Water Level Prediction on the Golok River Utilizing Machine Learning Technique to Evaluate Flood Situations

Pheeranat Dornpunya[1,], WatanasakSupaking[2], Hanisah Musor[2], Oom Thaisawasdi[2], Wasukree Sae-tia[2], Theethut Khwankeerati[2], and Watcharaporn Soyjumpa[2]*

[1] Thai Hydrologist Association & Technology & Hydro-Informatics Promotion Division, Hydro-Informatics Institute (Public Organization), Thailand
[2]Technology & Hydro-Informatics Promotion Division, Hydro-Informatics Institute (Public Organization), Thailand

**Corresponding Author: pheeranat.dor@gmail.com*

Abstract

During December 2022, the northeast monsoon, which dominates the south and the Gulf of Thailand, had significant rainfall that impacted the lower southern region, causing flash floods, landslides, blustery winds, and the river exceeding its bank. The Golok River, located in Narathiwat, divides the border between Thailand and Malaysia was also affected by rainfall. In flood management, instruments for measuring precipitation and water level have become important for assessing and forecasting the trend of situations and areas of risk. However, such regions are international borders, so the installed measuring telemetry system cannot measure the rainfall and water level of the entire area. This study aims to predict 72 hours of water level and evaluate the situation as information to support the government in making water management decisions, publicizing them to relevant agencies, and warning citizens during crisis events. This research is applied to machine learning (ML) for water level prediction of the Golok River, Lan Tu Bridge area, Sungai Golok Subdistrict, Su-ngai Golok District, Narathiwat Province, which is one of the major monitored rivers. The eXtreme Gradient Boosting (XGBoost) algorithm, a tree-based ensemble machine learning algorithm, was exploited to predict hourly water levels through the R programming language. Model training and testing were carried out utilizing observed hourly rainfall from the STH010 station and hourly water level data from the X.119A station between 2020 and 2022 as main prediction inputs. Furthermore, this model applies hourly spatial rainfall forecasting data from Weather Research and Forecasting and Regional Ocean Model System models (WRF-ROMs) provided by Hydro-Informatics Institute (HII) as input, allowing the model to predict the hourly water level in the Golok River. The evaluation of the predicted performances using the statistical performance metrics, delivering an R-square of 0.96 can validate the results as robust forecasting outcomes. The result shows that the predicted water level at the X.119A telemetry station (Golok River) is in a steady decline, which relates to the input data of predicted 72-hour rainfall from WRF-ROMs having decreased. In short, the relationship between input and result can be used to evaluate flood situations. Here, the data is contributed to the Operational support to the Special Water Resources Management Operation Center in Southern Thailand for flood preparedness and response to make intelligent decisions on water management during crisis occurrences, as well as to be prepared and prevent loss and harm to citizens.

KEYWORDS: eXtreme Gradient Boosting, Machine Learning, Golok River, Water Level Prediction, Flood Situation