C209

Flood Suspcpibilty Prediction using Machine Learning Techniques: Case study of Vu Gia-Thu Bon basin in Vietnam

○Mohamed SABER・Tayeb BOULMAIZ・Mawloud GUERMOUI・Karim ABDRADO・Sameh KANTOUSH・Tetsuya SUMI・Hamouda BOUTAGHANE・Tomoharu HORI・ Doan Van BINH・ Nguyen Quang BINH・ Emad MABROUK

Introduction

The devastating impacts of flash floods have been recorded and documented in developing and developed countries, however, they are more severe in developing countries, such as Vietnam. Causes of high flood risk in Vietnam include tropical cyclones, dense river networks, typhoons, and extended coastal areas. Most of the flood susceptible areas in Vietnam are densely populated; thus, there is a constant risk of loss of life and property in such areas. Flash flood risk mitigation requires precise and accurate flood monitoring measures to support hazard management. Recently, many machine learning techniques, have been recently used to improve the predictive accuracies of natural hazards. Therefore, the objective of this study is to use machine learning approaches such as Random Forest, LightGBM and CatBoost (Saber et al. 2021) for predicting FFS in humid environments, case of the the Vu Gia-Thu Bon (VGTB) basin in Vietnam and compare the results Rainfall Runoff Inundation Model (RRI).

Study Area

The VGTB River system (Fig. 1) is the largest river system in the Central Coast region of Viet Nam with an area of 10350 km2. The topography of the basin is mainly a hilly, mountainous area with an elevation of over 552m, accounting for 60% of the total area of the basin. The location of basin in the tropical monsoon climate, having dry summer (January-August) and rainy (September-December) seasons.

The combination of the intense rain and the steep terrain leads to flooding with high intensity and short occurrence time. Every year, the basin has 4 to 8 floods. The flood peak usually occurs in October and November due to different weather patterns such as typhoons, tropical depressions and cold air.

The VGTB basin has two sub-basins: the Vu Gia basin and the Thu Bon basin. Quang Hue River is connected between the Vu Gia and Thu Bon Rivers. Vu Gia River originates from the western slope in Kon Tum, flowing through Quang Nam province and Danang city and meeting the sea at Cua Han estuary.

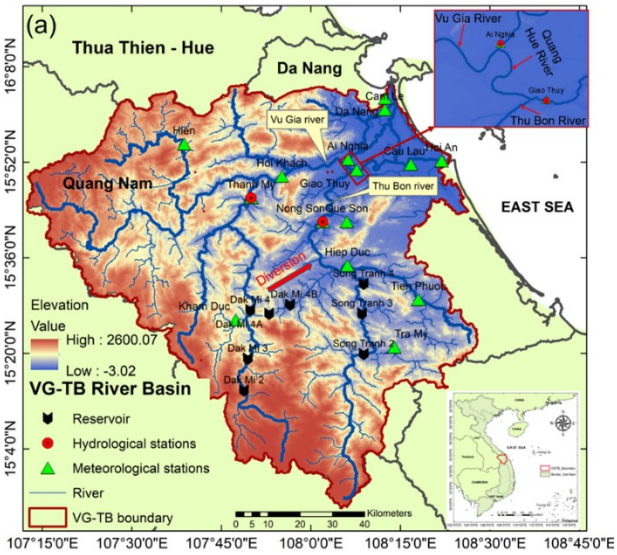

Fig. 1 (a) The location map of the VGTB Basin in Vietnam.

Approach and methods

Two main datasets were prepared for the flood susceptibility mapping using machine learning. First, the flood inventory map of the flood occurrences was prepared based on the post flood field survey.

Additionally, the non-flooded points were randomly selected throughout the catchment using the geographic information system (GIS) environment. Furthermore, a total of 10 influencing factors were also prepared for modelling such as Elevation, Slope, Aspect, Plan Curvature, Hillshade, Horizontal flow distance, SPI, Geology, Rainfall, Land use/Land cover. The dataset was divided into two groups of training (70 %) and testing (30 %) through a random selection scheme. Afterwards, two methods of information gain ratio and multicollinearity test (VIF) were used to check the importance of the factors in FFS within the study area (Fig. 2). Secondly, the implementation of the proposed machine learning approaches, the datasets have been divided into two main categories, training (70%) and validation (30%), and three algorithms were used namely, RF, LightGBM, and CatBoost. The results of the models were assessed for the accuracy using different measures including the most famous one AUC. Additionally, as we have compared the results of machine learning and the Rainfall Runoff Inundation Model.

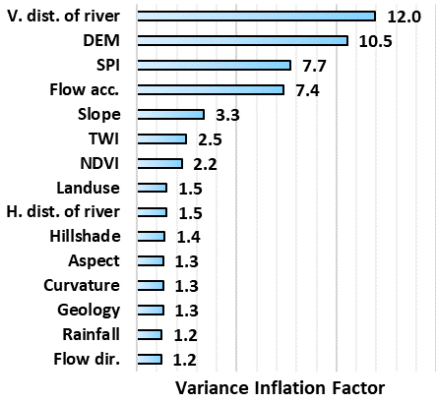

Fig. 2 Analysis of influencing factors using VIF for flood susceptibility.

Results and discussions

The results showed that the used three approaches are reasonably predicting the FSM in the study area with very high performance as the area under the receiver operating characteristic (ROC) curves for Random Forest, CatBoost and LightGBM models was more than 99%. The spatial Flood susceptibility mapping using Machine learning techniques shows acceptable agreement with the flood inundation maps developed by RRI model (Fig. 3) revealing that the Machine learning is a promising approach and can be used effectively as alternative approaches for the hydrological models.

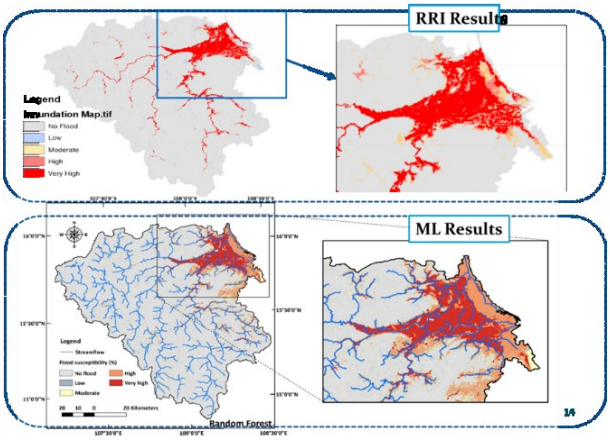

Fig. 3 Flood susceptibility map by the Random Forest.

Conclusion

Three ML models (Random Forest, LightGBM, and CatBoost) are successfully used in predicting of flooding susceptibility in a Humid area which experienced consecutive extreme typhoons. The results can be used as guidance for the planners and managers to mitigate the floods in the high prone flood susceptible regions such as VGTB, Vietnam.

Acknowledgment

This work was funded by APN "Asia-Pacific Network for Global Change Research" under project reference number CRRP2020-09MYKantoush (Funder ID: https://doi.org/10.13039/100005536).

References

1. Saber, M., Boulmaiz, T., Guermoui, M., Abdrado, K. I., Kantoush, S. A., Sumi, T., ... & Mabrouk, E. (2021). Examining LightGBM and CatBoost models for wadi flash flood susceptibility prediction. Geocarto International, 1-26.