AI-Based Flood Hazard Forecasting with Multi-Source Hydrometeorological Data in the Rhezala Basin (Tunisia)

Oral

Rim Ouachani*[1,2], Linda Mansouri[2], Hamouda Dakhlaoui[2]

[1]*University of Sousse/ Higher Institute of Transport and Logistics of Sousse, Sousse, Tunisia*

[2] *Université de Tunis El Manar/Ecole Nationale des Ingénieurs de Tunis (ENIT), Laboratoire de Modélisation en Hydraulique et Environnement (LMHE), Tunis, Tunisia*

* *rim.ouachani@enit.utm.tn*

ABSTRACT

Flood forecasting in semi-arid regions such as the Medjerda Basin in Tunisia remains highly challenging due to strong rainfall spatial variability, rapid hydrological response, and limited in-situ observations. This study addresses how artificial intelligence (AI) models, enriched with multi-source hydrometeorological data, can improve flood prediction and early warning in data-scarce catchments. The objective is to develop an integrated, data-driven framework that combines multi-source hydrological datasets, including in-situ measurements of precipitation and discharge (Q), and high-resolution climate reanalysis -derived soil information. This approach aims to improve flood hazard prediction in the Rhezala catchment, a key tributary of the Medjerda River. Two main hydrological drivers are incorporated into the modeling framework. The Antecedent Precipitation Index (API) is used as a dynamic proxy for basin wetness and runoff generation potential, while climate reanalysis derived soil moisture (SM) provides spatially distributed information on soil infiltration capacity and pre-event hydrological conditions. Together, these variables improve the representation of catchment state prior to flood events. The multi-step forecasting system is based on advanced artificial intelligence methods using recurrent neural networks, specifically Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) architectures. Both models are optimized using Bayesian hyperparameter tuning (Optuna) and systematically compared to assess their ability to capture nonlinear rainfall–runoff relationships under semi-arid conditions. Climate reanalysis derived soil moisture is directly integrated into the learning process, illustrating a strong fusion between high-resolution climate reanalysis data and AI-based hydrological modeling. To improve robustness and account for predictive uncertainty, an ensemble strategy is implemented, providing flood forecasts with confidence intervals. This allows for a more reliable representation of extreme events, which are particularly difficult to predict in semi-arid environments. Model performance is evaluated using hydrological metrics, including Nash–Sutcliffe Efficiency (NSE) and RMSE. Flood behavior is further characterized using key indicators such as peak discharge (Qmax) and high-flow percentiles (Q95, Q75). Results show that the LSTM model achieves good short-term forecasting performance, with NSE values exceeding 0.70 at a one-day lead time. However, its performance declines at longer horizons (seven-day lead time), with underestimation of flood peak magnitudes during extreme events. In contrast, the GRU model maintains stable performance across all forecasting horizons, achieving higher NSE values even at seven days and significantly improving flood peak detection and magnitude estimation. Beyond predictive performance, this study lays the foundation for a future operational Early Warning System aimed at supporting disaster hazard reduction by transforming model outputs into accessible alerts, thereby enhancing preparedness and community resilience. Overall, the proposed framework demonstrates that integrating AI models, hydrometeorological indicators, and high-resolution climate reanalysis products significantly improves flood prediction in data-scarce semi-arid catchments.

Keywords: Forecasting, Flood hazards, Climate reanalysis, .AI, GRU