Flood prediction System for Kelani River

HPSW Gunawardhana[1], RL Dangalla[2]

[1,2] *Department of Computing & Information Systems, Faculty of Applied Sciences,*
Sabaragamuwa University of Sri Lanka, Belihuloya, Sri Lanka
[#] Gunawardhana H.P.S.W.; <sanduniwasundara123@gmail.com>

Abstract—*A flood can be defined as the temporary inundation of an area of land. Physically, flooding can result when the total ingress of water to a locality becomes higher than total outflow from it. When the inflow to an area exceeds the outflow, it causes a surplus of water within the area. Step by step it continues to get worse becoming deeper and spreading. Kelani River is one of the most flooding rivers which affect human lives. Owing to the heavy rainfall of the upper catchment, the lower basin of the Kelani River is subjected to heavy floods.Flood damage can reduce by providing early warning messages to people. Main goals of flood prediction system are given as provide flood warning message to people, reduce the damage of the country economy. This research is to predict the maximum water level value for the next 24 hours. As a result of the research people can be aware of warnings, flooding areas and minimize the damage to their lives and properties.*

Flood model was designed by analysing data records of rainfall of the Kelani catchment. Knowledge discovery in databases is the process of discovering useful knowledge. Data mining techniques are used to data preparation and selection, data cleaning, incorporating prior knowledge on data sets and interpreting accurate solutions from the observed results.The decision tree machine learning algorithm was used to predict the water level of river according to the rainfall. Data set used in this study contains ten year time period daily maximum water levels and rainfalls. To train the algorithm used six river stations and its relevant rainfall that are most affected to the final result. This system has been shown to have an accuracy of more than 95%.

Keywords— daily maximum water level, prediction, decision tree algorithm.

I. INTRODUCTION

One-third of the world's land surface classified as arid or semi-arid, on the one hand, precipitation is rare and serious droughts occur frequently while on the other hand, floods may cause serious disasters. There is various kind of hazards in Sri Lanka. Flood, Drought, Famine, Landslide, and Water relates epidemics are some of them. Many Disaster Management organizations are looking to find ways to prevent such disasters and reduce their impact so some programs launched to manage emergencies through new technology. In many countries, the human and economic losses caused by extreme weather are on the rise, sometimes even faster than wealth. In some areas, natural disasters are now a major obstacle to poverty reduction at the household level (many studies show that disasters may cause families to fall into the poverty trap) and at the macro level (some areas are affected by regular activities and spend a large part of their resources for reconstruction rather than development).

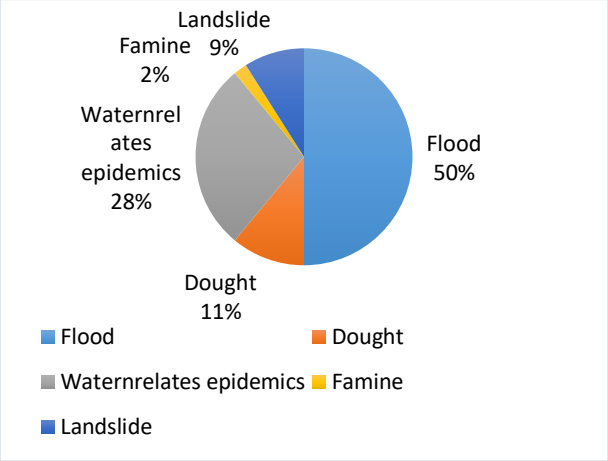

Figure1. Types of water related natural disasters

Figure1 describes most of the people affected by natural disasters, because of the flood. Economically the damage of people and properties are uncountable. Floods threaten human life and property not only for Sri Lanka. It happens worldwide. - Floods account for 15% of all deaths related to natural disasters.

A.Background

The Kelani is one of 103 major river systems in Sri Lanka. It takes the seventh place in respect of its extent of the watershed, 2340 km^2. However, it becomes third with respect to water resources aspect (4225 MCM average annual discharge) due to abundant rainfall in the catchment.

The river originates from the central hills near Adam's peak and traverses about 145 km through the South-western slopes of the Island to reach the sea near Colombo. The Kelani catchment is entirely situated in the wet zone of the country. The average annual rainfall of the catchment varies from 5700 mm in the upper catchment to 2300 mm in the lower basin. The major portion of the rainfall is received during the south-west monsoon period. However, the catchment remains wet throughout the year since it receives considerable amounts of rainfall during the North-East monsoon and inter-monsoonal periods. In earlier water absorption to the soil is high. But now water flowing paths are blocked with modernization. So that avoiding the flood is impossible. But people can reduce the damage to properties and lives. If the researcher can give a solution to reduce that damage it is very useful to the country.

When thinking this problem deeply the damage of flood is roundly affecting the economy of the country. In the flood period, people can't go for their jobs, people lose their daily income. The factories near the river will fail to achieve their production goals. Sometimes they are unable to export their expecting product level and it will become a loss to the exportation of the country.

B. Novelty

There is no flood model for entire Kelani River previously. There are some for Glencouse only. Prediction of maximum water level in meter msl value for next 24 hours is another novel point. Because the lead time is also important as well as the accuracy of prediction. When there is enough time to ready for the people nearby they can reduce the damage.

C. Significance

- Creating a system to predict the water level to the people who live near the river.
- They can reduce the damage to their lives and properties.
- People who travel nearby roads can aware of the flooding roads and use alternative roads.

D. Major goals

- Reduce the damage of people.
- Reduce the damage of properties.

Among all observed natural hazards, water-related damages are the most frequent and pose major threats to people and socio-economic development. People have big trouble with the flood in rainy days. Floods and landslides caused extensive damage to communities and their assets. It caused huge damage to the rural infrastructure such as roads and the damage to paddy cultivation and small-scale tea plantations had a direct impact on the livelihood of the people. The loss of labour opportunities due to floods has affected rural communities, thereby making them more vulnerable. If the researcher can predict Water level for the next 24 hours to one place, it is useful for a large number of people. Because they can go away from the places which the flood arrives. They can ready to reduce the damage for them.

II. METHODOLOGY

Figure 2 displays the research method and its flow in a diagram.

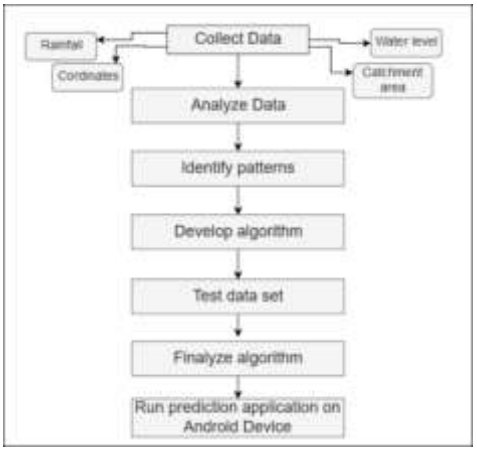

Figure 2. Research flow

As the first step researcher collects data which are needed for this study. Analyze the data as the next step using Weka software. Waikato Environment is a knowledge Analysis tool which is a suite of machine learning software written in Java. Fill missing values, Identify patterns, Design graphs and charts are done by using Weka.

Data parameters for the flood model are given below.

- Rainfall stations coordinates (x,y)
- Gauging stations coordinates (x,y)
- Rainfall (mm)
- Water level (mmsl)
- Catchment area (km2)

All the raw data are collected at the beginning. According to georeference, the different stations have special flood levels. Those values are given below in table1.

Station	Alert	Minor	Major	Unit
Hanwella	7.56	8.56	10.56	msl
Glencourse	14.88	16.38	18.88	msl
Holombuwa	44.87	47.97	49.57	msl
Deraniyagala	77.97	78.97	79.57	msl
Kithulgala	55.003	56.003	58.003	msl
Norwood	1099.26	1099.76	1100.76	msl

Table 1: Kelani river levels

Alert level denoted the maximum water level of the river in that coordinate. Minor flood is the level which inundates polders. Major flood level means the high flood which inundates roads and properties. The coordinates of the river stations are given below in below. X is longitude and Y is latitude.

Station	Coordinates	
	x	Y
N' Street	101112	195586
Holombuwa	144013	220455
Norwood	182051	182230
Kithulgala	160270	198922
Deraniyagala	152036	191688
Glencourse	135963	197396
Hanwella	123689	189980

Table 2. Coordinates

The rainfall is measured in mm. For this study daily rainfall was used. The format of rainfall is given below in figure 3.

Year	month	day	deraniyag	Glencouse	Hanwella	kenilwort	kithulgala	kotiyag
2007	1	1	0.00	0.00	0.00	0.00	0.00	0.00
2007	1	2	0.00	0.00	0.00	0.00	0.00	0.00
2007	1	3	0.00	0.00	0.00	0.00	0.00	0.00
2007	1	4	0.00	0.00	0.00	0.00	0.00	0.00
2007	1	5	0.00	0.00	0.00	0.00	0.00	0.00
2007	1	6	0.00	0.00	0.00	0.00	0.00	0.00
2007	1	7	0.00	0.00	0.00	0.00	0.00	0.00
2007	1	8	0.37	0.00	0.00	0.02	0.11	0.00
2007	1	9	0.00	0.01	0.14	0.25	0.28	0.62
2007	1	10	5.91	6.26	2.19	1.09	4.35	1.47
2007	1	11	0.64	0.67	0.31	2.06	0.98	1.34
2007	1	12	0.00	0.53	0.03	0.14	0.58	0.29
2007	1	13	0.00	0.00	0.29	0.00	0.13	0.00

Figure 3. Hanwella river station rainfall

Water level measured in meter msl. The water level stations are Norwood, Holombuwa, Kithulgala, Deraniyagala, Glencouse, Hanwella. Water level sample sheet is given below in the figure4. The daily maximum water level is used for prediction.

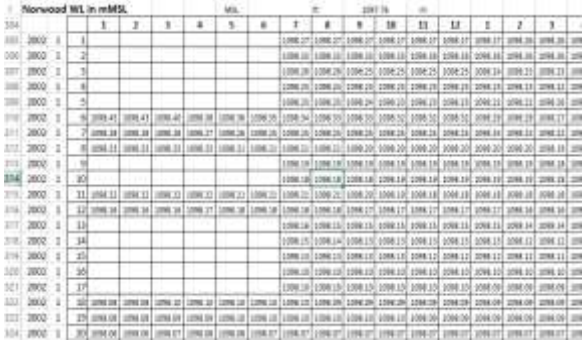

Figure 4. Norwood water level in mmsl

E. Data analysis tools and techniques

The affecting rainfall of stations and its weights can find out by using the Thiessen polygon method. The Thiessen Polygon approach has probably used the technique when there is more than one measurement in hydrometeorology for determining average precipitation over an area. The basic concept is to divide the watershed into several polygons, each one around a measurement point, and then take a weighted average of the measurements based on the size of each one's polygon, In here measurements within large polygons are given more weight than measurements within small polygons.

$P = (\sum_{i=1}^{n} PiAi)/ (\sum_{i=1}^{n} Ai)$

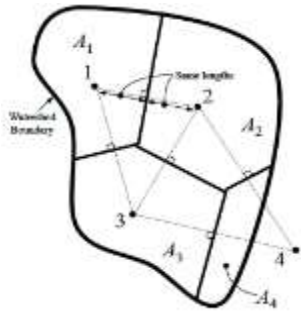

Figure 5. Thiessen polygon

Weighted average denoted by Pi and areas of each polygon denoted by Ai. In this given example display how Thiessen Polygons are made. First connect all of the measurement points, the dots in the figure are measurement points and they have been connected with dashed lines. As next step perpendicularly bisect each of the "connecting" lines and extend the bisecting lines until the either intersect the watershed boundary or another bisecting line. In this figure, there are four polygons which note that measurements made at point4 will be contributing less to the final average than, say, the measurements at point 2.

According to this method, Kelani river stations are mapped in Arc GIS and find out weights and catchment areas. Norwood catchment is given below as an example.

Figure 6. Norwood Thiessen map

Norwood and Kotiyagala rainfall stations affect to Norwood Water level according to geometry. After drawing this map it calculates the affecting catchment areas with weights. Those weights are given below.

WL Station	RF Station	Area(km²)	weight
Norwood	Norwood	24.08	0.77
	Kotiyagala	70.79	0.23
Holombuwa	Yatideriya	96.41	0.63
	Holombuwa	57.53	0.37
Kithulgala	Norton	66.43	0.15
	Kotiyagala	88.47	0.21
	Norwood	105.76	0.25
	Maussakele	108.38	0.25
	Kithulgala	16.95	0.04
	Kenilworth	42.92	0.10
Glencourse	Glencourse	187.45	0.12
	Norton	76.07	0.05
	Kotiyagala	88.52	0.06
	Yatideriya	250.77	0.16
	Holombuwa	140.63	0.09
	Labugama	29.47	0.02
	Deraniyagala	211.05	0.14
	Norwood	105.77	0.07
	Maussakele	114.82	0.08
	Maliboda	113.52	0.07
	Kithulgala	137.88	0.09
	Kenilworth	65.40	0.04
Hanwella	Glencourse	315.36	0.17
	Norton	76.07	0.04
	Kotiyagala	88.52	0.05
	Yatideriya	250.77	0.14
	Holombuwa	140.63	0.08
	Labugama	101.58	0.06
	Hanwella	105.53	0.06
	Deraniyagala	211.04	0.12
	Norwood	105.78	0.06
	Maussakele	114.82	0.06
	Maliboda	113.52	0.06
	Kithulgala	137.88	0.08
	Kenilworth	65.40	0.04
Deraniyagala	Norton	9.61	0.05
	Deraniyagala	28.99	0.16
	Maussakele	6.44	0.04
	Maliboda	110.00	0.61
	Kithulgala	24.10	0.13
	Kenilworth	0.92	0.01

Table 3. Thiessen weights

After arranging this data weighted rainfall is calculated by using weight and rainfall.
Weighted rainfall = weight*raifall

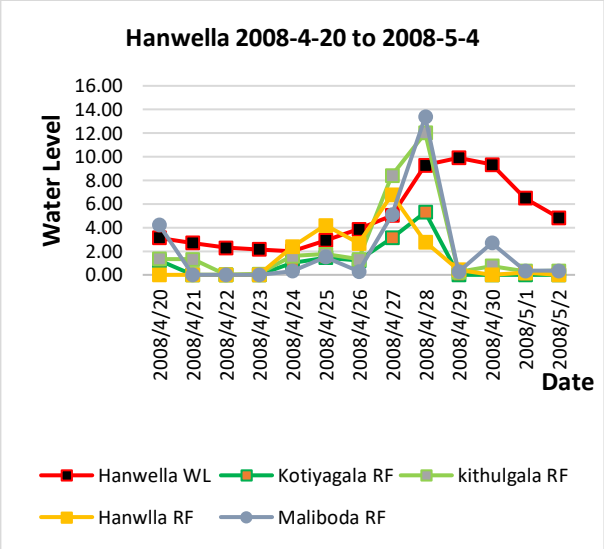

Figure 7. Hanwella rainfalls and water level

This figure 7 graph shows the rainfall of the previous day is affect to the next day maximum water level. In here Hanwella WL is depends on the Hanwella basin RF of previous day. The model is created in that condition. Six models were created for the entire river. One model is differ from another because the number of rainfall stations are vary from each other.

F. Implementation

Web page is designed to give the message to people and to collect rainfall. The web interface of the flood model developed using PHP, HTML, and MySQL. In the home interface in figure 8 view the six river stations that can view the warning message of prediction.

Figure 8. Home interface

When clicks on one station it moves to another interface with the rainfalls related to that particular station. In the figure 9 user can see after a click on "Hanwella", Hanwella catchment rainfalls appear. The user can enter rainfall in mm. For all other stations it displays station's relevant rainfall stations.

After filling all the values user can submit it and then the flood message will appear on the screen.

Figure 9. Station interface

III. RESULTS

Decision tree machine learning algorithm is used to create the model. The decision tree is the simple supervised algorithm mostly researches use. This is mostly used for classification problems. This works for both categorical and continuous dependent variables. This algorithm usually can divide the population into two or more homogeneous sets. The researcher used 70% training data and 30% testing data for model validation. 2007 to 2017 data set was used in this research.

In this python programme researcher used scikit-learn library. It shares a common basic API with a presumed library is designed to work with a set of mathematical and scientific packages centered on the NumPy and SciPy libraries. It is mainly researchers in various fields (computer science, neuroscience, astrophysics, etc.). This project also benefits from many occasional contributors who suggest small bug fixes and improvements. Scikit-learn's popularity can be gauged from various indicators such as the hundreds of citations in scientific publications, successes in various machine learning challenges, and statistics derived from our repositories and mailing list. Here, uses some of its modules such as train_test_split, Decision Tree Classifier, and accuracy_score. Basically, the Pattern is trained using all the data and can get answers using the trained patterns. In below table 4 display the evaluation accuracy of the models.

Station	Result accuracy
Norwood	0.9489
Holombuwa	0.9569
Kithulgala	0.9423
Glencourse	0.9715
Hanwella	0.9727
Deraniyagala	0.9849

Table 4. Accuracy of results

Accuracy and the warning message are the main outputs of this system.

IV. CONCLUSION

Developing an effective Flood Warning System is not an easy task so that at the present there is no effectively functioning Flood Warning System in Sri Lanka. The challenge is to incorporate or raise the level of operation of other components most directly related to eliciting appropriate protective behaviour, especially from communities threatened with floods. Interestingly, making a better Flood Warning System is not highly expensive: apart from the forecasting components, most elements are more labour than capital-intensive.

ACKNOWLEDGEMENT

I wish to record my deep sense of gratitude and profound thanks to my research supervisor Mr.R.L.Dangalla for his keen interest, inspiring guidance and constant encouragement with my work during all stages, to bring this research into fruition.

REFERENCES

[1] Eng.Sivapalasundaram, Eng.P.Hettiarachchi, and Eng.K.K.A.Piyasena, "Hydrological Annual 2015/2016," 2016.
[2] J. K. Beven, Rainfall–Runoff Modelling, vol. 15, no. 4. 2004
[3] L. Wang and Q. Cheng, "Web-based hydrological modelling system for flood forecasting and risk mapping," vol. 7145, p. 71450A, 2008.
[4]"mathworks.com." [Online]. Available: https://www.mathworks.com/matlabcentral/. [Accessed: 20-Sep-2018].
[5] M. Butts, "Process-based hydrological modelling framework MIKE SHE for flood forecasting on the Upper and Middle Odra," ICID 21st Eur. Reg. Conf. 2005, vol. 5, no. May, pp. 1–8, 2005.
[6] P. K. Thakur and A. Sumangala, "Flood Inundation Mapping and 1-D Hydrodynamic Modeling Using Remote Sensing and GIS Technique," no. September 2006, 2006.
[7] S. Lecturer and S. Lanka, "Flood risk zone mapping of Kelani River basin in North of Colombo City The analysis was undertaken to shows the probable extent of flood water in case of failure of embankment on the left sides at Kelani River."