Province, located at the western tip of Madura Island, bordering Sampang Regency to the east, Gresik City to the north, Surabaya City to the south and Madura Strait to the west. There are several ways to overcome flooding, such as managing river flows, not littering, reforesting forests and creating a monitoring and warning system in flood-prone areas.

An early warning system is a series of systems that will notify of a natural disaster. Early warning of natural disasters so that the community can respond when a natural disaster is about to occur. Delays in handling natural disasters can result in great losses to the community [2]. The early warning system gives potential victims time to take the necessary actions when a natural disaster occurs.

Researchers took a river sample, namely the Kali Tunjung River because the map shows that flood-prone areas are located around the Kali Tunjung River basin. In this study, an Arduino-based flood detection system has been created that utilizes water level and water flow sensors combined with the k-nearest neighbor (KNN) algorithm, Arduino is a technology in the form of a circuit board that contains a microcontroller that can be programmed via a computer as needed, KNN is used in intelligent systems as an algorithm that determines the class of an event that occurs [3]. So that in this study, Arduino and the KNN algorithm are combined to be able to predict the arrival of floods quickly and accurately so that potential flood victims can take appropriate action before the flood comes.

2. LITERATURE

2.1. Flood

Explain the chronology of the research, including research design, research procedures (in the form of algorithms, Pseudocode or others), testing methods and data acquisition [3]. The description of the course of research must be supported by references, so that the explanation is scientifically acceptable. Tables and Figures are presented in the centre, as shown below and cited in the manuscript.

a. Rainfall

   In the rainy season, high rainfall will cause flooding in rivers and if it exceeds the river bank, flooding or inundation will occur [5].

b. Physiographic effect

   Physiography or physical geography of rivers such as the shape and slope of the River Basin Area (DPS), river slope, hydraulic geometry (cross-section shape such as width, depth, longitudinal section, river bed material), river location.

c. Erosion and Sedimentation

   Erosion in the DPS affects the river's storage capacity, because the eroded soil in the DPS when carried by rainwater into the river will settle and cause sedimentation. Sedimentation will reduce the capacity of the river and when there is a flow that exceeds the river's capacity, it can cause flooding River capacity.

d. River capacity

   The reduction in flood flow capacity in rivers is caused by sedimentation originating from excessive erosion of river beds and riverbanks, due to the absence of vegetation cover.

e. Effect of high tide

   Sea water slows down the flow of rivers to the sea. During floods along with high tides, the height of the inundation/flood becomes higher because of backwater.

The effects of flooding are many, namely spreading germs, loss of property, damaged facilities and infrastructure and difficulty in obtaining clean water [6]. There are several ways to overcome flooding, namely not littering, reforestation, dredging rivers and early warning systems for flooding [7].

In Bangkalan Regency, there are several main factors that Bangkalan Regency often experiences waterlogging which causes Bangkalan to be submerged in water, including Bangkalan Regency has a topography with a land slope of 0-2 º (flat) and a land slope of 3-15 º (wavy) with a stretch of lowland at an altitude of between 2-100 meters above sea level [8]. From this condition, Bangkalan Regency often experiences waterlogging during the rainy season, especially when it rains together with the rising tide. In general, the topography of Bangkalan Regency is divided into two parts, namely:

- The western part towards the north is relatively low close to the coastal coast.
- The eastern part has hilly topography such as in Geger Sub-district.

According to BPBD Bangkalan Regency, the causes of flooding in Bangkalan Regency are as follows, namely:

No.	Parameter	Indicator	Data Source	Monitoring Month
1.	Disaster prone areas	Potential for flooding > 80% of 18 sub-districts and 281 villages	BPBD	December - January
2.	Rainfall	>700 mm for several consecutive times	BMKG Maritime	December – January
3.	Sea level rise	The west monsoon is characterized by wind speeds reaching >25 knots in the Madura Strait waters and above average sea wave heights, namely reaching 2 - 2.5 m.	BMKG Maritime	December – January - February
4.	Other factors that are considered dominant	The emergence of anomalous events	BPBD - BMKG	December – January - February

Fig. 1 Causes of Flooding in Bangkalan Regency

Then followed by rainfall data for East Java Regency, namely:

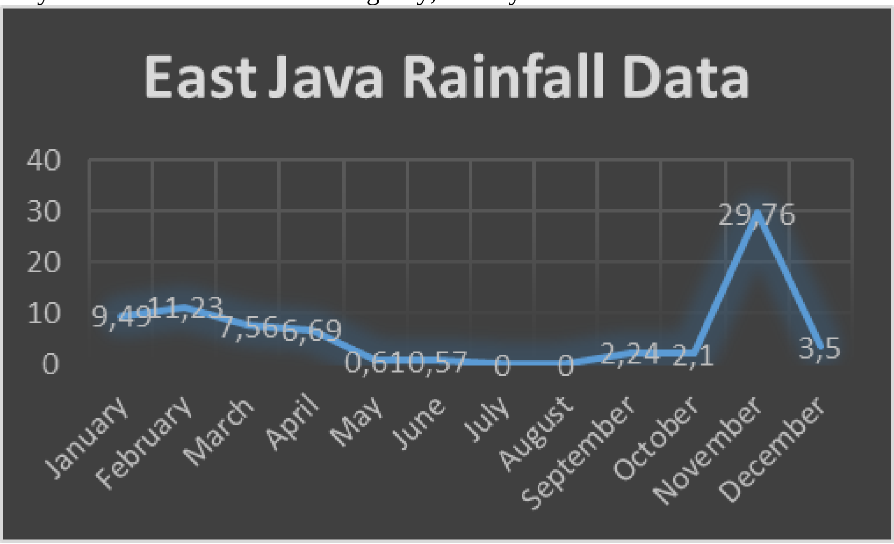

Fig. 2 Rainfall in East Java province in 2023

Then followed by rainfall data for Bangkalan Regency, namely:

Fig. 3 Rainfall in Bangkalan District in 2023

It can be concluded that the highest flooding possibility in Bangkalan city is in December, January, and February. Then followed by data on sea level rise in Bangkalan city stating that the sea level rise in the Madura Strait is between 5 - 10 mm / year. There are two potential flash floods in Bangkalan Regency, if simultaneously there is an increase in rainfall capacity above the average of 600 / mm, the highest tide or full moon. Then the BPBD of Bangkalan Regency also stated that there are ten characteristics of flood threats, including the following:

- Influence of Rainfall
- Changes in land use in River Basin Areas (DAS)
- Influence of Tides and Global Warming
- Influence of Surface Elevation
- Drainage System
- Community Habits of Throwing Away Garbage
- Pressure on Land Use
- Sedimentation
- Flood Disaster Control System
- Impact of Rising Sea Levels

2.2. Flood Detection System

Flood detection system is a system used to detect the arrival of floods. This system is usually combined with an early warning system so that potential victims can take the action before a flood disaster occurs so that they can minimize losses caused by flooding [9]. Flood detection systems are installed on river flows or places that are inundated by water so that when there is an overflow of water, it can be known and can be acted upon in accordance with applicable operations.

2.3. Sensor

Sensors are devices that can detect and respond to several types of input from the physical environment. Specific inputs can be light, heat, motion, humidity, pressure, movement and infrared [10]. There are many types of sensors such as ping sensors, pressure sensors, rotational speed sensors, temperature sensors, infrared sensors, smoke sensors and humidity sensors. Sensors are very useful as a substitute for human senses in carrying out their functions, data sent from sensors can be processed according to the needs needed, for example when you want to detect a fire, sensors related to fires are needed such as temperature sensors, gas sensors and light sensors.

In Arduino, sensors are used as tools to obtain data around the desired processing. In this simulation, two types of sensors are used, namely water level and water flow sensors [11]. five cm in length. The water level sensor uses analog signals on the wemos to communicate data, the analog value of this sensor is 0-1024 where the higher the water that has been there, the smaller the resolution value of this sensor.

a. Sensor Water Flow

Water Flow Sensor is a sensor that functions to capture the speed of water flow, this sensor is shaped like a water tap. This sensor works on a digital signal, where each water flow value will be calculated through the interrupt function on the wemos. The working principle of this flow meter sensor is to measure water flow by calculating the rotation of the wheel in the water flow if there is water flowing through it, a motor is embedded in the wheel that has a magnet and when it rotates it will produce a magnetic field based on the hall effect principle. From the event there is a magnetic field and the absence of a magnetic field that is repeated will produce an output in the form of a wave, this signal is used to calculate the speed of water flow.

b. Sensor Water Level

Water level sensor is a sensor that functions to capture the value of the height of water in a container. The shape of this sensor is small, which is five cm in length. The water level sensor uses an analog signal on the wemos to communicate data, the analog value of this sensor is 0-1024 where the higher the water that has been there, the resolution of the value of this sensor is getting smaller.

2.4. K-Nearest Neighbour

KNN is one of the machine learning algorithms that works by finding a number of k patterns closest to the input pattern and then determining the class based on the most patterns among the k patterns [12]. The KNN training process produces a k value that provides the highest accuracy in generalizing future data. This k value training process is useful for seeing the k values and their results until the most optimal k value is produced.

KNN is widely used for data mining applications, statistical pattern recognition, image processing and many other things. The purpose of this algorithm is to classify new objects based on attributes and training samples. KNN is commonly used for predictions, predictions are used for the value of future results. The accuracy of the KNN algorithm is greatly influenced by the presence or absence of irrelevant features, or if the weight of the feature is not equivalent to its relevance to the classification.

The KNN algorithm has several advantages, namely it is not too affected by training data that has a lot of noise and is effective if the training data is large [13]. Meanwhile, the weakness of KNN is that KNN needs to determine the value of the K parameter (the number of nearest neighbors). The KNN Algorithm Sequence is as follows:

- Determine the parameter K (number of nearest neighbors)
- Calculate the square of the euclidean distance of each object to the given sample data
- Sort the objects into groups that have the smallest euclidean distance
- Collect the category Y (nearest neighbors classification)
- Output the results / voting based on the specified K value.

The output of the KNN algorithm will be in the form of a class of the searched data or the searched sample data and will be calculated based on the most data occurrences according to the previously set K value.

2.5. System

According to James o'Brien, a system is a collection of interconnected components, working together to achieve a goal by receiving input and then producing output in a regular transformation process [14]. It can be concluded that the system consists of input, output and processes that are interrelated with each other to achieve the goal of forming a system. Examples of systems used in life are information service systems, ticket booking systems, banking systems, academic information systems, and inventory systems. The purpose of creating a system is to overcome a problem that has been determined, then the system can also make it easier for its users to reach information.

2.6. Simulation

Simulation according to Shannon is the process of designing a model of a real system followed by conducting experiments to study the system or evaluate strategies. Then according to Emshoff and Shimun, simulation is a system model where the components are presented by arithmetic processes and algorithms run by a computer to estimate the dynamic properties of the system [15]. It can be concluded that simulation is a representation of real models through an experiment to estimate or model the real system. The simulation approach begins with the construction of a system model that already exists or approaches the real system, so that it can represent real components into a smaller form but still describes the behavior of the real system. Some reasons for conducting simulations are that real systems are difficult to observe directly, cost very high costs, take a long time, and will damage existing systems [16]. Examples of simulations that are commonly available are such as world globes, driving license (SIM) simulator systems, flight simulation systems.

3. METHODOLOGY

This research is an experimental research. In this study, a simulation experiment was conducted to examine the benefits and uses of a flood detection system for early warning processed using the KNN method. The research method used in this study is the waterfall method and the steps are as follows.

3.1. Data collection stage

The researcher has collected data on flood information and rainfall information. Data that has been collected from BPBD and BMKG Bangkalan Regency. The data collected are in the form of a Flood Incident Map in Bangkalan Regency, a disaster-prone map of Bangkalan Regency, a map of river basins in Bangkalan Regency, an evacuation location map of Bangkalan Regency, an elevation map of Bangkalan Regency, a flood planning report for Bangkalan Regency from BPBD, while from BMKG Malang City climate report data from 2018 to 2023 was obtained.

3.2. Analysis stage

At this stage, the researcher analyzes the data that has been collected and draws conclusions to be used in the design stage as training data. The data that is considered suitable for use and relevant in this case is data concerning the causes of flooding, namely as follows.

3.3. Design stage

At this stage, the researcher will prepare the equipment that will be needed and the researcher will design a system using the data that has been obtained previously. The equipment used in this study are two water flow sensors and one water level sensor. Then this sensor will function as follows. The water flow sensor will record the flow of water in and out, this functions to simulate river flow, the inflow of water will be recorded as a simulation of water entering the river, the outflow of water will be recorded as a simulation of water flowing out of the river, then the water level sensor will function to record the water level.

3.4. Testing stage

At this stage, researchers and managers will conduct testing on the system with an experimental method, namely direct testing of the system[17]. This testing stage has been simulated on the tools and systems, resulting in varied results, on the tool the incoming sensor data is recorded. Then the system carries out the prediction process from the data that has entered the system.

3.5. Evaluation stage

At this stage, researchers will evaluate the system and fix system errors that exist in previous tests. The following is a context diagram

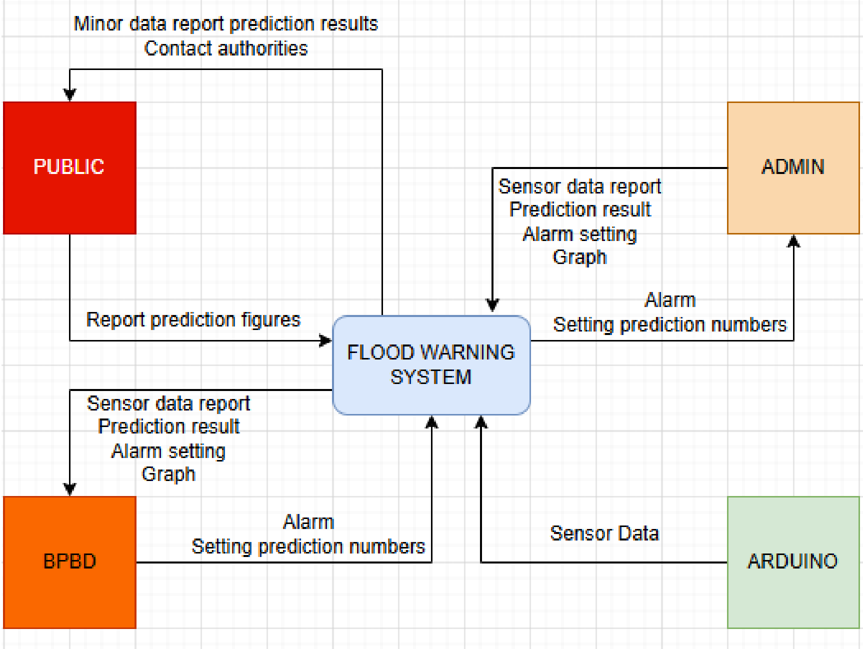

Figure 4. Diagram Context System

The following is a context diagram:

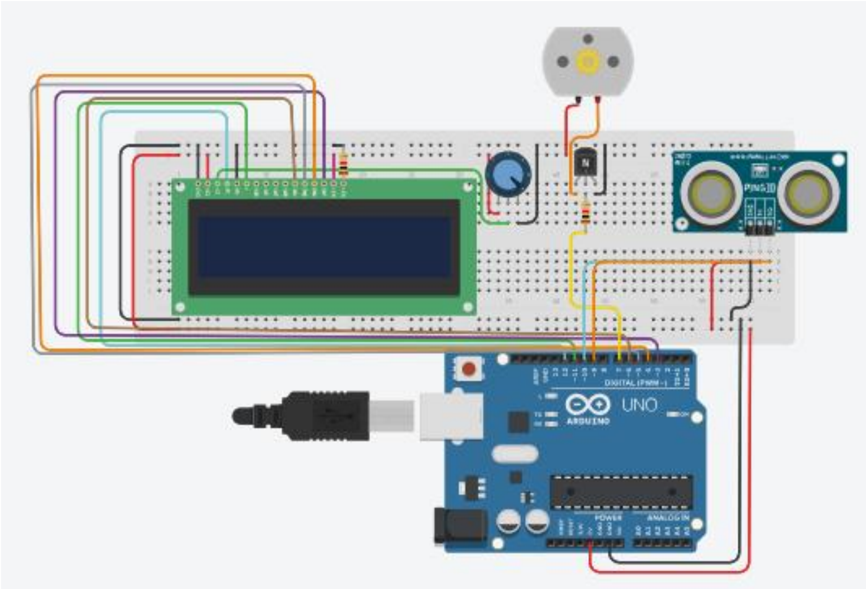

Figure 5. Arduino Circuit

In the circuit using Wemos D1 R2, two water flow sensors, one water level sensor, one green, yellow and red LED each, and a buzzer. The water level sensor in the circuit functions to capture water level data in the simulation container [18]. The water flow sensor in the circuit functions to capture the flow of water entering and leaving the container. The LED and buzzer in the circuit function to provide information regarding the status of the water level.

Implementation :

In this flood detection system, there is a server that is tasked with receiving data from Wemos, processing data to produce predictions, and displaying data[19]. An example of a request captured by the web server based on the apache access.log from Wemos:

192.168.43.36 - -
[03/Oct/2018:00:10:28 +0800] "GET
/arduino/add.php?water_flow_speed_in_=104.00&water_flow_speed_out_=0.00&water_height=266
HTTP/1.1"
200 - "-" "-"

The data sent by Wemos will then be processed by the add.php file on the web server. In add.php, the data is saved to the database. The data will be saved every five seconds.

4. RESULT AND DISCUSSION

4.1. Wemos Testing

Testing	Expected results	Observation result
Sending Wemos sensor data to the system	Data stored in database	Data stored in database

Figure 6. Wemos Test Table.

4.2 Algorithm Testing

KNN testing is done in two ways, namely the first is to find a group of k objects into the training data that are closest or closest or similar to the object to the testing data and test it with the Random Subsampling validation technique. The KNN algorithm test aims to test the prediction results in order to predict the status of the river. In the KNN algorithm, there are two types of data that will be used, namely the current weather data testing data and training data from the weather data that has been collected[20]. In this system, training data will be taken from data stored in the database, then testing data will be taken from user input.

In this prediction system, there are six variables that will be used in predicting floods, namely minimum temperature, maximum temperature, average temperature, and average humidity, rainfall, and duration of exposure. These six variables will be calculated with user input testing data. An example in this test will use the following data. As a method of data collection, researchers use weather data from 2018 to 2023 then the data is given a safe or flood status based on flood reports that occur. Then, noise cleaning is carried out on the data, such as unrecorded data or invalid data such as the numbers 9999 and 8888, and unused attributes are removed, leaving 53 data on the weather, the distance of which will then be calculated.

K Value	a	b	c	d	e	Accuracy
1						40%
3						20%
5						60%
7						40%
9						60%
11						60%
13						60%
15						60%
17						60%
19						60%

Figure 7. K accuracy table of each data

4.3 Simulation Results and System Testing

Simulation Results
Alert 4 At Initial Water Level, **Alert** 3 At 12 Seconds, **Alert** 2 At 30 Seconds
Alert 4 At Initial Water Level, **Alert 3** At 10 Seconds, **Alert 2** At 14 Seconds
Alert 4 At Initial Water Level, **Alert 3** At 7 Seconds, **Alert** 2 At 15 Seconds
Alert 4 at initial water level, **Alert 3** at 14 seconds

Figure 8. System Simulation Experiment Results

Two of the four events produce accurate results, but the other two events produce less accurate results, so it can be concluded that the accuracy of the algorithm and system is fifty percent (50%) and the results vary greatly depending on the available test data.

Then run the test with one of the white box methods to test the possible hypotheses in the system using the limit value method as shown in the following figure.

5. CONCLUSION

The conclusion reached by the system and simulation in this study is, this study aims to create a flood detection system and simulation for early warning using the KNN (K-Nearest Neighbor) method based on Arduino. This system was developed to help people living in flood-prone areas to take preventive measures before flooding occurs. This system uses a water level sensor to measure water levels, and a water flow sensor to measure the speed of water flow in and out of the river basin.

Rainfall, temperature, humidity, and other factors are used as variables in the application of the KNN algorithm to predict flood potential. Testing is carried out by comparing the results of the system prediction with actual flood event reports. The results of the study show that this system has a relatively low accuracy, which is around 60%, due to the lack of variation in the training data used. However, this study contributes to the development of a flood early warning system that can help people reduce the risk and impact of flood disasters.

REFERENCE

[1] T. J. Su, T. S. Pan, Y. L. Chang, S. S. Lin, and M. J. Hao, "A Hybrid Fuzzy and K-Nearest Neighbor Approach for Debris Flow Disaster Prevention," *IEEE Access*, vol. 10, pp. 21787–21797, 2022, doi: 10.1109/ACCESS.2022.3152906.

[2] M. Khalaf *et al.*, "IoT-Enabled flood severity prediction via ensemble machine learning models," *IEEE Access*, vol. 8, pp. 70375–70386, 2020, doi: 10.1109/ACCESS.2020.2986090.

[3] J. Ibarreche *et al.*, “Flash flood early warning system in colima, mexico,” *Sensors (Switzerland)*, vol. 20, no. 18, pp. 1–26, 2020, doi: 10.3390/s20185231.

[4] P. Muñoz, J. Orellana-Alvear, J. Bendix, J. Feyen, and R. Célleri, “Flood early warning systems using machine learning techniques: The case of the tomebamba catchment at the southern Andes of Ecuador,” *Hydrology*, vol. 8, no. 4, 2021, doi: 10.3390/hydrology8040183.

[5] T. A. Khan, Z. Shahid, M. Alam, M. M. Su’ud, and K. Kadir, “Early Flood Risk Assessment using Machine Learning: A Comparative study of SVM, Q-SVM, K-NN and LDA,” *MACS 2019 - 13th Int. Conf. Math. Actuar. Sci. Comput. Sci. Stat. Proc.*, 2019, doi: 10.1109/MACS48846.2019.9024796.

[6] S. Liu, R. Liu, and N. Tan, “A spatial improved-knn-based flood inundation risk framework for urban tourism under two rainfall scenarios,” *Sustain.*, vol. 13, no. 5, pp. 1–19, 2021, doi: 10.3390/su13052859.

[7] J. Ren, B. Ren, Q. Zhang, and X. Zheng, “A novel hybrid extreme learning machine approach improved by K nearest neighbor method and fireworks algorithm for flood forecasting in medium and small watershed of Loess region,” *Water (Switzerland)*, vol. 11, no. 9, 2019, doi: 10.3390/w11091848.

[8] H. Lumbantobing, I. Ratna Avianti, K. Harisapto, and S. Suharjito, “Flood Prediction based on Weather Parameters in Jakarta using K-Nearest Neighbours Algorithm,” *Eduvest - J. Univers. Stud.*, vol. 4, no. 6, pp. 5055–5065, 2024, doi: 10.59188/eduvest.v4i6.1339.

[9] S. Van Ackere, J. Verbeurgt, L. De Sloover, S. Gautama, A. De Wulf, and P. De Maeyer, “A review of the internet of floods: Near real-time detection of a flood event and its impact,” *Water (Switzerland)*, vol. 11, no. 11, pp. 1–26, 2019, doi: 10.3390/w11112275.

[10] M. I. K. Alfahadiwy and A. Suliman, “Flood Detection using Sensor Network and Notification via SMS and Public Network,” *Student Conf. Res. Dev. (SCOReD 2011)*, no. June, pp. 1–7, 2011.

[11] M. Esposito, L. Palma, A. Belli, L. Sabbatini, and P. Pierleoni, “Recent Advances in Internet of Things Solutions for Early Warning Systems: A Review,” *Sensors*, vol. 22, no. 6, 2022, doi: 10.3390/s22062124.

[12] D. S. Rani, G. N. Jayalakshmi, and V. P. Baligar, “Low Cost IoT based Flood Monitoring System Using Machine Learning and Neural Networks: Flood Alerting and Rainfall Prediction,” *2nd Int. Conf. Innov. Mech. Ind. Appl. ICIMIA 2020 - Conf. Proc.*, no. Icimia, pp. 261–267, 2020, doi: 10.1109/ICIMIA48430.2020.9074928.

[13] V. V. Krzhizhanovskaya *et al.*, “Flood early warning system: Design, implementation and computational modules,” *Procedia Comput. Sci.*, vol. 4, pp. 106–115, 2011, doi: 10.1016/j.procs.2011.04.012.

[14] S. Sankaranarayanan, M. Prabhakar, S. Satish, P. Jain, A. Ramprasad, and A. Krishnan, “Flood prediction based on weather parameters using deep learning,” *J. Water Clim. Chang.*, vol. 11, no. 4, pp. 1766–1783, 2020, doi: 10.2166/wcc.2019.321.

[15] W. Joko, “Flood Early Warning System Develop at Garang River Semarang using Information Technology base on SMS and Web,” *Int. J. Geomatics Geosci.*, vol. 1, no. 1, pp. 14–28, 2010.

[16] Universitas Prima Indonesia. Fakultas Teknologi & Ilmu Komputer, Institute of Electrical and Electronics Engineers. Indonesia Section. CSS/RAS Joint Chapter, and Institute of Electrical and Electronics Engineers, “MECnIT 2020 : International Conference on Mechanical, Electronics, Computer, and Industrial Technology : June 25-26, 2020, Universitas Prima Indonesia, Medan, Indonesia,” pp. 30–35, 2020.

[17] C. Chen, Q. Hui, W. Xie, S. Wan, Y. Zhou, and Q. Pei, “Convolutional Neural Networks for forecasting flood process in Internet-of-Things enabled smart city,” *Comput. Networks*, vol. 186, p. 107744, 2021, doi: 10.1016/j.comnet.2020.107744.

[18] C. Moreno *et al.*, “Rivercore: IoT device for river water level monitoring over cellular communications,” *Sensors (Switzerland)*, vol. 19, no. 1, 2019, doi: 10.3390/s19010127.

[19] I. Suwarno, A. Ma’arif, N. M. Raharja, A. Nurjanah, J. Ikhsan, and D. Mutiarin, “IoT-based Lava Flood Early Warning System with Rainfall Intensity Monitoring and Disaster Communication Technology,” *Emerg. Sci. J.*, vol. 4, no. Special issue, pp. 154–166, 2020, doi: 10.28991/ESJ-2021-SP1-011.

[20] M. Anbarasan *et al.*, “Detection of flood disaster system based on IoT, big data and convolutional deep neural network,” *Comput. Commun.*, vol. 150, no. November 2019, pp. 150–157, 2020, doi: 10.1016/j.comcom.2019.11.022.