EGU24-13359, updated on 12 Feb 2025
https://doi.org/10.5194/egusphere-egu24-13359
EGU General Assembly 2024

Machine learning models for stream-level predictions using readings from satellite and ground gauging stations

Cristiane Girotto[1], Farzad Piadeh[2], Kourosh Behzadian[1,3], Massoud Zolgharni[1], Luiza Campos[3], and Albert Chen[4]

[1]University of West London, School of Computing and Engineering, London, UK
[2]Centre for Engineering research, School of Physics, Engineering and Computer Science, University of Hertfordshire, Hatfield, UK
[3]Dept of Civil, Environmental and Geomatic Engineering, University College London, London, UK
[4]College of Engineering, Mathematics and Physical Sciences, University of Exeter, Exeter, UK

Abstract

While the accuracy of flood predictions is likely to improve with increasing gauging station networks and robust radar coverage, challenges arise when such sources are spatially limited [1]. For instance, severe rainfall events in the UK come mostly from the North Atlantic area where gauges are ineffective and radar instruments are limited to it 250km range. In these cases, NASA's IMERG is an alternative source of precipitation estimates offering global coverage with 0.1-degree spatial resolution at 30-minute intervals. The IMERG estimates for the UK's case can offer an opportunity to extend the zone of rainfall detection beyond the radar range and increase lead time on flood risk predictions [2].

This study investigates the ability of machine learning (ML) models to capture the patterns between rainfall and stream level, observed during 20 years in the River Crane in the UK. To compare performances, the models use two sources of rainfall data as input for stream level prediction, the IMERG final run estimates and rain gauge readings. Among the three IMERG products (early, late, and final), the final run was selected for this study due to its higher accuracy in rainfall estimates. The rainfall data was retrieved from rain gauges and the pixel in the IMERG dataset grid closest to the point where stream level readings were taken.

These datasets were assessed regarding their correlation with stream level using cross-correlation analysis. The assessment revealed a small variance in the lags and correlation coefficients between the stream-level and the IMERG dataset compared to the lags and coefficients found between stream-level and the gauge's datasets. To evaluate and compare the performance of each dataset as input in ML models for stream-level predictions, three models were selected: NARX, LSTM, and GRU. Both inputs performed well in the NARX model and produced stream-level predictions of high precision with MSE equal to 1.5×10^{-5} while using gauge data and 1.9×10^{-5} for the IMERG data. The LSTM model also produced good predictions, however, the MSE was considerably higher, MSE of 1.8×10^{-3} for gauging data and 4.9×10^{-3} for IMERG data. Similar

performance was observed in the GRU predictions with MSE of 1.9×10^{-3} for gauging data and 5.6×10^{-3} for IMERG. Nonetheless, the results of all models are within acceptable ranges of efficacy confirming the applicability of ML models on stream-level prediction based just on rainfall and stream-level information. More importantly, the small difference between the results obtained from IMERG estimates and gauging data seems promising for future tests of IMERG rainfall data sourced from other pixels of the dataset's grid and to explore the potential for increased lead time of predictions.

[1] Piadeh, F., Behzadian, K. and Alani, A. (2022). A critical review of real-time modelling of flood forecasting in urban drainage systems. *Journal of Hydrology*, 607, p.127476.

[2] Foelsche, U., Kirchengast, G., Fuchsberger, J., Tan, J., Petersen, W. (2017). Evaluation of GPM IMERG Early, Late, and Final rainfall estimates using WegenerNet gauge data in southeastern Austria. *Hydrology and Earth System Sciences*, 21(12), pp. 6559-6572.