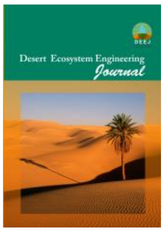
DEEJ (2025) 14 (1) 1-2

University of Kashan

Desert Ecosystem Engineering Journal

Journal homepage: http://deej.kashanu.ac.ir



Zoning of Flash Flood Generation Probability Using Machine Learning Techniques (A Case Study: Karaj Dam Watershed, Iran)

Abdolvahed Kordi,[1] Amir Sadoddin,[2]* Gahangir Mohammadi,[3] VahedBerdi Sheikh,[4] Omid Asadi Nalivan[5]

Received: 17/04/2025 Accepted: 13/07/2025

Extended Abstract

Introduction: Flooding stands as a major hydro-climatic extreme event and one of the most severe natural disasters, posing significant threats to socio-environmental systems. Given its multifaceted and dynamic nature, floods are influenced by numerous factors, with both natural and anthropogenic agents collectively impacting their magnitude, intensity, extent, and duration. A fundamental step in managing flood-prone areas is to identify the spatial extent of potential floods. By spatially analyzing and investigating floodplains, a risk-based management framework can be established, leading to effective preventive measures. Therefore, flood prevention and management necessitate the initial delineation of high flood-potential areas using modeling methods, followed by the identification of the underlying drivers of these events. In recent years, Iran has experienced an intensifying frequency and impact of flash floods, resulting in considerable damage. The Karaj Dam Watershed in Alborz Province, in particular, has been severely affected, creating substantial challenges for travelers, tourists, and local residents. This ongoing issue underscores the urgent need for enhanced flood management strategies. Consequently, this research focuses on the spatial analysis and susceptibility assessment of flash flood-prone areas. Our aim is to provide essential and practical insights to inform planning, decision-making, and the implementation of various flood mitigation measures by relevant authorities and stakeholders.

Materials and Methods: To investigate the probability of flash flood generation across the Karaj Dam Watershed in Alborz Province, we first compiled a comprehensive map of past flash flood events. This involved field visits, consultation with local resources, and integrating data from various sources within the ArcGIS 10.8 software environment. Based on an extensive review of relevant literature, we identified fourteen variables hypothesized to influence flash flood occurrence. To ensure an optimal combination of predictors, we applied the Variant Inflation Factor (VIF) test for feature selection, screening out redundant variables. For model development and validation, the flash flood event data was divided using a 10-fold cross-validation method, with 70% allocated for training and 30% for validation. We then processed and analyzed the spatial data, including environmental variables and hydrologic information, using three distinct machine learning algorithms: Maximum Entropy, Random Forest, and

1. PhD Student, Department of Watershed Management, University of Agricultural Sciences and Natural Resources, Gorgan, Iran. E-mail: Vahed.kordi@yahoo.com

2. Corresponding Author, Professor, Department of Watershed Management, University of Agricultural Sciences and Natural Resources, Gorgan, Iran. E-mail: amir.sadoddin@gau.ac.ir

3. Associate Prof., Department of Forest, University of Agricultural Sciences and Natural Resources, Gorgan, Iran. E-mail: Mohamadi.jahangir@gau.ac.ir

4. Professor, Department of Watershed Management, University of Agricultural Sciences and Natural Resources, Gorgan, Iran. E-mail: V.sheikh@yahoo.com

5. Assistant Prof., Department of Water Engineering, Faculty of Agriculture, University of Maragheh. Iran. E-mail: O.asadi@maragheh.ac.ir

DOI: 10.22052/deej.2025.256422.1098

Support Vector Machine. The accuracy of each susceptibility model was rigorously evaluated using comprehensive criteria, including the Relative Operating Characteristic (ROC) curve and the Kappa index. The model demonstrating the most acceptable accuracy was subsequently selected to generate a spatial probability map of flash flood generation for the study area. The insights derived from this research offer valuable contributions to flood hazard assessment and management efforts.

Results: Our analysis revealed that the Support Vector Machine (SVM) model proved to be the most suitable algorithm for determining flash flood susceptibility within the study area. This model achieved an impressive Area Under the Curve (AUC) value of 0.88 and a Kappa coefficient of 0.68, demonstrating strong predictive capability and outperforming both the Maximum Entropy and Random Forest models. The resulting flood susceptibility map indicates that 53% of the study area, approximately 46,200 hectares, exhibits moderate, high, or very high flash flood potential. Conversely, the remaining 47% of the watershed, about 38,860 hectares, falls into the low and very low flood probability classes. Furthermore, our research identified the key contributing factors to flash flooding in the region. In order of significance, elevation, drainage density, slope, and average annual rainfall emerged as the most important drivers influencing flood occurrence.

Discussion and Conclusions: Historically, the Karaj Dam Watershed has experienced a unique interplay of climatic and geomorphological factors that amplify flash flood risk. Extreme cold conditions at higher altitudes lead to substantial winter snow accumulation. When this is followed by heavy rainfall in spring and summer, it causes a rapid increase in water flow, resulting in significant flash flooding. This aligns perfectly with the watershed's natural characteristics: its steep slopes limit the soil's capacity to absorb rainwater, and when combined with high drainage density and intense, short-duration rainfall, these factors significantly elevate the risk of flash flooding.
The findings from this research are crucial for both improving our understanding of flood phenomena and for practically identifying flood-prone areas. By pinpointing the key contributing factors to flash flood occurrences, we can prioritize management measures tailored to the region's specific conditions, ultimately aiming to reduce flood-related damages. The generated flood susceptibility map serves as a vital tool, providing a roadmap for implementing effective preventive measures. Ultimately, this research offers an effective framework for developing robust flood hazard management plans. It can significantly contribute to reducing human losses and economic damages by guiding informed decision-making and planning processes. We strongly recommend that responsible organizations and agencies utilize these insights, including the detailed zoning of flood susceptibility and the identified effective factors, to optimize financial resource allocation and ensure efficient use of public resources in flood mitigation efforts.

Keywords: Data mining, Flood occurrence probability, Flood generation, Flood zonation.