The role of predisposing factors in determining the rainfall intensity necessary to cause flash floods in Portugal

Caio Villaça[1] (caiovnv10@gmail.com), José Luís Zêzere[1,2], Pedro P. Santos[1,2]
[1]Centre for Geographical Studies, Institute of Geography and Spatial Planning, University of Lisbon
[2]Associated Laboratory TERRA

This presentation participates in OSPP – Outstanding Student & PhD candidate Presentation contest

Sharing is encouraged

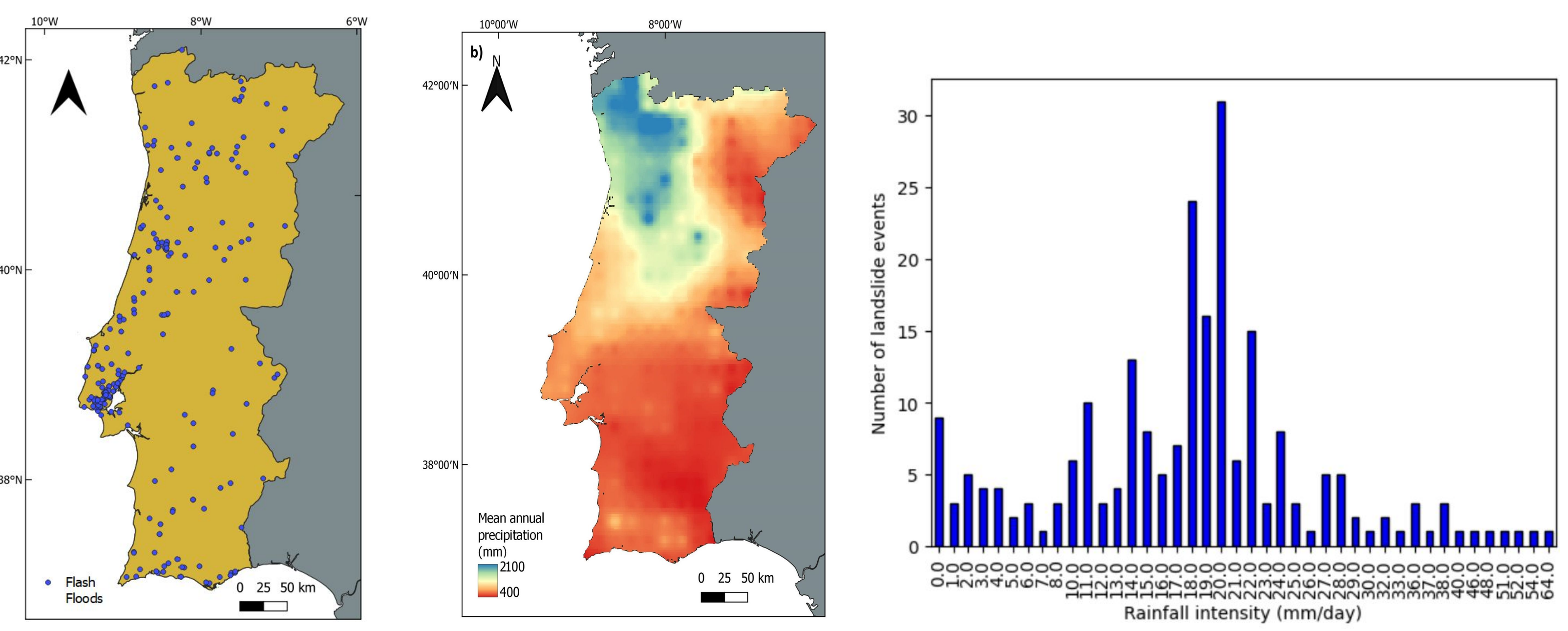
Introduction and Goals

Flash floods are often responsible for deaths and damage to infrastructure in many parts of the world. Modelling flash floods occurrence is important to try to mitigate their effects. The general objective of this work is to analyze the correlation between predisposing factors and the rainfall intensity threshold responsible for triggering flash floods and evaluate the spatial variation of the rainfall intensity threshold in the continental area of Portugal.

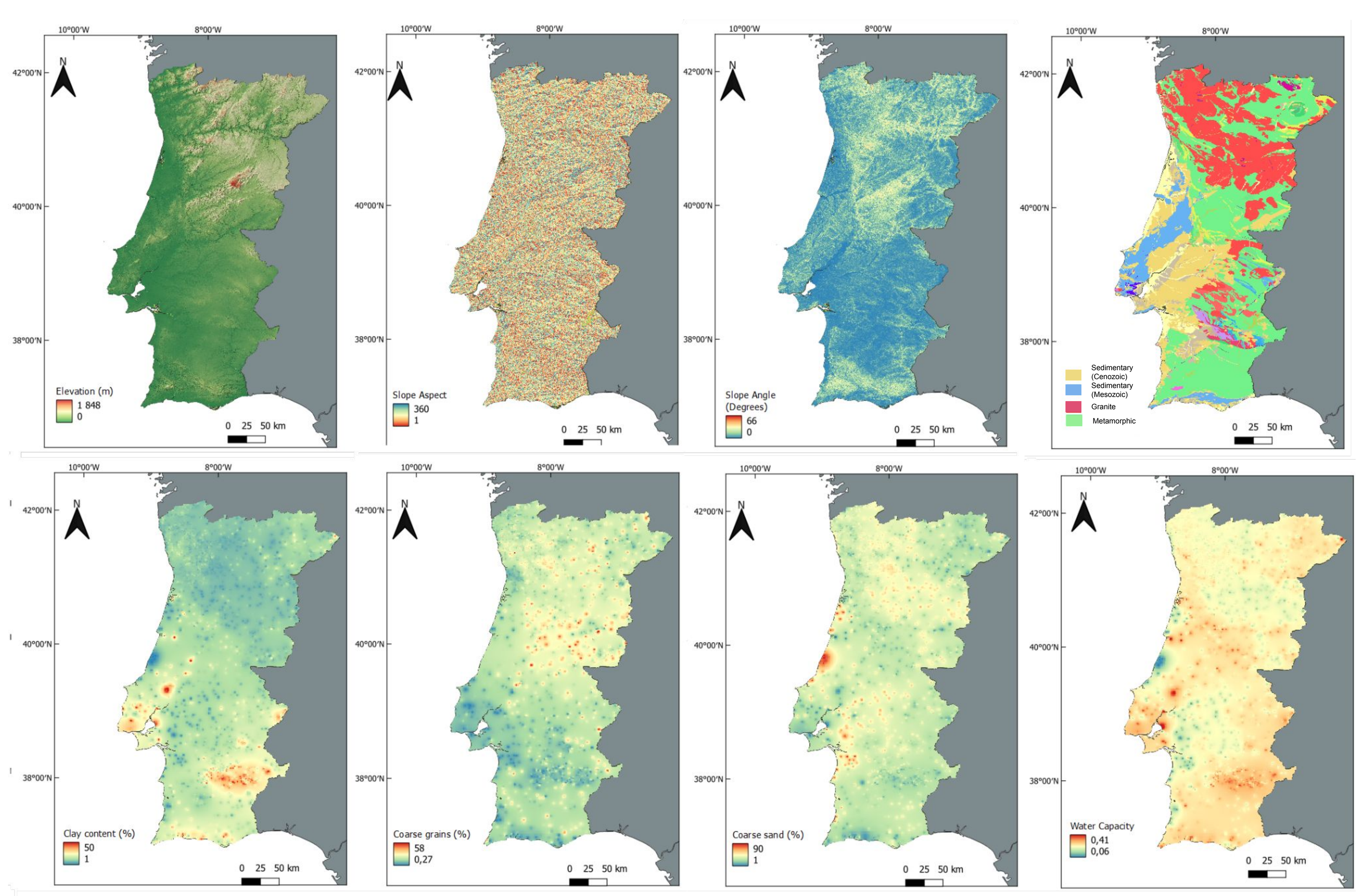
Methodology

The area of study is the continental region of Portugal. The flash flood inventory used in this work was extracted from the database DISASTER (Zêzere et al., 2014) built with information found on newspapers and only considers events that caused any disturbance in people's lives. The databases contain information about the day and coordinates of 261 occurrences (the highest precision is associated with the 1:1000 scale and the lowest with the 1:25000 scale).

Historical daily rainfall data was collected automatically from the Copernicus database (E-OBS daily gridded meteorological data for Europe) from the year 1950 to 2022. We extracted the accumulated precipitation for 3 days preceding each event and calculated the rainfall intensity. Due to the highly imbalanced data, only occurrences with rainfall intensity between 10 and 20 mm/day were used.

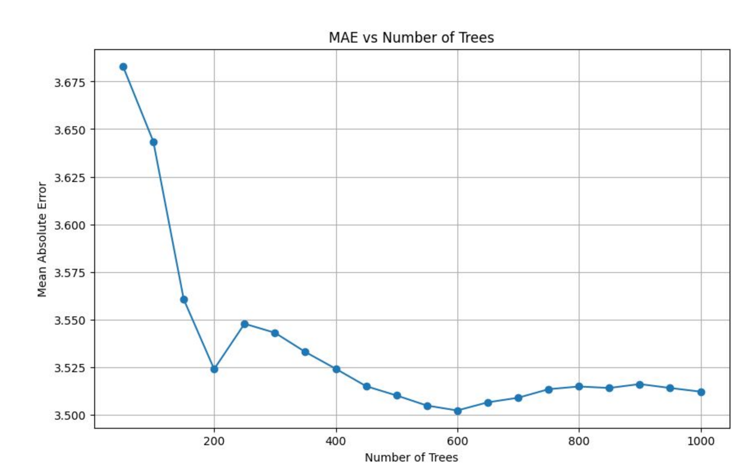


The predisposing factors were extracted considering the whole basin that corresponds to each flood event (Santos and Reis, 2018; Santos et al., 2019). The analyzed predisposing factors were: I) Morphology (accumulated flow, slope angle, average elevation, slope aspect, topographic wetness index (TWI) and stream power index (SPI)) collected using a DEM - 30 m/px; II) Lithology collected from the LNEG Geological map (1:500.000); III) Soil properties (percentage of clay, coarse sand and coarse elements and field capacity) collected from the INFOSOLO database.



The Random Forest algorithm and the K-fold cross-validation technique were used to evaluate the model's performance and create a final model that identifies the relationship between the predisposing factors (independent variables) and the different rainfall intensities (dependent variable) related to each flash flood occurrence.

The data was divided into a training set (80%) and a test set (20%). K = 10 was used to evaluate the training set performance, exploring various numbers of trees to determine the optimal parameter value. The number of trees were defined as 500 since it is the value where the errors start stabilizing.



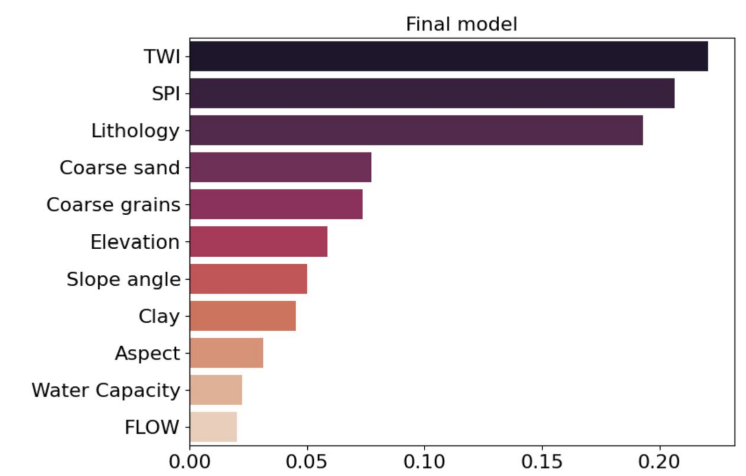
Results

Accuracy:

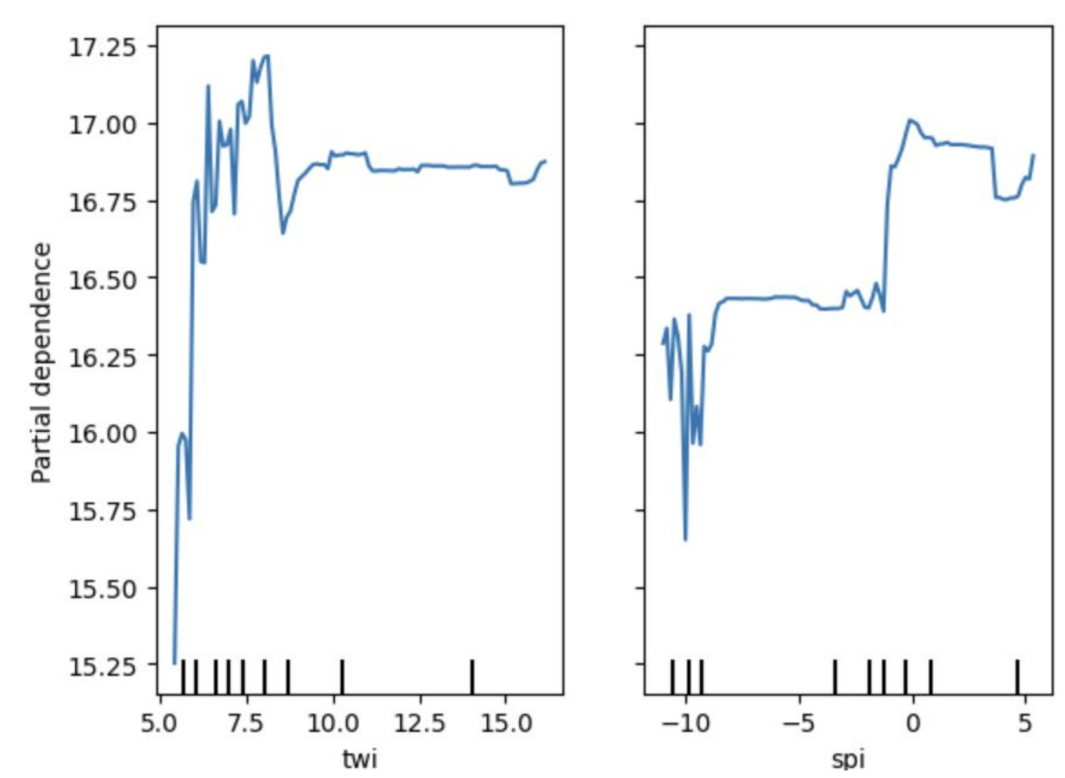
The final model obtained a root mean square error (RMSE) value of 3, a mean absolute error (MAE) of 2.6 , and a mean absolute percentage error (MAPE) of 17%. The accuracy is acceptable for the objectives of the work.

Feature ranking:

The TWI, SPI and Lithology were defined as the most important factors in the model for defining the amount of rainfall needed for flash floods to occur in mainland Portugal. Flow accumulation, field capacity and slope aspect were classified as the lowest important features for the model



It was also analysed the partial dependence (PDP) of the two most important features of the final model. The PDP show how the dependent variable varies according to the change of one feature while the others are fixed. It is shown by the PDP that the value of the two most important features change proportionally (contributes positively) to the dependent variable for the most part.

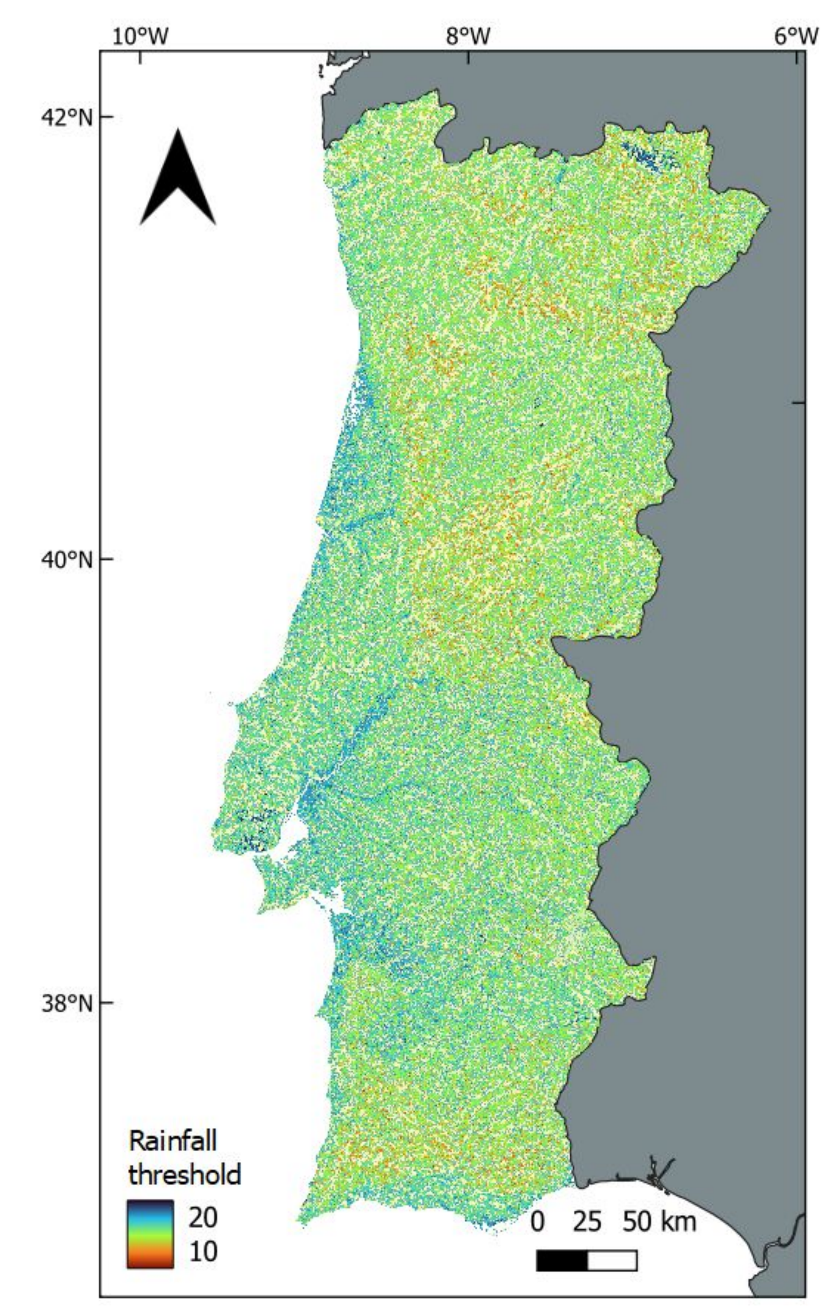


Discussions and Conclusions

This study elaborated RF models to evaluate the influence of each predisposing factor in assessing the rainfall intensity threshold for flash floods and to analyse the spatial variation of the rainfall intensity thresholds across mainland Portugal. Is it not possible to conclude with absolute certainty that the predisposing factors ranked by RF are the ones conditioning the spatial distribution of rainfall intensity thresholds (Stegar et al., 2021), but the model analysed on this work provides some insights of possible correlations. Given the constraints of this study, particularly the limited dataset available for model development, we recommend several future validation measures to bolster our findings. These include acquiring an expanded dataset on flash floods, employing alternative machine learning or statistical methodologies, and applying our approach to different databases and geographical regions. Moreover, the accuracy and validation of the prediction values confirm that the derived map is effectively reliable. The main achievement of this work are as follows:

1. The flash flood predisposing factors that were more useful to assess the rainfall intensity threshold in Portugal are the TWI, SPI and lithology.
2. Medium rainfall intensity represents most part of the continental region while high and low rainfall intensities are concentrated in some areas.

The final map represents a detailed distribution of the rainfall intensity threshold for flash floods in Portugal. The heterogeneous distribution of rainfall intensity values across the map is the key to identify the role of the predisposing factors in quantifying the rainfall intensity that triggers the flash floods. These results show the complexity of the terrain and how a variety of factors can influence the occurrence of flash floods. It should be highlighted that the final map do not analyse flood susceptibility. Each pixel's rainfall intensity value acts as a critical threshold, whereby surpassing this threshold with observed rainfall is likely to trigger flash floods.



References:
Santos, P. P., & Reis, E. (2018). Assessment of stream flood susceptibility: a cross-analysis between model results and flood losses. *Journal of Flood Risk Management*, *11*, S1038-S1050.
Santos, P. P., Reis, E., Pereira, S., & Santos, M. (2019). A flood susceptibility model at the national scale based on multicriteria analysis. Science of the total environment, 667, 325-337.
Steger, S., Mair, V., Kofler, C., Pittore, M., Zebisch, M., & Schneiderbauer, S. (2021). Correlation does not imply geomorphic causation in data-driven landslide susceptibility modelling–Benefits of exploring landslide data collection effects. *Science of the total environment*, *776*, 145935.
Zêzere, J. L., Pereira, S., Tavares, A. O., Bateira, C., Trigo, R. M., Quaresma, I., ... & Verde, J. (2014). DISASTER: a GIS database on hydro-geomorphologic disasters in Portugal. *Natural hazards*, *72*, 503-532.

This work is financed by National Funds through the Portuguese funding agency, FCT - Fundação para a Ciência e a Tecnologia within project: 2022.14473.BD