Real Time Flood Prediction for Sungai Isap Residence Using Support Vector Machine

Khoo Chun Keong, Mahfuzah Mustafa, Ahmad Johari Mohammad, Mohd Herwan Sulaiman

Faculty of Electrical & Electronics, Universiti Malaysia Pahang, Pahang, MALAYSIA

***Abstract*— A flood is an extremely dangerous disaster that can wipe away an entire city, coastline, and rural area. The flood caused extensive damage to life and property that has great erosive power and can be extremely destructive. This research explores the use of Support Vector Machine (SVM) method to predict the flood in 2015. Total investigated 29 months data covering from January 2013 until May 2015 in the Sungai Isap Kuantan, Pahang, Malaysia. Rainfall, temperature and river level are considered as a factor of flood. From the experimental, it can be concluded the SVM technique is an efficient flood prediction. The proposed of SVM prediction models is for implemented in real-time flood warning system to let the residence to take action immediately.**

Keywords*— *extremely dangerous disaster , flood, extensive damage, extremely destructive, SVM

I. INTRODUCTION

Floods are caused by rains, river overflow, temperature, strong winds in coastal areas and dam breaking. The floods hit Malaysia in 2014 that make more than 200,000 people affected while 21 killed in the floods [1]. This flood has been described as the worst floods in these decades. The floods can have been devastating consequences and can have an effect in the economy, environment and people.

The flood effect that can divide by two types, there are primary effects and secondary long-term effect [2]. The primary effect of flooding included loss of life, damage on building and structures, damage power generation, loss of clean drinking water treatment and water supply, and raise the risk of waterborne diseases. Flood waters typically inundate farmland, making the land unworkable and preventing crops from being planted or harvested, which can lead to shortages of food both for humans and farm animals [3]. Secondary long-term effects, that affect economic hardship due to a temporary decline in tourism, rebuilding costs, and food shortages leading to a price increase common after the effect of severe flooding.

There are three major types of flood, there are flash floods, rapid on-set floods, and slow on-set floods [4]. Flash flood is kind occurs within a very short time from two hours to six hours and sometimes within a few minutes that mostly caused as a result of heavy rain or dam break. The rapid on-set floods are taking slightly longer to develop, and the flood can last for one or two days only. For rapid on-set flood's type, people can quickly save their valuable things and escape before flood getting very dangerous. The slow on-set flood is a type that usually that results of water bodies over flooding their banks. They tend to develop slowly and can last for days and week. They usually spread over many kilometers and occur more in flood plains.

This research explores the use of Support Vector Machine (SVM) models to predict the floods. Rainfall, temperature, and river levels are considered as the factor of a flood, and a number of SVM architectures were evaluated as flood prediction models. The data from the factor was noted from January 2013 until May 2015.

II. STUDY AREA

In order to examine the performance of SVM and the impact of the prediction, Sungai Isap, Kuantan (Latitude. 3.7833°, Longitude. 103.3000°), the capital city of Pahang state is located in the east of Malaysia, known as 'gateway to the east coast', is a fast growing commercial city in the east coast of peninsular Malaysia with the population of more than 400,000 peoples. Figure 1 and Figure 2 show the picture of the satellite and maps at Sungai Isap, Kuantan that clearly see effect flood at Sungai Isap if from the river Sungai Kuantan. For the past few years, Kuantan has faced with the worst flood caused by the destruction and loss. Figure 3 shows that picture flood on December 2013. The Municipal Council of Kuantan (MPK) has carried out many solutions to these problems, especially in the famous area such as Sungai Isap, including the construction of a man-made lake as catchment during heavy rainfall or monsoon period bet the scenario became even worse from year to year.

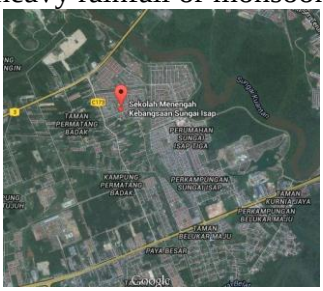

Figure 1 : Satellite view

Figure 2 : Maps View


Figure 3 : Flood 2013

III. METHODOLOGY

The flood prediction flow chart based on SVM is shown in Figure 4 . Flood conditioning factor data set was constructed by factors from Jabatan Pengaliran & Saliran (JPS) is temperature, river level, rain level, dew point, humidity, sea level pressure, visibility, wind and precipitation from January 2013 until May 2015. These factors were gathered to form a visualize data histogram, then from the visualize data histogram found the factors of significant data and clearly that the data temperature, river level and rain level is most significant. Due to the large difference in the order or magnitude of the value, the available samples are scaled in 0-1 using the normalization preprocessing method. SVM is a technique is supervised machine learning technique and is an efficient and reliable tool in flood predict [5]. The Radial Basic Function (RBF) kernel function of the SVM model is used in this paper. The training part uses 70% of the data that is from January 2013 until September 2014 (21 months) and for the testing part is use 30% of the data that start from October 2014 until May 2015 (8 months).

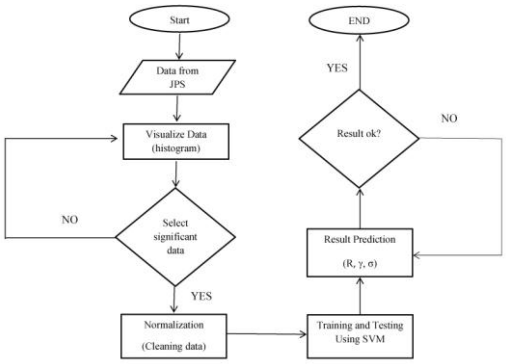

Figure 4: Flow chart SVM prediction

IV. RESULT AND DISCUSSION

The performance of the SVM model is evaluated by prediction of the flood happening. All of the calculation SVM results with 100 times of calculation training. Figure 5, the meaning of 'R' is the percentage of SVM prediction shows, the highest percentage prediction is 94.785%, and the lowest is 85.229%. From the Figure 6 the graph of value γ, the max of the value is 4.55E-04 and the minimum value is 3.19E-08.

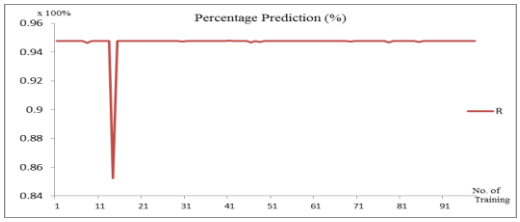

Figure 5 : Graph percentage prediction

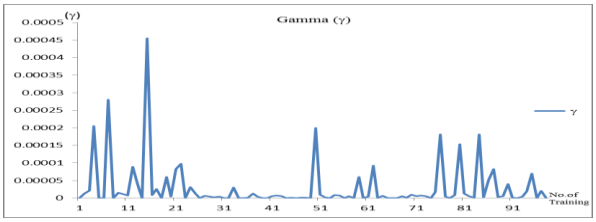

Figure 6: Graph gamma (γ)

From the Figure 7, the graph of value σ, the max of the value is 19458009, and the minimum value is 31.6192. The reading of highest percentage prediction from the 100 times SVM calculations is 94.785% and got 23 times same reading, so choose the best of five readings with the gamma reading is not large and sig2 is not too small to prevent over-fitting problems.

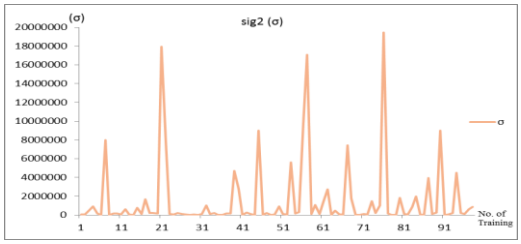

Figure 7: Graph sig2 (σ)

	R (%)	γ	σ
1	94.785	2.92E-07	4707446
2	94.785	4.28E-07	9812685
3	94.785	1.35E-06	9009867
4	94.785	5.36E-06	17095028
5	94.785	5.57E-06	1777336

Table 1: Top 5 best readings

V. CONCLUSION

The flood is the most damaging catastrophic phenomena in the world. Over the last decade, the flood becomes the hot topics in the world, even the literature, because this evaluation is a tough and nonlinear problem. Many methods have been tried by researchers in literature, but each method has their weakest point. Flooding on year 2013 caused serious damage in Sungai Isap, Kuantan, Pahang, Malaysia. The goal of this study was to do the accuracy flood prediction to assist having proper management over the effected area. For use SVM calculation produced 94.785% of the rate prediction. The information from current research can help JPS, Malaysian Meteorological Department, and government to perform the proper take action immediately.

REFERENCES

[1] A. F. Presse. (2014, 9 march). *Floods kill 21 in Malaysia, waters recede*. Available: http://news.asiaone.com/news/malaysia/floods-kill-21-malaysia-waters-recede

[2] C. C. Attribution-ShareAlike. (2015, 29 September). *Flood*. Available: https://en.wikipedia.org/wiki/Flood#cite_note-1

[3] L. B. Stephen Bratkovich. (1993, 3 March). *Flooding and its effect on trees*. Available: http://www.na.fs.fed.us/spfo/pubs/n_resource/flood/cover.htm

[4] eSchooltoday. (2010, 1 April). *Your cool facts and tips on flooding*. Available: http://eschooltoday.com/natural-disasters/floods/types-of-floods.html

[5] M. S. Tehrany, B. Pradhan, S. Mansor, and N. Ahmad, "Flood susceptibility assessment using GIS-based support vector machine model with different kernel types," *Catena*, vol. 125, pp. 91-101, 2015.