

PREDICTING THE URBAN FLOOD RISK FOR THE KELANI RIVER BASIN

N. D. R. Perera^{a*}, U. P. Liyanage^a, K. K. W. H. Erandi^a, and S. S. N. Perera^a

^a*Centre for Mathematical Modelling, Department of Mathematics, Faculty of Science, University of Colombo.*

**danularan@gmail.com*

Abstract

The Kelani River Basin supports more than 25% of Sri Lanka's population, including the Colombo metropolitan area. Due to climate change, irregular monsoons, and rapid urbanization, over 1.2 million people downstream of Hanwella remain highly vulnerable to floods. Traditional hydraulic models, such as the Hydrologic Engineering Centre's River Analysis System (HEC-RAS), struggle with real-time adaptability, motivating the need for data-driven forecasting approaches. This study develops machine learning models to predict maximum water levels at the Hanwella gauge station over 24-hour and 24–48-hour horizons. A streamlined workflow was designed to minimize reliance on scarce topographic and land-use inputs by using only rainfall, streamflow, and water level data. An extensive dataset of over 80,000 hourly records (2013–2024) from six stations was pre-processed with validated Piecewise Cubic Hermite Interpolating Polynomial (PCHIP) interpolation and enriched with lagged features of water level, streamflow, and cumulative rainfall. Six regression algorithms were evaluated, with Extreme Gradient Boosting (XGBoost) emerging as the best-performing model by effectively capturing nonlinear hydrological dynamics and ensuring robust generalization. The models achieved predictive accuracy exceeding 90%. The 24-hour forecast delivered a ± 0.3 m error margin sufficient for timely evacuation planning, while the 24–48-hour forecast provided reliable lead time for reservoir management despite increased uncertainty at longer horizons. This first application of XGBoost to Hanwella water level forecasting highlights its potential for lean, data-driven early-warning systems. Further improvements may be achieved through advanced feature engineering and hybrid models such as XGBoost coupled with Long Short-Term Memory (LSTM) networks.

Keywords: *Flood Prediction, Kelani River Basin, Machine Learning, XGBoost, Early-Warning System*