京都大学 防災研究所
Disaster Prevention Research Institute
Kyoto University 2023LS-04

Collaborative Research Project Report

Project No. 2023LS-04

May, 2024

Coordinator: Prof. Tetsuya SUMI

[Project Report]

Date: May, 9, 2024

☐ General Collaborative Research

☐ International Collaborative Research

☑ **Long/Short-term Research Visit**

☐ New Exploratory Research

To: Director of the Disaster Prevention Research Institute, Kyoto University,

Project Title:

Development a Global Model for Flood Susceptibility Mapping Using Machine Learning

Approaches.

Principal Investigator : Emad_Mabrouk

Affiliation : Department of Computer Science, Faculty of Computers & Information, Assiut University, Egypt

Name of the DPRI Contact Person : Prof. Tetsuya SUMI

Research Period / Duration of Stay : Aug. 17 – Sep. 16

Research Location / Location of Stay : Sumi Lab, WRRC, DPRI/ Stay in Uji, kyoto

Number of the Participants in the Project: 4 (DPRI: 3 / non-DPRI:1)

Research Report

(1) Purpose

The main research objective is to develop a machine learning (ML) model capable of generating a global function or program for FSM, addressing fundamental drawbacks of classical approaches in previous applications, such as locality, limitations of black-box models, and data-scarce regions.

(2) Research Approach: Machine Learning Algorithms

In this project, a machine learning algorithm, called Memetic Programming (MP), was employed to predict flood depth in FSM and generate a new function or program that can be trustily used for different dataset from different regions.

The search process of the MP algorithm is initiated by generating a fresh population of programs, represented as trees with all nodes being chosen randomly. Subsequently, the algorithm iterates through the following strategies: selection, mutation, and crossover. The algorithm picks a set of programs (parents) and employs mutation and crossover operations to create new offspring to form a new population. Following this, the MP algorithm uses various local search procedures to enhance the elite programs within the current population. These steps for generating populations are iterated until a termination condition is met. At that point, the algorithm returns the best program discovered. The main flowchart of the MP algorithm is illustrated in Figure 1a, and some examples of MP trees are presented in Figure 1b. Herein, each tree generated by MP is compiled into a mathematical formula to represent the function or relation between the input and the output.

To evaluate the machine learning model's ability to predict the flood depth, we employed a range of criteria to assess its accuracy and reliability. These criteria, including the mean absolute error (MAE), correlation coefficient (R), and normalized Nash-Sutcliffe coefficient (NNSC), offer comprehensive insights into different aspects of the model's predictive capabilities.

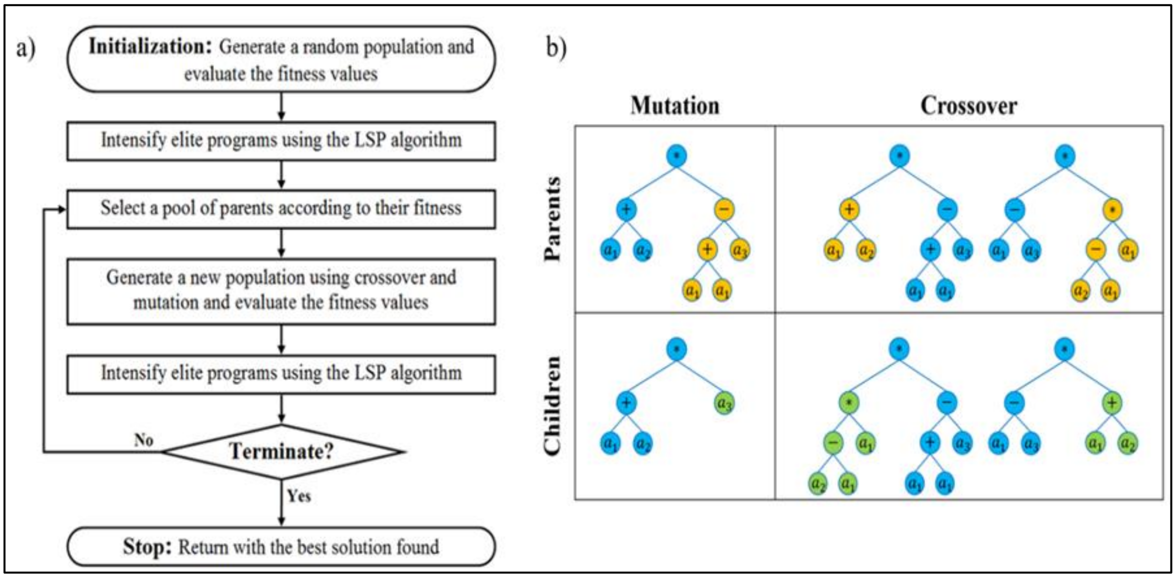

Fig. 1 a) The MP flowchart, b) The MP main operations; Mutation and Crossover.

(3) Summary of Research Findings

The developed MP model has been successfully applied for two case studies from Saudi Arabia (KSA) and Vietnam (VT). The algorithm provides a set of programs, each of which was translated as a mathematical function, with good accuracy and high correlation coefficients values with $R > 0.7$. Then, extra experiments have been done to test these programs through the cross-validation stage, where the program generated by Vietnam dataset is used to predict the output of KSA dataset, and vice versa. Results are shown in Table 1 for Vietnam and KSA datasets. The results showed promising performance of these programs with the cross-validation datasets as illustrated in Figure 2.

Table 1. MP results for VT and KSA datasets.

Dataset	Points	MSE	RMSE	MAE	MBE	R
Training (VT data)	1177	0.147	0.384	0.237	0.006	**0.808**
Testing (VT data)	505	0.158	0.397	0.249	0.013	**0.792**
Cross-Validation (KSA)	492	0.393	0.627	0.276	0.269	**0.521**
Training Set (KSA data)	344	0.138	0.372	0.207	0.014	**0.771**
Testing Set (KSA data)	148	0.189	0.435	0.244	0.004	**0.76**
Cross-Validation (VT)	1682	0.242	0.492	0.331	0.008	**0.656**

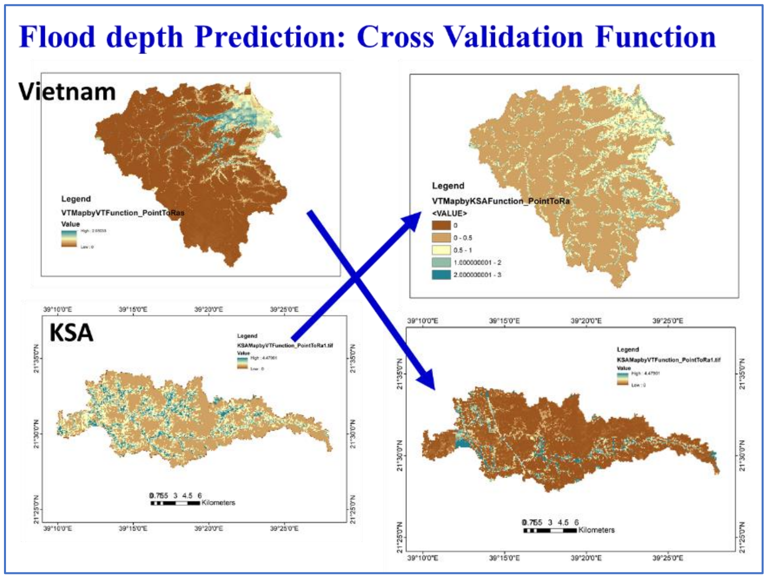

Fig. 2 Shows the selected basisn from Vietnam and Saudi Arabia as cross-validation cases.

Moreover, the programs generated using datasets from Vietnam and KSA have been applied to additional datasets from Algeria, Japan, and Syria. The results were very satisfactory, and an example from the Syria case study, as an ungauged basin, is shown in Figure 3.

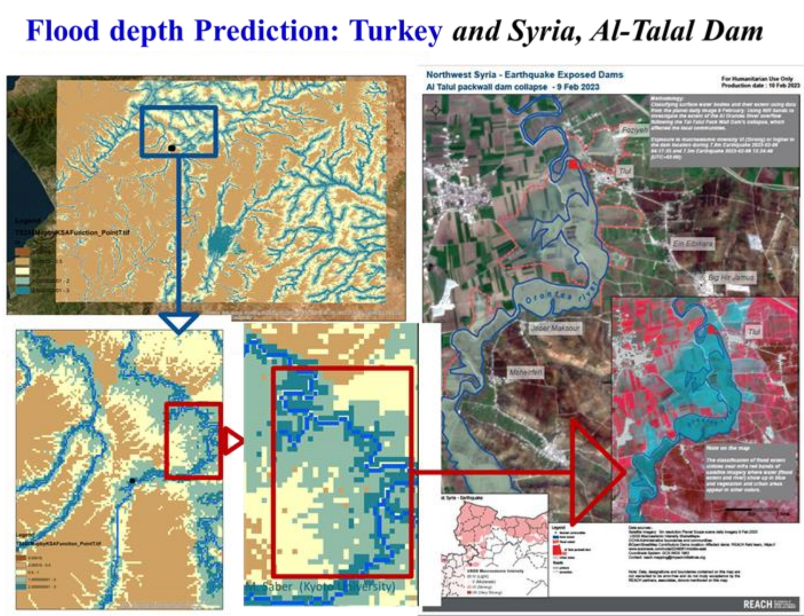

Fig. 3 The validation application of the developed function of the machine learning in Syria.

Therefore, the most important finding of this research is the development of programs and functions that can be applied independently to ungauged basins without the need for training datasets. These experiments and their results pave the way for developing a global map for flood susceptibility mapping.

(4) Important outcomes and future work

We are now collaborating with various researchers to develop additional functions using different datasets from multiple case studies across different regions worldwide, as shown in Figure 4. Additionally, we will consider using new machine learning algorithms to enrich the experiments and perform deep analyses and comparisons of the results.

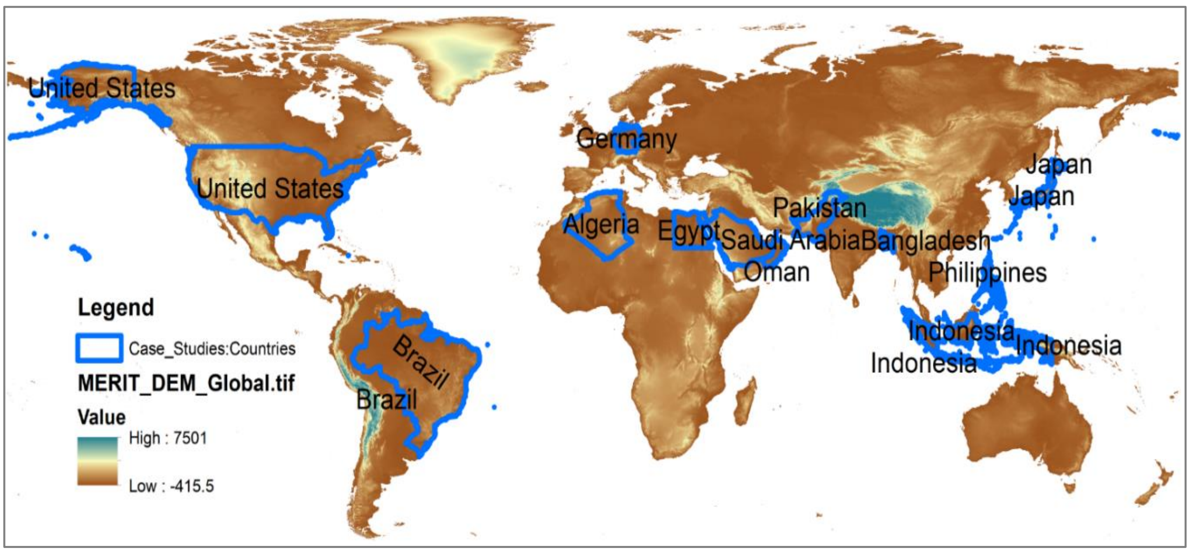

Fig. 4 Selected case studies for the Global ML model test and training

(5) Additional research activities and training during the Visit to DPRI

Along with the above findings, Prof. Emad conducted a training course for the students at Sumi Lab, and he also supported many students in Sumi Lab to use Machine learning in their work for examples, the following topics:

1. Sediment erosion prediction using Machine learning in Medjredah Basin, Algeria.
2. Tunnel Abrasion prediction using Machine learning in Japan.
3. Suspended sediment prediction in Oman and Japan.

(6) Publications of Research Findings

There are several publications under preparation for journal submissions, and we have already presented some related work in local and international conferences and published one related journal paper.

1. Saber, M., Boulmaiz, T., Guermoui, M., Abdrabo, K. I., Kantoush, S. A., Sumi, T., ... & Mabrouk, E. (2023). Enhancing flood risk assessment through integration of ensemble learning approaches and physical-based hydrological modeling. Geomatics, Natural Hazards and Risk, 14(1), 2203798.

2. Mohamed Saber*, Sameh A. Kantoush, Tetsuya Sumi, Emad Mabrouk. Machine Learning Approaches and Hydrological Modeling for Flood Risk Assessment, 2nd International Conference Water Resources Management & Sustainability: Solutions for Arid Regions, Dubai, UAE, Feb. 2024

3. Saber et al. Kantoush, S. A., Sumi, T., ... & Mabrouk, E. (2024). Real-Time Prediction of Suspended Sediment Concentration using Machine Learning: Case Study of Miwa Reservoir, Tenryu River in Japan, (in Review for JSCE).