Multi-Hazard Susceptibility Mapping in Europe Using Deep Learning

submitted in partial fulfillment for the degree of master of science

Jeremy Palmerio
15788792

Master Information Studies
data science
Faculty of Science
University of Amsterdam

Submitted on DD.MM.YYYY

	UvA Supervisor	External Supervisor
Title, Name	Maarten Marx	Timothy Tiggeloven
Affiliation	UvA Supervisor	External Supervisor (VU)
Email	M.J.Marx@uva.nl	timothy.tiggeloven@vu.nl

ABSTRACT

GITHUB REPOSITORY

1 INTRODUCTION

2 RELATED WORK

3 METHODOLOGY

The overarching goal of this research is to create a multi-hazard susceptibility maps of the European continent. To achieve this, single-hazard susceptibility maps are first generated using statistical and deep learning methods, on which a comprehensive multi-hazard map can be trained. These models will be trained on pre-engineered features and optimized using state-of-the-art hyperparameter tuning. The MYRIAD-HES will be used a ground truth for the final map and will be validated using performance metrics such as AUC, F1-Score and MAE. Interpretability techniques like SHAP values will be used to analyze feature importance and identify key susceptibility patterns across Europe, supporting disaster preparedness efforts.

This study presents a novel approach to hazard susceptibility mapping by combining high-resolution spatial data ($300m^2$) with large-scale coverage across the entire European continent. Unlike many previous studies which are limited to national or regional scales, this work applies convolutional neural networks (CNNs) with spatial attention mechanisms to learn fine-grained susceptibility patterns across diverse geographic and climatic zones. The integration of spatial attention blocks enhances interpretability and allows the model to focus on the most relevant environmental features within each local context.

3.1 Data and Resources

The single-hazard maps will be generated from feature engineered raw data, which has been collected by the European Union's Copernicus Programme [1]. The feature engineering will follow the methodology from Tiggeloven et al. and and will be generated from topological and environmental data including Digital Elevation Maps (DEM), roads, rivers, coastlines, faultlines, soil, lithology, landcover, Normalized Difference Vegetation Index (NDVI), wind, temperature and precipitation maps. The feature engineered data, as well as the hand-labeled ground truths for the deep learning approaches will be provided by the project supervisor and the IVM. All input maps—referred to as signals—are consistent in shape and resolution, with dimensions of 15,560 × 26,650 pixels, covering the area from 25°W to 46°E and from 27°N to 73°N.

The multi-hazard map will use the single-hazard maps for training and validate against the primary ground truth provided by the MYRIAD dataset, which covers eleven different hazards (e.g., floods, droughts, landslides, earthquakes) for the period 2004–2017.

tabularx booktabs

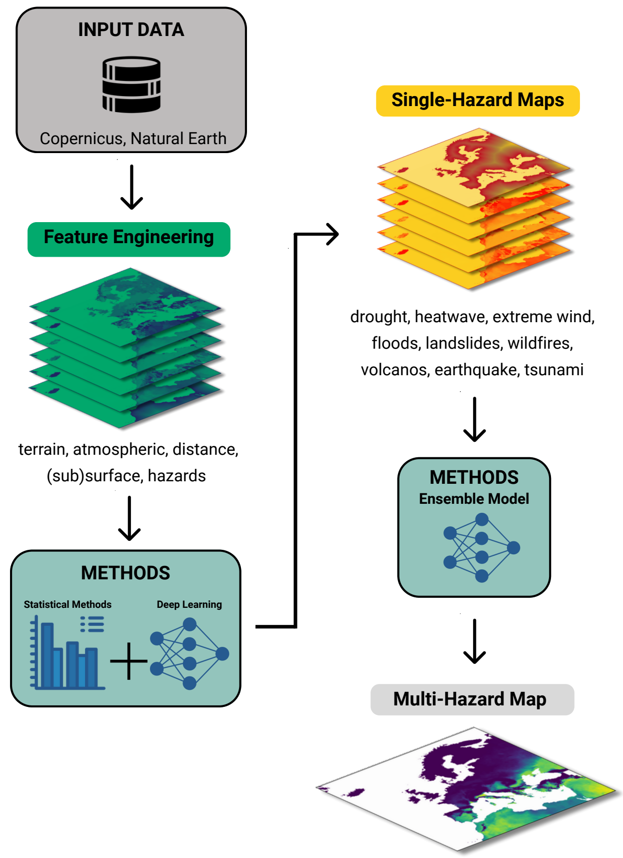

Figure 1: High level methodology overview. Various Copernicus and Natural Earth data is pre-processed into consistent size and resolution feature maps as described in table 1. Some hazard are mapped statistically, while floods, landslides and wildfires are mapped with a deep learning approach described in Section 3.4 Finally the hazard maps obtained are used to train a simple MLP ensemble model for the final multi-hazard map.

3.2 Data Pre-processing and Exploratory Data Analysis (EDA)

Although feature engineering has been done, the datasets will be verified for consistency and completeness. Features will be normalized and resampled to a common resolution of 10 arc-seconds (3̃00m a the equator). Additionally, a thorough EDA will be conducted to identify residual inconsistencies or outliers, and to examine the distribution and inter-correlations of the pre-engineered features.

3.3 Statistical Individual Hazard Modeling

Following the methodology from the Japan study, statistical methods will be used for extreme wind, droughts, heatwaves and wildfire

Signal	Description
accuflux	Accumulated water flux derived from topographic gradients.
aspect	Compass direction that a slope faces.
coastlines	Distance to the nearest coastline.
curvature	Curvature profile of the terrain, showing how slope changes across the landscape.
elevation	Height above sea level.
fire_weather	Composite index reflecting weather conditions favorable for wildfires.
GEM	Global Earthquake Model or related seismic hazard data.
GLIM	Global Lithological Map, providing rock and soil type data.
heatwave	Occurrence or intensity of extreme heat events.
HWSD	Harmonized World Soil Database: detailed soil types and properties.
landcover	Categorical data representing land use (e.g., urban, forested).
NDVI	Normalized Difference Vegetation Index: vegetation health and density.
pga	Peak Ground Acceleration: max ground shaking during an earthquake.
precipitation_daily	Daily rainfall averages.
rivers	Distance to river.
slope	Gradient or steepness of terrain.
strahler	Strahler stream order: classifies stream hierarchy.
temperature_daily	Daily temperature averages.
wind_direction_daily	Predominant daily wind direction.
wind_speed_daily	Average daily wind speed.

Table 1: Descriptions of input variables used for single hazard susceptibility mapping

hazard susceptibility maps. Earthquake and volcanoes are dealt with based on occurrence and distance metrics.

3.3.1 Extreme Wind. Extreme wind can be statistically mapped by choosing a threshold (Beaufort scale 6, 10.8m/s) and summing the days with a wind speed above said threshold. A normalization can be applied to obtain a usable range.

3.3.2 Droughts. Droughts can be modeled using the Standardized Precipitation Evaporation Index (SPEI, Vicente-Serrano et al., 2010), and normalizing for consistency.

3.3.3 Heatwaves. A heatwave signal can be generated from consecutive periods of high temperatures. Given a threshold (95%), periods above this can be flagged as heatwaves. Similarly to the extreme wind signal, heatwave periods can be statically summed and normalized for each cell.

3.3.4 Tsunamis. The tsunami signal is not considered for the European analysis, as there is relatively no incidence of these hazard on the continent.

3.3.5 Earthquake. Though earthquakes are uncommon in Europe, the east of the region is susceptible. The susceptibility can be generated from distance to fault lines, and based on historical data.

3.3.6 Volcanoes. Volcanic susceptibility can be modeled by identifying historical volcanically active sites and estimating the affect area scaled by the Volcanic Explosivity Index (VEI) following Shi Kasperson (2015)

3.4 Deep Learning Framework for Individual Hazard Mapping

Floods, wildfires, and landslides require more complex modelling to capture their underlying spatial dynamics. For these hazards, convolutional neural networks (CNNs) are used to learn spatial patterns from the input features in a classification setup. Each hazard is modelled independently, and the resulting susceptibility maps are combined in an ensemble framework to produce the final multi-hazard map, capturing potential interactions between hazards. The single-hazard susceptibility maps will be trained per hazard, while the overall multi-hazard susceptibility map will use an ensemble approach to combine the outputs of all single-hazards and accounts for the interaction effects between various models.

3.4.1 Input and Hazard Signals. Each CNN is trained on hazard-specific input features, selected based on expert knowledge and prior literature 2. To avoid scale biases, all input features are scaled to the [0, 1] range using a MinMax scaler. The label for each input patch is derived from historical hazard occurrence data between 2008 and 2017 and binarized to indicate whether a hazard occurred at least once in that cell. This setup allows the model to learn conditions that make a location susceptible to the hazard.

Hazard	Inputs
Wildfire	temperature_daily, NDVI, landcover, elevation, wind_speed, fire_weather, (root or surface)soil_moisture
Landslide	slope, curvature, aspect, accuflux, NDVI, landcover, GLIM, GEM, HWSD
Floods	slope, curvature, aspect, accuflux, NDVI, landcover, precipitation_daily, GLIM, rivers, HWSD, pga

Table 2: Input signal for each hazard mapped using the deep learning framework. The description of variables can be found in table 1.

3.4.2 Spatial Partitioning. Framing this problem as a pixel wise classification task reveals the difficulty involved. Instead, we assume the features and labels exhibit local properties, wherein only the features in the local area of a cell affect its susceptibility to a given hazard. These local patches allow for finer grained patterns to emerge from the convolution layers, and also provide a consistent size input to the network. The training pipeline will ingest these patches of size $2n + 1$ centered around each cell as inputs and the susceptibility label as target. This method is preferred because it allows for the neighboring cells' features to contribute to the prediction of each cell.

For splitting the data into training, validation, and test sets, a spatial partitioning strategy based on national borders is used. This reduces the risk of spatial leakage and provides a realistic evaluation of the model's ability to generalise to unseen geographic regions. This approach is inspired by the partitioning strategy from Tiggeloven et al., where susceptibility models were evaluated on non-overlapping regions to mitigate spatial autocorrelation bias.

3.4.3 Architecture. The architecture used for wildfire, flood, and landslide susceptibility mapping is a lightweight Convolutional Neural Network (CNN), adapted from (Source Tiggeloven). The model takes as input a spatial patch of size 5×5 pixels centered on the target cell, where each pixel has multiple feature channels (e.g., NDVI, slope, precipitation). These patches capture local spatial patterns that influence hazard susceptibility.

The network consists of three convolutional layers with ReLU activation functions and batch normalization. Each convolution is followed by a max-pooling layer to reduce dimensionality while preserving spatial features. A spatial attention layer is introduced after the final convolution to assign weights to different spatial locations in the input patch, thereby increasing interpretability and model focus. After flattening, the feature maps are passed through a dense layer with dropout for regularization, followed by a final output layer with sigmoid activation to produce a continuous susceptibility score in the range [0, 1].

3.4.4 Loss Function. Given the goal of mapping wildfire susceptibility — defined as the probability of at least one wildfire occurrence over a historical period — the model is trained as a binary classifier. Each input patch is assigned a binary label: 1 if the central cell experienced at least one wildfire during the observed period (2008–2017), and 0 otherwise. This binarization simplifies the task into a spatial susceptibility classification problem, a standard approach in hazard mapping literature.

The model is trained using Binary Cross-Entropy (BCE) loss, which is appropriate for probabilistic binary classification tasks. BCE allows the model to output continuous susceptibility values that reflect the probability of hazard occurrence, without requiring a hard decision threshold. These continuous outputs are interpreted directly as susceptibility levels.

For comparison and extension, the framework also supports regression-based training using Mean Absolute Error (MAE) when labels are available as wildfire counts or intensities. However, for the primary model presented in this study, BCE remains the default due to its alignment with the susceptibility mapping literature.

3.4.5 Evaluation. Each CNN model is evaluated on held-out spatial regions, using country-based partitioning to split the data into training, validation, and test sets. This spatial partitioning ensures that model performance reflects generalization to entirely unseen geographic areas, rather than memorization of local patterns

The evaluation metrics include:

- Area Under the Receiver Operating Characteristic Curve (ROC-AUC): to assess the model's ability to distinguish between susceptible and non-susceptible cells across all thresholds.
- F1 Score: to balance precision and recall in binary classification.
- Mean Absolute Error (MAE): used as a calibration metric for the continuous output of the classifier.

To enhance transparency and interpretability, SHapley Additive exPlanations (SHAP) are applied post-training. SHAP values quantify the contribution of each input feature to a model prediction, both locally (per-cell) and globally (across the map), helping to identify key environmental drivers of wildfire susceptibility in different regions of Europe.

3.4.6 Hyperparameter Optimization. To improve model performance and ensure optimal architecture configuration, hyperparameter optimization was conducted using the Weights & Biases (W&B) platform. A Bayesian search strategy was implemented to systematically explore combinations of architectural and training-related hyperparameters. See table 3 Each trial was trained using

Hyperparameter	**Search Space**
Number of convolutional layers	{2, 3, 4, 5}
Number of filters per layer	{16, 32, 64, 128}
Dropout rate	Uniform distribution in [0.0, 0.5]
Learning rate	Log-uniform distribution in [1e-5, 1e-2]
Batch size	{32, 64, 128}
Optimization strategy	Bayesian search (Weights & Biases)
Evaluation metric	ROC-AUC on validation set
Number of runs	20–30 trials per hazard model

Table 3: Hyperparameter search space for wildfire susceptibility CNN model.

Binary Cross-Entropy loss and evaluated on the validation set using the AUC-ROC metric. The optimization was run for 20–30 iterations per hazard model, depending on computational budget, on the Snellius supercomputer.

Following best practices, early stopping was applied with a patience of 5 epochs, and the learning rate was scheduled to reduce on plateau. The W&B platform facilitated live tracking of training metrics and enabled visual comparison between runs to avoid overfitting and ensure stable convergence.

4 RESULTS

5 DISCUSSION

6 CONCLUSION

REFERENCES

[1] Copernicus Climate Change Service. 2019. ERA5-Land hourly data from 1950 to present. https://cds.climate.copernicus.eu/datasets/reanalysis-era5-land?tab=overview

Appendix A FIRST APPENDIX

Put your appendices here.