P34

Stormwater management modeling and machine learning for flash flood susceptibility prediction in Wadi Qows, Saudi Arabia

○Fahad Alamoudi, Mohamed Saber, Sameh Kantoush, Tayeb Boulmaiz, Karim Abdrabo, Tetsuya Sumi

Predicting flash flood-prone areas is essential for proactive disaster management. This study aims to compare machine learning models (random forest, LightGBM, and CatBoost) and the Personal Computer Storm Water Management Model (PCSWMM) hydrological model to predict flash flood susceptibility (FFS) in an arid region (Wadi Qows in Saudi Arabia). Eleven flood-controlling factors were identified and analyzed for their relative importance in forecasting flood occurrences based on considerable literature research. Approximately 300 flash flood sites were identified through a post-flood survey after the extreme flash floods of 2009 in Jeddah city.

Machine learning techniques are widely applicable in water-related applications; for instance, these methods can accurately predict flash flood susceptibility in arid regions (Saber et al. 2021). Flash floods are becoming more common as a result of changes in violent storm patterns and global climate change (Hirabayashi et al. 2013). Flash flood susceptibility Mapping is one of the most critical metrics according to researchers and governments throughout the globe (Ali et al. 2020). Over the past few decades, the frequency of extreme flood events has increased in the MENA region (Abdrabo, Saber, et al. 2022). The last two decades, flash floods in Saudi Arabia have increased. For instance, flash floods occurred in Jeddah city in 2009 and 2011. The number of "Jeddah drowning" victims reached 113 in 2009 (Youssef et al. 2016). Therefore, in this study, we compare machine learning models (random forest, LightGBM, and CatBoost) and the PCSWMM model to predict flash flood susceptibility (FFS) in an arid region.

The methodology of this study consists of two main parts. In the first part, we use machine learning techniques to perform FFS mapping, and in the second part, we use the PCSWMM to obtain a flood inundation map. First, a flood inventory map (Figures 1a, b) is generated based on 300 inundated sites. These places were determined based mostly on post-flood.
Topographic, hydro-logical, geological, and landform variables were all considered. We used elevation, aspect, slope, hillshade, flow accumulation, horizontal flow distance, vertical flow distance, stream power index, rainfall, land use/land cover, and topographic wetness index to analyze the linear connection between the FFSFs and other variables. With the help of a random selection strategy, the dataset was split in two: 70% for training and 30% for testing.

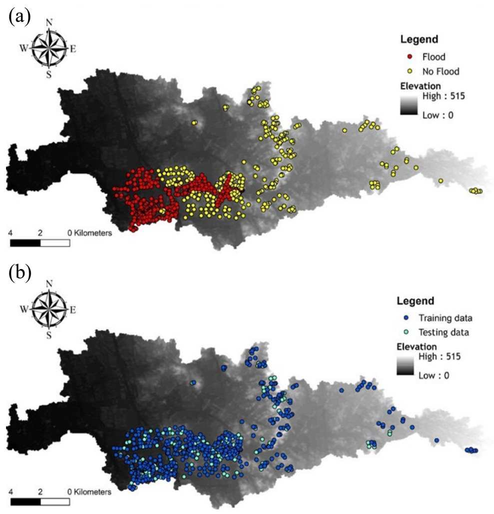

Figure 1. (a) flooded and nonflooded locations, and (b) the flood inventory map used to construct the training and testing datasets.

The assessment system of measurement of the newly assessed algorithms (ML) validated their high overall performances when predicting flooding in an arid environment. Accordingly, those techniques were employed to estimate flood susceptibility maps for Wadi Qows. The three FSMs developed using these three ML (CatBoost, LightGBM, and RF) techniques were then compared with the flood inundation map obtained with the PCSWMM, as shown in Figure 3. The FFSMs developed by the ML methods show reasonable spatial distributions. They agree well with the actual situation after 2009 and 2011 in Jeddah city, revealing that the model efficiently predicted flash floods in the study area. Different flood levels, from low to very high, are categorized to show the difference in the FFSMs. Most downstream areas dominated by high populations are affected mainly by high and very high flood levels (Figure 2). Figures 2 and 3 indicate that the FFSMs developed by these three models are comparable with the PCSWMM results. The flood susceptibility levels developed by the two methods (ML and PCSWM) show acceptable agreement. For instance, the low-flood results agree well. Nevertheless, other flood levels, such as the estimated areas of high and extremely high flood risk zones, comprise 40% (RF), 43% (CatBoost), and 41% (LightGBM) of the study area, larger than the same area indicated in the flood inundation map developed by PCSWMM at approximately 24%.

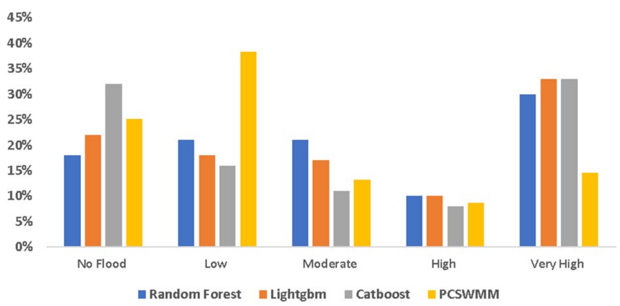

Figure 2. Affected area of the flood susceptibility with three ML methods and flood inundation map of PCSWMM model.

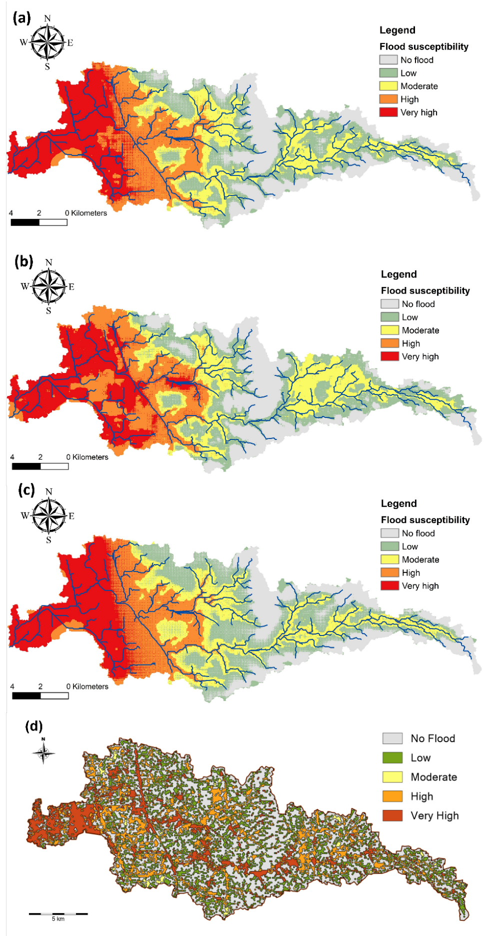

Figure 3. Flood susceptibility maps by LightGBM (a), Random Forest (b), CatBoost (c). PCSWMM (d).

Therefore, the PCSWMM should be further calibrated to be perfectly with the ML methods. However, the results are reliable and acceptable, especially concerning the spatial coverage of the flooding levels.

References

1. Abdrabo et. al., 2022, Applications of Remote Sensing for Flood Inundation Mapping at Urban Areas in MENA Region: Case Studies of Five Egyptian Cities. 307–30.

2. Ali, et. al., 2020, GIS-Based Comparative Assessment of Flood Susceptibility Mapping Using Hybrid Multi-Criteria Decision-Making Approach, Naïve Bayes Tree, Bivariate Statistics and Logistic Regression: A Case of Topľa Basin, Slovakia. Ecological Indicators 117: 106620

3. Saber et. al., 2021, Examining LightGBM and CatBoost Models for Wadi Flash Flood Susceptibility Prediction. Geocarto International, 1.

4. Hirabayashi et. al., 2013, Global flood risk under climate change. Nature Climate Change. 3: 816–2.