ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

DOTTORATO DI RICERCA IN

SCIENZE DELLA TERRA, DELLA VITA E DELL'AMBIENTE

Ciclo 38

Settore Concorsuale: 04/A3 - GEOLOGIA APPLICATA, GEOGRAFIA FISICA E GEOMORFOLOGIA

Settore Scientifico Disciplinare: GEO/05 - GEOLOGIA APPLICATA

MULTISCALE LANDSLIDE SUSCEPTIBILITY ANALYSIS AND RISK ASSESSMENT ADAPTABLE TO LINEAR INFRASTRUCTURE DESIGN USING INNOVATIVE METHODOLOGIES

Presentata da: Rodolfo Rani

Coordinatore Dottorato

Barbara Cavalazzi

Supervisore

Matteo Berti

Co-supervisore

Marco Sciarra

Esame finale anno 2026

Abstract

Landslides in complex geological settings are widespread and hazardous natural phenomena that, in the context of climate change, pose significant challenges for infrastructure and communities. Despite knowledge of triggering mechanisms and monitoring techniques, conducting comprehensive landslide susceptibility studies remains challenging. Landslides present particular risks to linear infrastructure, such as railways, due to their rigid construction and critical transport function. Susceptibility mapping could be a valuable tool during the feasibility phase of railway projects, helping identify high-risk areas and inform mitigation strategies.

This PhD project, conducted in collaboration with ITALFERR s.r.l., addresses practical challenges in landslide susceptibility analysis, proposing innovative methodologies. The research initially focuses on two aspects often overlooked in the literature: differentiating landslide types, each with unique conditioning factors and mitigation requirements; and producing outputs that are easily interpretable by decision-makers. These challenges were addressed in the Serra San Quirico area (Marche, Italy), along the Fabriano-Castel Planio railway line. Different landslide types were treated as separate datasets, producing five type-specific susceptibility maps, evaluated using AUROC metrics. The maps were then combined into an overall susceptibility map using a complementary probability approach, meanwhile assessing each type's contribution to overall risk (sensitivity). Threshold reliability was improved through an ensemble reclassification model of six methods, applying the statistical mode to define objective class boundaries. Both a statistical model (Weight of Evidence, WoE) and a machine learning model (Generalized Additive Model with Boosting, GAMB) were implemented and compared. Finally, comparison with Synthetic Aperture Radar ground motion data revealed no significant correlation, highlighting the temporal disconnect between long-term susceptibility and short-term deformation.

The catastrophic rainfall events of May 2023 in Emilia-Romagna, 14 days apart with a combined return period exceeding 500 years, triggered thousands of landslides and demanded rapid response. A multi-institutional collaboration generated a detailed inventory of 80,000 polygons from high-resolution aerial imagery, classified into five macro failure and propagation mechanisms. This inventory served as ground truth for landslide-type-specific probability modeling, integrating triggering and preparatory factors. Unlike conventional approaches that aggregate rainfall over fixed windows, rainfall was treated as a continuous time series and embedded into a Transformer Neural Network (TNN), coupled with a Dense Neural Network (DNN) for static predictors. Model interpretability was ensured using SHAP-based Expected Gradients, revealing the temporal and spatial influence of rainfall. Although event-specific, this work lays the foundation for generalized spatiotemporal landslide forecasting and contributes toward unified rainfall thresholds for early warning and susceptibility assessments under varying climate scenarios.

Finally, the PhD demonstrates that desk studies and cost-effective targeted surveys can identify areas at high risk from debris flows even where historical landslide information is lacking. The

methodology combines morphometric parameters, empirical approaches for transported volume, and publicly available exposure and vulnerability data to evaluate susceptibility, expected debris flow magnitude, and potential impact on built environments. A test application in Central Val Camonica (Italian Alps) illustrates its value for linear infrastructure planning.

In conclusion, this research represents a step toward comprehensive and reliable landslide susceptibility analysis. The use of advanced models, such as TNNs, enhances prediction quality. While the reliability of landslide inventories, the treatment of different landslide types, and the effective communication of complex outputs to decision-makers, significant for real-world applications, were addressed through a unified methodology, which includes a landslide-type-specific target analysis and an ensemble reclassification models.

Funded by the European Union - Next Generation EU, Mission 4, Component 2, Investment 3.3 (Ministerial Decree 117/2023) CUP J33C22001540009, and by ITALFERR S.p.A.

Contents

1. Chapter 1: General introduction

1.1 Preface

This first chapter of the thesis introduces the world of landslide susceptibility, presenting the general concepts, recent advancements, and real-world applications for land-use and infrastructure planning. This chapter also outlines the PhD research goals, aimed at addressing gaps in the literature and in the practical application of landslide susceptibility analyses.

Firstly, a general introduction is provided in section 1.2, starting with the definition of landslide susceptibility and tracing its historical evolution up to the current state of analyses, which now include dynamic probability computations rather than the traditional static susceptibility concept. All key elements involved in landslide susceptibility are discussed, including conditioning factors, landslide inventories, and methods for partitioning the study area.

Followed by the section 1.3, which delves into the modelling framework providing an overview but focusing more on the Statistical and Machine Learning models, describing the mostly used and innovative ones. This section also covers current evaluation techniques, discussing their advantages and limitations, and addresses concerns related to reclassifying output maps for decision-makers.

In section 1.4 a systematic and comprehensive literature review of implementations of landslide susceptibility and risk modelling for actual efficient land use planning by national government and building company. Special attention is given to linear infrastructure design, which can significantly benefit from susceptibility analyses.

Finally, section 1.5 outlines the research questions that guided this thesis and provides an overview of the following chapters, all with the aim of producing reliable, real-world applications of landslide susceptibility analyses.

1.2 Landslide Susceptibility: General Introduction

1.2.1 Definition and Historical Overview

Landslide susceptibility can be expressed as the likelihood of a landslide occurring in a specific area, given the geological and local terrain characteristics (Brabb 1972; Reichenbach et al. 2018). It is an estimate of where a landslide is more likely to occur (Guzzetti et al. 2005). Landslides can cause significant damage to infrastructure and, in unfortunate circumstances, result in loss of human life.

Alongside Early Warning Systems (EWS), which aim to avoid catastrophic losses by issuing alerts when a landslide may be triggered, landslide susceptibility provides a priori information that may help prevent such situations. While EWS are concerned about when a landslide (or other natural hazard) might occur, susceptibility focuses on where. Thus, susceptibility mapping can be a valuable tool in territorial planning, helping to avoid development in areas prone to landslides in favour of more stable locations.

Landslide susceptibility should not be confused with landslide hazards, a term that, in early literature, was often used synonymously. Thanks to Varnes' clarifications in 1984, a clearer distinction between susceptibility, hazard, and risk was established. Landslide susceptibility provides information about the spatial probability of landslide occurrence in a given area. Hazard, by contrast, includes both the spatial component (i.e. susceptibility) and temporal or intensity-related components. These may be expressed as the expected frequency, the volume involved, or the landslide's velocity (Guzzetti et al. 2005a). Although hazard is more difficult to quantify, especially due to the lack of reliable temporal data, it is essential for producing risk maps, which combine hazard information with exposure elements (e.g., buildings, roads, railways). Notably, Dahal et al. (2024) developed a reliable and sophisticated method to directly produce the hazard map.

The concept of landslide susceptibility originated in the 1970s, when Brabb of the U.S. Geological Survey (USGS) created the first landslide susceptibility map for San Mateo County, California (Brabb et al., 1972 - https://pubs.usgs.gov/mf/0360/plate-1.pdf), followed by a susceptibility map of the entire U.S. territory (Radbruch-Hall 1977). At that time, methodologies were primarily expert-based, relying on overlays of lithology, slope classes, and past landslide occurrences.

In 1976, Neuland proposed one of the first statistical approaches to landslide susceptibility, applying Principal Component Analysis (PCA), an F-Test, and bivariate analysis to identify the most influential factors. A key turning point came with Carrara et al. (1983), who established the foundations for spatial multivariate analysis in susceptibility mapping through a case study in the Calabria region (Italy).

During the 1990s and early 2000s, there was a rapid expansion of statistical methods and case studies focused on landslide susceptibility (Reichenbach et al. 2018). This growth was driven by the emergence of Geographic Information Systems (GIS) (Goodchild 2010), which enabled the integration of large spatial datasets, automated computation, and the production of high-resolution susceptibility maps (Huabin et al. 2005). Among the most influential works from this period are the study by Carrara et al. (1991), which represents one of the earliest applications of logistic regression integrated into GIS, and the work by Atkinson and Massari (1998), who developed a bivariate analysis using Generalised Linear Models (GLM). In the same years, Aleotti and Chowdhury (1999) conducted a comprehensive review of existing models, highlighting in particular the emerging potential of Neural Networks. Similarly, Guzzetti (1999) carried out a critical evaluation of susceptibility modelling approaches, with a particular focus on the implications of different mapping units.

Moving to the 2000s, the Ayalew and Yamagishi (2005) work demonstrated the capability of logistic regression analysis, establishing it as the most widely implemented method during that period (Reichenbach et al. 2018). The 2005–2015 decade marked a transition from regression-based models to hybrid approaches incorporating machine learning techniques, as in the study by Pradhan and Lee (2009), which compared neural networks with logistic regression, and Tsangaratos et al. (2017), who applied a random forest model. The majority of landslide susceptibility models developed after 2020 integrate machine learning algorithms or ensemble

approaches combining statistical and data-driven methods (Ado et al. 2022, Alvioli et al. 2024). Recently, the literature focuses on how to examine the resulting susceptibility from a ML model applying Explainable Artificial Intelligence (XAI) methods (Gunning & Aha 2019), in order to understand the importance of each conditioning factors (Dahal & Lombardo 2023).

In Figure 1.1, the entire and common landslide susceptibility workflow is conceptualised. In summary, it usually starts with the collection of the landslide information of the study area, then the factors that influence the landslide presence, divided into static, preparatory and triggering ones. From the data, it moves to the modelling part, which can be held by physically based models or data-driven ones, which can be statistically based or machine learning. The resulting susceptibility map is generated, with values that range from 0 to 1 (indicating the probability of occurrence of a landslide). This result is usually evaluated with several metrics using the original landslide inventory as a ground truth for comparison. Then, to obtain a clear visualisation and usable map, the susceptibility map is reclassified into classes following the reclassification methods.

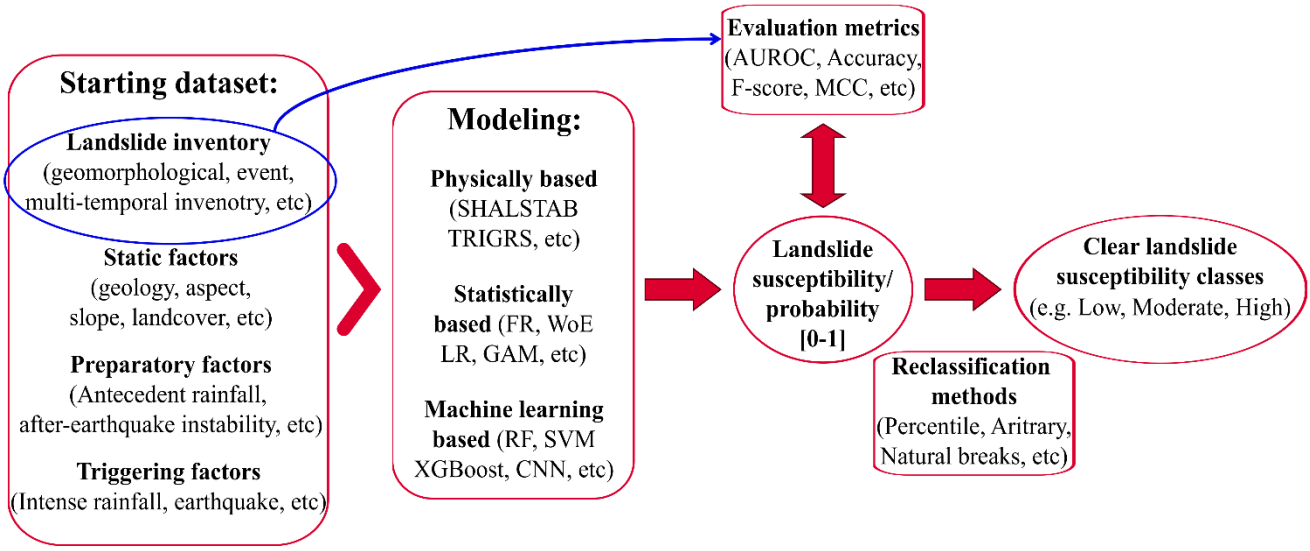


Figure 1.1. Conceptualisation of the entire workflow of landslide susceptibility/probability modelling. For each step (dataset, model and methods) are also indicate some general used example in the literature.

1.2.2 Conditioning Factors

The conditioning factors are all the geological, topographical, and environmental information of the terrain that can contribute or somehow being related to landslides. In simple words they are the independent or explanatory variables (Ayalew & Yamagishi 2005, Budimir et al. 2015), which can be called predictors or features inside the ML context.

They can be divided in three main categories based on their temporal evolution, visualized in Figure 1.2 (Steger et al. 2024, Berti et al. 2025b): the predisposing factors which are considered invariable on the observation scale (e.g. geology, elevation, slope, landcover); the preparatory factors that show changes or cyclical trends during the same period (e.g. NDVI, soil moisture, predecessor rainfall, after-earthquake instability); the triggering factors which are considered

the processes that acts in a very short and well-defined time and actually initiating factor (e.g. intense rainfall, earthquake).

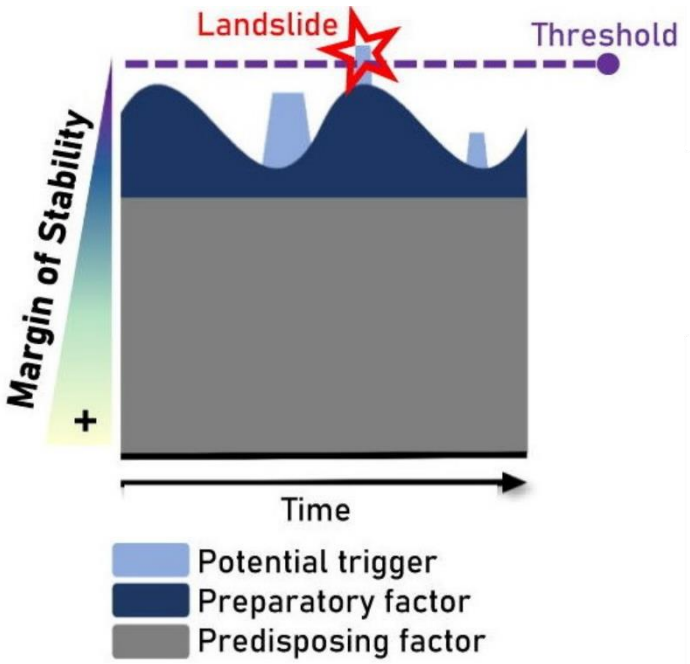


Figure 1.2. Landsliding is initiated ("unstable") due to an interplay of static predisposing factors and dynamic preparatory and triggering factors. (Steger et al. 2024, modified).

The predisposing factors define the long-term likelihood of landslide formation based on terrain characteristics and remain unchanged regardless of triggering conditions. Predisposing factors are usually grouped with preparatory ones to form the so-called static factors, visualised in Figure 1.3. Commonly used static variables could be grouped into 5 categories (Reichenbach et al. 2018):

- Topographic/morphological factors: slope angle, aspect, curvature, elevation, slope length, topographic wetness index (TWI), stream power index (SPI), relief, and terrain ruggedness.
- Geological-structural factors: lithology, rock strength, soil depth, proximity to faults or lineaments, drainage density.
- Land-cover and environmental factors – normalized difference vegetation index (NDVI), land use/land cover (LULC), soil texture, geomorphologic units, and proximity to rivers.
- Hydrological factors: groundwater table, soil moisture, soil hydraulic conductivity, and depth to bedrock, less frequently used, but gaining interest (Parker et al. 2016).
- Anthropogenic factors: distance to roads, infrastructure density, urbanization, engineering activities, and excavation sites.

Regarding the triggering factors they can be expressed as:

- Natural triggers: rainfall amount/intensity/duration, cumulative precipitation, seismic activity (peak ground acceleration), snowmelt, and volcanic activity.

- Anthropogenic triggers: road cuttings, construction, deforestation, mining, irrigation, and land management practices.

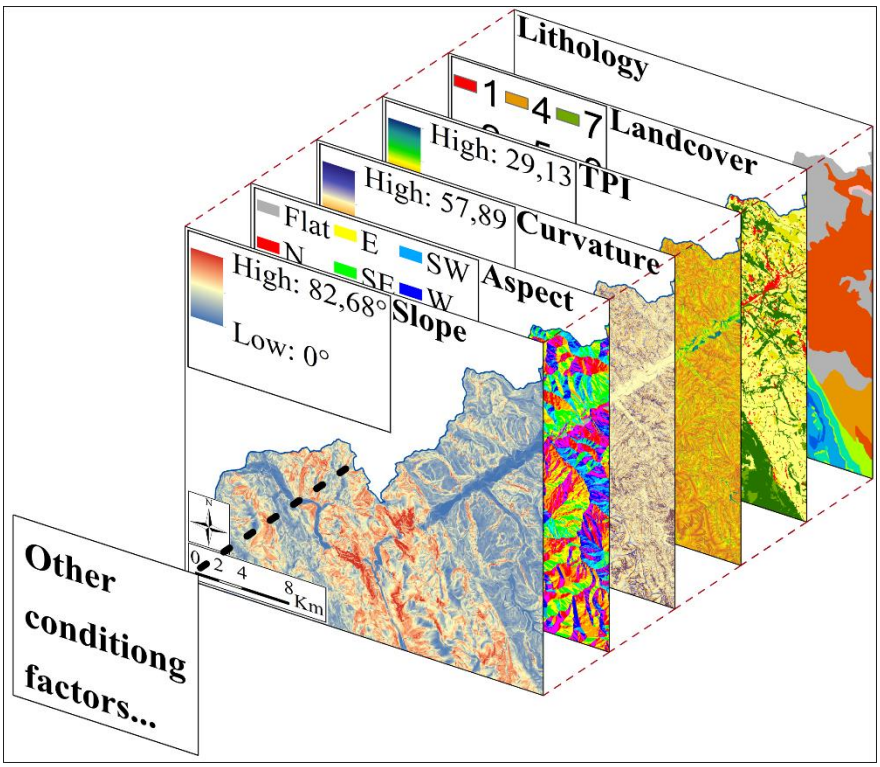


Figure 1.3. Example of conditioning factors (static) stacked and aligned.

GIS software is set as the best tool to extrapolate topographical factors, mainly from Digital Elevation Models (DEM), but also helpful for retrieving information about digitalized geology, land cover, and anthropogenic factors. Other information can be obtained from satellite images, like NDVI or soil moisture, or even triggering rainfall. However, it is not straightforward to determine which factor is mostly influencing the landslide presence (static) (Kavzoglu et al. 2015) and occurrences (temporal) (Steger et al. 2022b, Dal Seno et al. 2024). The availability and the quality of this terrain information is drastically connected to the study area/region/county (Van Westen et al. 2003, Woodard et al. 2023). Moreover, it is not just the changing of the available information and their resolution (Jebur et al. 2014) to differentiate each study area results, but also the real terrain/environmental condition and landslide types that change itself, thus, change their relationship (Van Westen et al. 2003).

Based on the model, the independence of the conditioning factors is a prerequisite, leading few authors to study directly the influence of the factors and the selection of the most influent and independent factors (Liao et al 2022). Nevertheless, with the rising of the ML methods, which learn by themselves the most important features and their dependence; as demonstrated by Meena et al 2022, removing the least influential factors does not improve the ML model's prediction. With the development of XAI it is becoming easier to understand this relationship between results and predictors, thus making it more manageable (Dahal & Lombardo 2023).

The triggering factors are even more difficult to access, due to their time-dependence, as it is complex to attest the exact date/year of occurrence of a certain landslide and thus correlate to

the exact triggering event (Lombardo & Tanyas 2020). Therefore, something that has been commonly done with rainfall data is to take the triggering phenomena not as single event but as a year average in a large period of time, in order to eliminate the single event and landslide correlation and extract a more general features of the territory in the long term (Broothaerts et al. 2012, Yano et al. 2019). In this way the rainfall, which is the trigger, is treated like environmentally preparatory and static information. This is not anymore suitable given the climate change conditions, which brought general increase of intense rainstorms or very wet conditions in a short period and prolongate drought conditions (Donat et al. 2016, Kikstra et al. 2022).

1.2.3 Landslide inventory

Landslide inventory is the spatial information of the present and actual landslides in a given territory, it could be more or less complete, incorporating polygonal information (Santangelo et al. 2015) or just the point of the landslide occurrences (Trigilia et al. 2010). It can be further detailed with the type of movement associated with the polygon (Ardizzone et al. 2023), or the date and triggering event (Santangelo et al. 2023). Nevertheless, the landslide inventory is usually produced by photointerpretation and field investigation given the geomorphological features, thus it refers to a huge span of time given by geomorphological evidence of different climate and environmental conditions (Reichenbach et al. 2018). Guzzetti et al. (2012) groups the landslide inventory in four classes: geomorphological inventories (the majority); event inventories associated with a specific triggering event; multi-temporal inventories associated with multiple triggering event or periods; and historical inventories which account historical records and alerts. In this context we would like to acknowledge the Italian case of "Inventario dei Fenomeni Franosi in Italia" (IFFI) by Trigilia et al. 2010 and "Piano di Assetto Idrogeologico" (PAI), that make available geomorphological/historical landslide inventories with attached the movement type information from Varnes 1994 for the whole national territory. As anticipated in the previous paragraph the multi-temporal inventories are complex to create/obtain (Van Westen et al. 2006), given the difficult attribution of the triggering event to the landslide inventory, even though, are the richest in information (Galli et al. 2008). However, emerging automated mapping methods are addressing these gaps, and some advances now include landslide type (Bhuyan et al. 2024).

Landslide susceptibility modelling is intrinsically connected to the quality and completeness of the landslide inventory (Guzzetti et al. 2000; Fell et al. 2008). The well-known geological idiom "the past is the key to the future" is particularly relevant here. Regardless of the methodology employed (statistical, machine learning or even expert-based) any susceptibility assessment must be learned from past occurrences. Understanding where landslides have taken place, and equally important, how they occurred (e.g. their morphologies, triggering mechanisms, material types), is essential for interpreting the spatial relationship between landslides and terrain characteristics, a feature often overlooked in the literature. In ML context the landslide inventory would become our labels, meaning our target.

However, it is important to stress that susceptibility maps are not simply replications of landslide inventory maps. Their purpose is not to replicate past landslides, but to extrapolate

from them, capturing the terrain's inherent predisposition to fail under certain conditions. Ideally, a robust susceptibility map should highlight not only areas where landslides have already occurred but also identify similar terrain units that may be equally or more prone to future failures. Therefore, while existing landslide areas must score high in susceptibility, the value of the model lies in its ability to generalise beyond the training inventory, extending predictive insight to non-landslide areas. Especially useful in regions where few or no landslides have yet been documented but the conditioning factors suggest instability.

The landslide inventory used as ground truth, independently from our chosen model, is extremely influencing on the results (Zêzere et al. 2009, Gaidzik & Ramírez-Herrera 2021). Landslide with polygonal shapes that are erroneous or too simplified or their triggering position not properly mapped are misleading for the construction of our susceptibility maps, because the model (statistical or ML) learns on an incorrect assumption. Therefore, a reliable inventory is fundamental for a good susceptibility map, which is often not available, depending on the study area. Moreover, the failure mechanism information should be taken into account (Reichenbach et al. 2018), as every landslide type has unique conditioning factors features, as understandable form the Figure 1.4. Few authors incorporated different landslide types knowledge in their study, providing a more comprehensive result (Loche et al. 2022, Yu et al. 2024, Regmi et al. 2014).

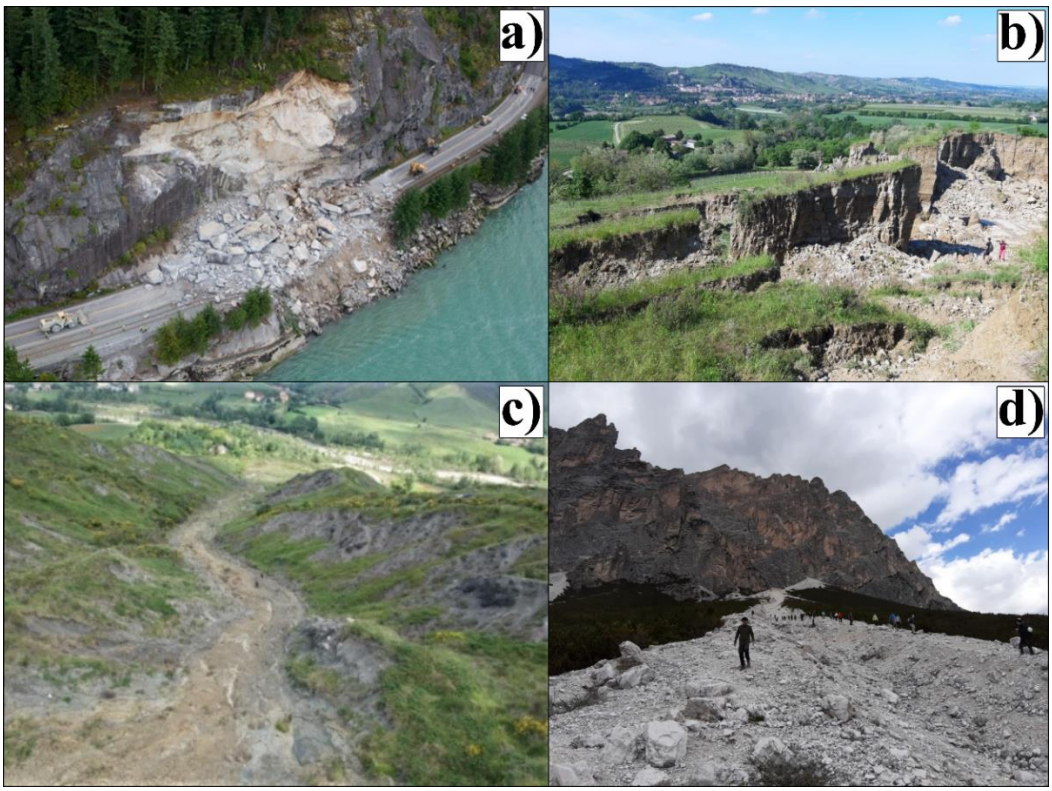


Figure 1.4 Photos of different landslide types: ***a) Rockfall****, Porteau Cave, BC, Canada (Canadian Broadcasting Corporation);* ***b) Rockslide****, Brisighella, RA, Italy (Rodolfo Rani);* ***c) Earthflow****, San Clemente, BO, Italy (Rodolfo Rani);* ***d) Debris flow****, Cortina, BL, Italy (Rodolfo Rani).*

Another important factor influencing landslide susceptibility modeling is the sampling strategy used to define what constitutes a landslide in the inventory. In the literature, two main sampling strategies are commonly employed: (a) sampling observations from the entire landslide body, and (b) sampling observations specifically from the landslide release areas (i.e., scarp locations), as illustrated in Figure 1.5. However, the effects of these different sampling locations on the resulting model predictions are seldom analyzed (Clerici et al. 2006, Regmi et al. 2014, Hussin et al. 2016, Dou et al. 2020, Guo et al. 2024). Even fewer studies have focused on spatial prediction patterns and their geomorphological implications, comparing different landslide zones (e.g., release vs. deposit areas) (Steger et al. 2016, Lima et al., 2023). In general, a landslide susceptibility model focused on initiation zones of flow-type movements can effectively capture the topographic controls on landslide initiation (Steger et al. 2022a). However, this approach may be less effective for assessing exposure to buildings and infrastructure located in flatter areas at the foot of the slope. Models trained on landslide release zones are tailored to predict the location of landslide initiations, whereas models trained on the entire landslide body may offer a broader prediction of the potential impact area, which could be more appropriate for areas of gentler terrain (Lima et al. 2023).

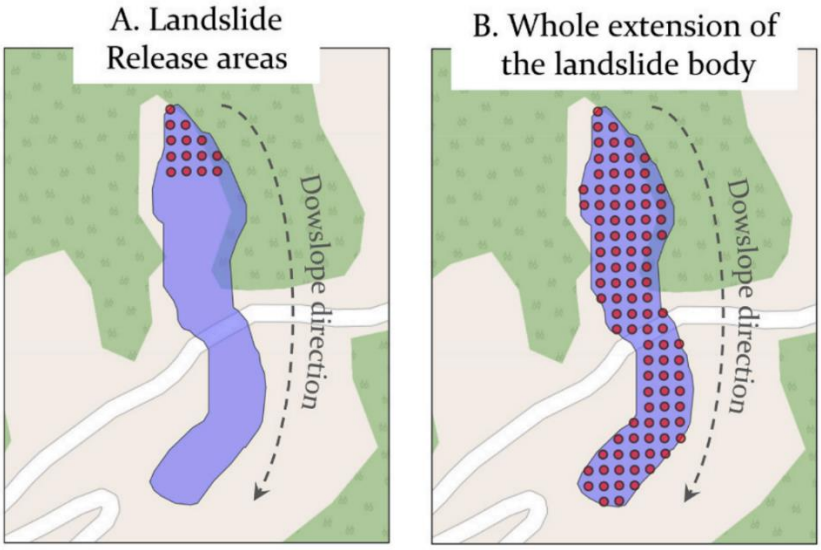


Figure 1.5. Schematic landslide sampling technique usually adopted (landslide release and body sampling) (Lima et al. 2023, modified).

Notice, that the best would be having a reliable and detailed landslide inventory and contextually have the terrain data exactly the moment before landslide (DEM, NDVI, landcover, etc). Because given the goal of understanding where landslide is not present, we should compare this with the terrain condition before the landslide presence where now we observe it (Reichenbach 2018). The landslide itself changed the terrain conditions (Curvature, NDVI, landcover in long term), thus it could not be a proxy of totality of the conditioning's factors. This concept had already arisen by Atkinson & Massari (1998), but not yet addressed sufficiently, for two reasons: fist the lack of data, before and after a landslide or landslide inventorying is done; second, the fact that is generally seen with a good approximation using the conditioning factors the same through time.

1.2.4 Area partitioning units

Landslide susceptibility modelling requires the definition of a mapping unit, which acts as the spatial reference for retrieving and analysing all geo-environmental information. The choice of mapping unit significantly influences the resulting zonation, model interpretation, and applicability (Guzzetti et al. 1999).

Mapping units are generally grouped into three categories: grid cells (pixels), slope units (SUs), and unique condition units (UCUs) (Reichenbach et al. 2018).

- Grid cells are square raster units, typically derived from DEMs, and are the most widely used due to their computational simplicity and compatibility with GIS platforms.
- Slope units represent hydrological and geomorphological entities bounded by drainage divides and channels, aligning more closely with the natural geometry of hillslopes and landslide processes (Carrara et al. 1991).
- Unique condition units are obtained by intersecting thematic layers (e.g., slope, lithology, land cover), typically refining slope units based on environmental attributes.

Grid cells are flexible and easy to integrate with raster datasets, but their lack of geomorphological meaning, especially at coarse resolutions, can limit interpretability. Moreover, high-resolution grids may introduce class imbalance (few landslides vs. many non-landslide cells), and their outputs can appear fragmented, making operational use difficult (Jacobs et al. 2020).

Slope units are gaining increasing interest, particularly in regional-scale assessments, due to their ability to accurately reflect natural slope morphology and catchment dynamics (Jacobs et al. 2020; Lei et al. 2024). They are now easier to generate, thanks to open-source tools (e.g. Alvioli et al., 2016) and are less computationally intensive (Camilo et al. 2017). Additionally, slope units have been shown to provide more accurate results in performance evaluations (Liu et al. 2025). SUs are particularly well-suited for modelling large, deep-seated, or complex landslides, as they often encompass the entire unstable slope area. However, slope units may oversimplify terrain at local scales and thus be less appropriate for urban planning, where more detailed resolution is required, as they do not capture fine-grained variations in landform or infrastructure (e.g. Figure 1.6 a and b SUs are too large to fully reflect the landslide distribution). In such contexts, susceptibility models benefit from incorporating full landslide polygon information in a pixel-based grid, since the entire landslide body (from the crown concavity to the toe camber) can provide valuable morphological insight (Lima et al. 2023). Furthermore, not all landslide types are compatible with the assumption of a single representative area within a slope unit (e.g., the crown) to label landslide presence. Debris flows, for instance, have well-defined triggering areas, while their depositional zones are primarily controlled by downslope topography. When analysed at the scale of a full slope unit, the averaged conditioning factors may not accurately reflect the conditions at the initiation point, and vice versa for the deposit zone (Amato et al. 2019, Chang et al. 2023). Conversely, slope units are particularly appropriate for landslide types such as rotational slides in clays or slow-moving complex landslides, where triggering and depositional areas evolve

simultaneously above a shared slip surface. In such cases, instability is often distributed across the entire slope, making slope units a meaningful and geomorphologically coherent unit for susceptibility mapping.

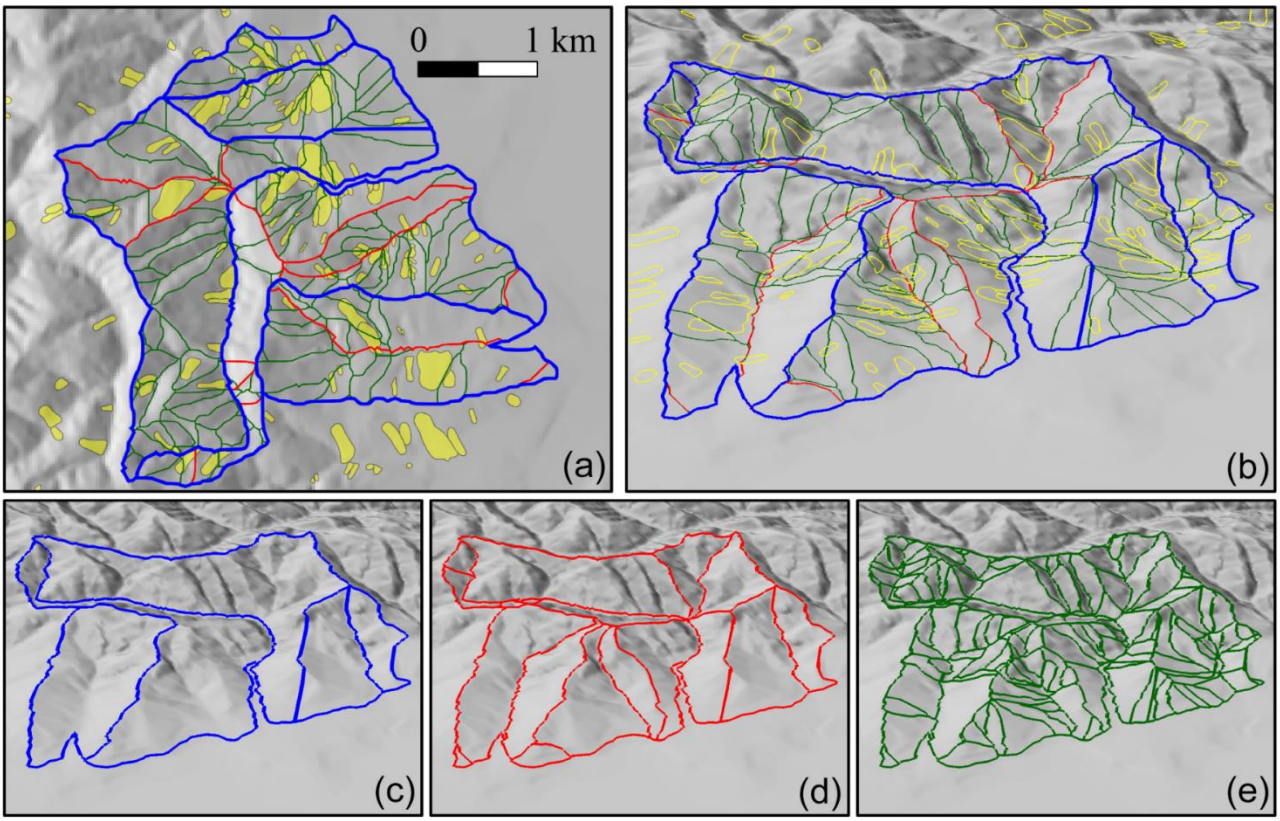


Figure 1.6. Example of subdivisions into SUs for a portion of the study area. Legend: blue, red, and green lines show boundaries of SUs of increasing density and corresponding decreasing average size. Yellow areas are landslides. The five maps show the same area in plan view ***(a)*** *and in perspective view* ***(b, c, d, e)*** *(Alvioli et al. 2016).*

UCUs offer a middle ground; by capturing unique combinations of conditioning factors, they can improve thematic precision, especially when slope units are further subdivided by lithology or land cover. However, they can result in a very large number of small and difficult to interpret units and require classification of continuous variables, which introduces uncertainty.

Ultimately, the selection of a mapping unit should be tailored to the landslide type, scale of the analysis, data resolution, and the intended use of the susceptibility map (Guzzetti 2006). For operational applications such as territorial planning, pixel-based or refined UCU-based approaches often offer better spatial detail. In contrast, SUs remains ideal for broader scale assessments and flow type landslides, where hydrological influence dominates (Wang et al. 2022), but without systematically improving their performance (Martinello et al. 2021, Ma et al. 2023).

Finally, in specific cases, administrative units have been adopted as mapping units particularly when the outputs are intended for civil protection use or resource allocation. Such units allow alignment with municipal boundaries (Fang et al. 2025) or facilitate communication in early warning systems (Segoni et al. 2025).

1.2.5 From susceptibility to probability

Recently, the research focus within the landslide community has shifted from producing landslide susceptibility maps to developing probability maps. The key difference is that susceptibility maps are based solely on static predisposing factors (e.g., geology, land cover, topography) and are therefore considered time-invariant. On the other hand, the term probability is broader and typically refers to the likelihood of a landslide occurring given certain environmental conditions, including triggering or preparatory factors that vary over time.

While traditional susceptibility mapping aims to produce a general-purpose map valid across time, it deliberately excludes time-dependent factors. Nevertheless, some studies (Broothaerts et al. 2012, Yano et al. 2019) have incorporated averaged triggering variables over long time spans, in an attempt to represent a "static" environment of influence. However, under evolving climate conditions, this assumption is no longer valid (Donat et al. 2016, Myhre et al. 2019). For example, Ma et al. (2025) demonstrate a direct correlation between climate change and increased landslide occurrences, primarily driven by the growing frequency of extreme rainfall events. Furthermore, Gariano & Rianna (2025) reviewed 50 studies on the impact of climate change on landsliding in Europe. They concluded that climate change is expected to increase the frequency of shallow landslides, while some projections indicate a possible reduction in the activity of deep-seated movements.

As a result, there is a growing demand for landslide probability mapping. However, this requires highly detailed and well-structured multi-temporal landslide inventories, which are often difficult to obtain. These inventories must include not only spatial information (e.g., points or polygons), but also temporal data (e.g., yearly or monthly occurrence, or exact dates), and preferably, information on the landslide mechanisms (Bhuyan et al. 2024). Such multi-event inventories are still rare in many countries, yet they are essential for advancing toward more dynamic and predictive models of landslide hazard (Guzzetti et al. 2012).

The importance of a multi-event landslide inventory lies in its ability to capture the temporal variability of triggering and preparatory factors, allowing the model to learn not only from static conditioning variables, but also from the dynamic environmental conditions associated with landslide occurrences (Bordoni et al. 2021) associated in a climate change context (Kim et al. 2021). By incorporating information from multiple landslide events across different time periods, the model can identify patterns linked to specific meteorological or hydrological scenarios, such as short, intense rainfall bursts or prolonged wet conditions, and relate them to the spatial configuration of the terrain (schematic workflow in Figure 1.7). This temporal dimension enables the development of predictive models capable of simulating future landslide scenarios under varying climatic conditions, thus transitioning from a purely spatial assessment (susceptibility) to a truly spatiotemporal framework (probability). In essence, the model does not simply learn where landslides have occurred, but when and under what conditions, making it a more powerful tool for both risk forecasting and early warning applications (Fang et al. 2023). Nevertheless, evaluating landslide susceptibility remains a valid and essential approach, particularly when the objective is long-term territorial planning or land-use regulation.

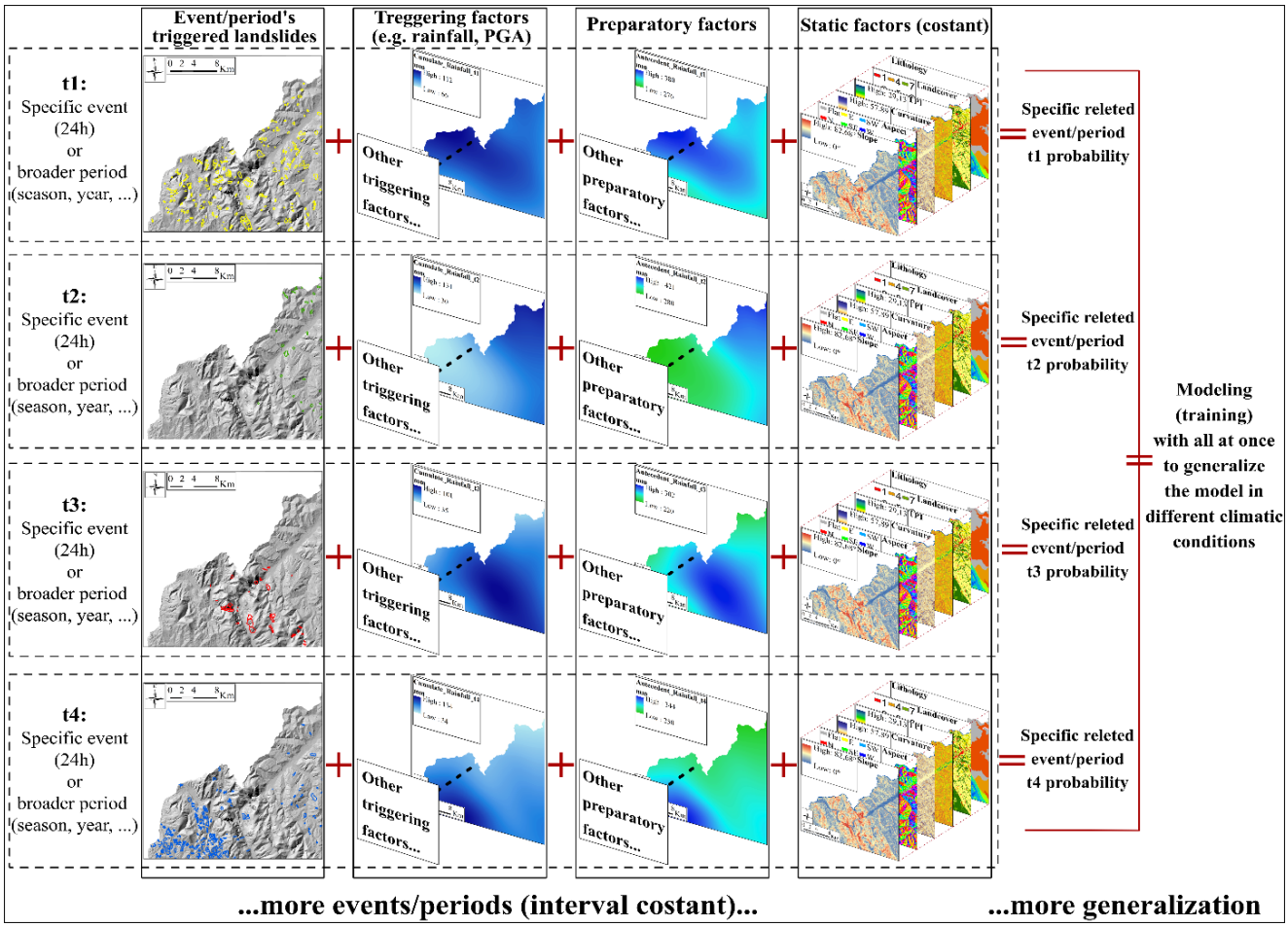


Figure 1.7. Conceptual scheme of a landslide probability modeling with multi-event/temporal landslide inventories and related conditioning factors through time and space.

1.3 Modelling overview

Landslide susceptibility modelling is widely recognized as one of the most critical steps in producing accurate and reliable maps (Reichenbach et al. 2018). However, as previously discussed in this thesis, the quality of the source data, the conditioning factors and particularly the landslide inventory, as well as the choice of mapping units, also plays a fundamental role in the overall process (Liu et al. 2025). Therefore, the reliability of susceptibility or probability maps depends not only on the modelling technique itself, but on the integrity and coherence of the entire processing chain (Reichenbach et al. 2018).

Nonetheless, modelling remains a central component and should not be underestimated. The literature presents an extensive array of modelling approaches, which can be broadly classified into four categories: expert-based, physically based, statistically based, and machine learning methods. Setting aside the primordial heuristic expert-based method, already mentioned in the first paragraph, here we briefly discuss the physically based methods, and then in detail some statistics and ML methods.

Physically based models simulate the actual processes that lead to slope failure by applying mathematical formulations of slope stability. These models require detailed geotechnical, hydrological, and topographic data, making them particularly suitable for small to medium-

scale studies where such data are available and relatively homogeneous. Unlike statistical or machine learning approaches, physically based models do not depend on past landslide inventories. Instead, they assess slope stability directly through the calculation of the factor of safety (Fs) or similar threshold-based indicators. Two widely used examples are SHALSTAB (Shallow Landsliding Stability Model; Montgomery & Dietrich 1994) and TRIGRS (Transient Rainfall Infiltration and Grid-Based Regional Slope-Stability Model; Baum et al. 2008). These tools evaluate slope stability under both steady-state and transient hydrological conditions.

A complementary category of physically based models includes runout models, which simulate the post-failure dynamics, thus, how landslide material travels downslope from the initiation zones. When these models are integrated with susceptibility maps that define potential triggering areas, they provide a more comprehensive view of landslide hazard. This combined approach has proven particularly valuable for debris flows and other fast-moving landslides (e.g. Kritikos & Davies 2015, Mergili et al. 2019, Lima et al. 2023, Bornaetxea et al. 2025), bridging the gap between susceptibility and full hazard mapping.

Most statistical models applied in landslide susceptibility are classification models, which means their goal is to assign each mapping unit (e.g., pixel, slope unit) to a binary class: landslide (presence = 1) or non-landslide (absence = 0). In this context, the model does not predict how much deformation or displacement will occur (which would be a regression problem), but rather where landslides are likely or unlikely to occur. This classification concept is usually applied also on the ML models.

However, recent studies have examined the relationship between ground deformation (displacement) and both predisposing and triggering factors (site-specific: Nava et a. 2023, Zhou C. et al. 2024), although in a form of regional probability (comparable to susceptibility maps) only by He et al. (2025), whose fascinating results are shown in Figure 1.8. While these studies provide valuable insights, they diverge from traditional susceptibility models, which aim to describe the terrain over a long-time span using either time-independent predisposing factors or landslide probability derived from multi-temporal inventories along with related preparatory and triggering factors. Ground deformations, often obtained from satellite-based measurements (e.g., InSAR, SBAS), are typically time-constrained to a limited period, usually ranging from 5 to 15 years, depending on data availability. This time constraint can lead to misrepresentation of the full distribution of ground movements, as it does not account for the potential variability in deformation caused by changing climatic conditions over a longer period. Consequently, this approach may fail to capture the full range of landslide hazards, as we pointed out in the next chapter of this thesis.

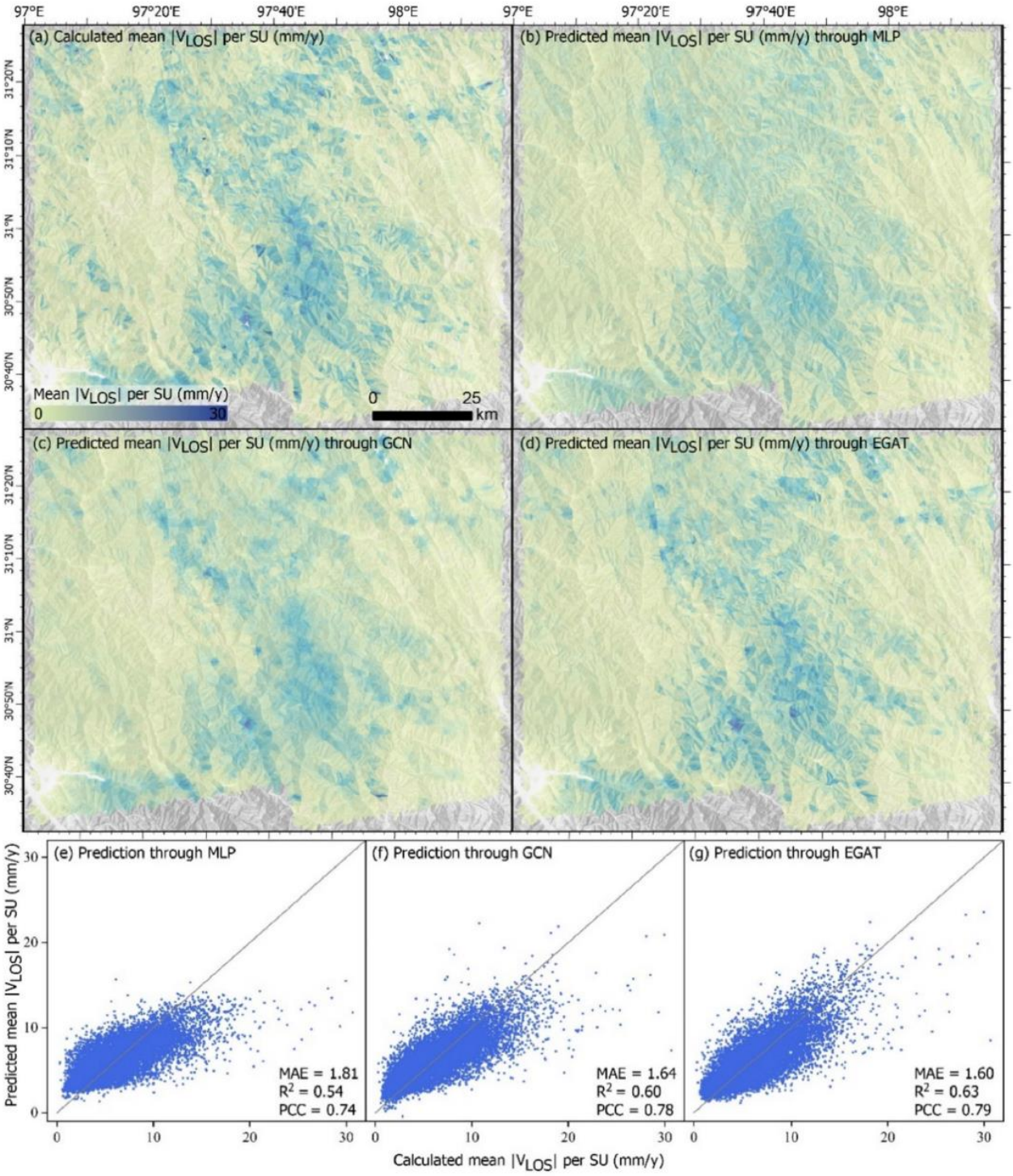


Figure 1.8. Mean V_{LOS} maps associated to ***(a)*** *InSAR calculations, Multi Layer Perceptron (MLP)* ***(b),*** *graph convolution network (GCN)* ***(c)*** *and Edge-featured graph attention network (EGAT)* ***(d).*** *The association between observed and predicted values is respectively presented in panels* ***(e)****,* ***(f),*** *and* ***(g).***

1.3.1 Statistically based models

I present here some of the most commonly used statistical models and their applications. Statistically based models produce predictions based on the distribution of the data, specifically, the relationship between landslide occurrences (the ground truth) and the conditioning factors.

The primary goal of these models is to quantify the correlation between landslide presence and each predictor variable, relying on different mathematical concepts. These models rely entirely on observed data and do not simulate physical processes; instead, they infer spatial likelihoods of landsliding through correlations and patterns within the data, from which the synonym "data-driven" for these models.

Frequency Ratio (FR)

Frequency Ratio is a bivariate approach (assess each conditioning factor independently) commonly used in landslide studies due to its straightforward implementation (Bonham-Carter 1994). It measures how often landslides occur in areas characterised by a specific class of a conditioning factor, compared to how much that class occupies the entire study area (Lee & Sambath 2006). For example, if 30% of landslides occur on slopes between 30–35 degrees, and such slopes make up only 10% of the total area, the frequency ratio for that slope class would be high, indicating a strong association with landslide occurrence. The method is easy to compute and interpret, however, a major limitation is its assumption that all factors are statistically independent, which may not hold true in natural settings where many terrain attributes are correlated (e.g., slope and elevation) (Ozdemir & Altural 2013). Another limitation is that it requires all continuous predictors to be discretized into meaningful classes (overtaken by Li et al. 2017), which may introduce subjectivity and affect model accuracy.

Weight of Evidence (WoE)

Weight of Evidence is also a bivariate method rooted in Bayesian probability theory (Bonham-Carter 1994). It was originally developed for mineral exploration (Agterberg 1992) and later adapted to natural hazard studies (Lee & Choi 2004). WoE works by evaluating how strongly each category of a conditioning factor (for instance, a specific lithology or land cover type) is associated with the presence or absence of landslides. The method calculates two key values for each class of each predictor: a positive weight, which indicates how strongly a certain terrain condition is associated with the presence of landslides, and a negative weight, which expresses its relationship with non-landslide areas. The difference between these weights, known as contrast (see example in Figure 1.9), reflects the relative influence of that class on slope instability. All the weights are then combined to estimate the probability of landslide occurrence for any location, based on the characteristics of that location. WoE has been appreciated for its conceptual simplicity, intuitive logic, and ease of implementation in GIS environments. However, like FR, it evaluates each factor separately and therefore cannot account for interactions or cumulative effects between variables. WoE, like FR, need the discretization in classes of continues values, incorporating subjectivity.

Logistic regression (LR)

Logistic Regression is a multivariate statistical model (assess multiple conditioning factors simultaneously) that is widely recognized for its robustness and interpretability (Ayalew & Yamagishi 2005). It aims to model the probability that a given location is affected by a landslide based on a combination of multiple predictors. Unlike bivariate models, LR evaluates how all the variables act together, controlling for their mutual effects. It also allows for statistical testing

to determine the significance of each variable, providing insight into which factors are most influential. LR is based on a log-odds relationship between predictors and outcomes, and the coefficients of formula are generally obtained through statistical maximum likelihood estimation. This model works well with both continuous variables (e.g., slope gradient) and categorical variables (e.g., geology), and can accommodate non-linear relationships through data transformation. It has become a cornerstone in landslide susceptibility mapping thanks to its balance between performance and ease of interpretation (Reichenbach et al. 2018). The main limitations of LR include the assumption of linearity between predictors and the log-odds of the outcome, which might not capture complex terrain processes (Hong et al 2017, Woodard et al. 2023).

Generalized Additive Model (GAM)

Generalised Additive Models (GAM) represent an evolution of LR by offering greater flexibility in modelling non-linear relationships between predictors and landslide occurrence (Hastie & Tibshirani 1986). Rather than forcing each predictor to influence landslide susceptibility in a log-linear way like LR, GAMs allow the data to "speak for themselves" through the use of smooth curves or splines achieving a real data-driven model (Wood 2017). For example, a GAM can identify that landslide probability increases with slope angle up to a certain point and then levels off or even decreases, behaviors that a standard LR would struggle to capture (Moreno et al. 2023). This flexibility allows GAMs to model threshold effects, optimal ranges, and other complex patterns that are common in geomorphological processes. As such, they are particularly suitable for regional-scale applications where environmental gradients are complex and non-linear. GAMs have been shown to outperform traditional methods in many landslide-prone areas, plus providing good interpretability (see Figure 1.9) (Brenning 2008, Ahmed et al. 2023) are gaining popularity in the scientific community (Alvioli et al. 2024). As shown in Figure 1.9 the interpretability of the influence of GAM on the result is clear, like the one express by the contrasts of WoE (notice that are different y-scale and meaning). However, the GAM does not have to be discretised a priori, making it clearly more flexible, capturing more precise nuances than the WoE. Although they are flexible due to their use of smooth functions to model non-linear relationships, their flexibility is inherently constrained by the choice and tuning of spline parameters, which can limit the model's ability to capture highly complex interactions compared to more adaptive ML models (Goetz et al. 2015).

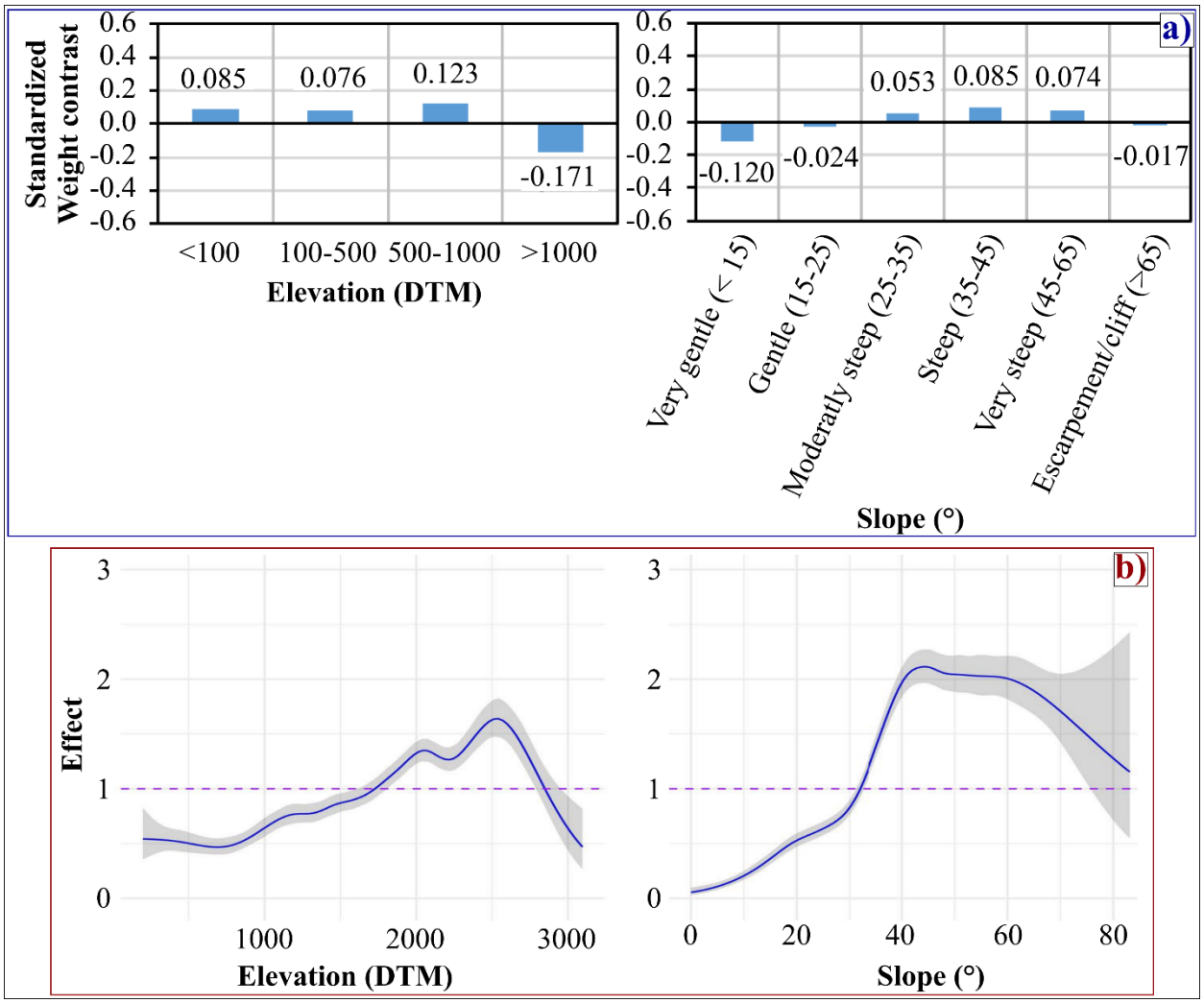


Figure 1.9. Comparison of WoE and GAM visualizations of the influence by the factors (DTM and Slope) on the landslide susceptibility (different areas). ***a)*** *WoE example from Batar & Watanabe (2021) (modified),* ***b)*** *GAM example from Patanè et al. (2025) (modified).*

1.3.2 Machine learning models

Machine learning (ML) has become an increasingly relevant approach for landslide susceptibility and probability modelling due to its ability to learn from large datasets and continuously improve its predictions. Unlike traditional statistical models, which attempt to capture the relationship between landslides and their predictors through predefined formulas, ML algorithms adapt to new data and can uncover complex, non-linear relationships without explicit programming.

The strength of ML lies in its dynamic learning capacity, particularly in contexts where traditional rule-based systems struggle. For example, in fields like spam detection, manually crafted rules can quickly become outdated, whereas ML models can adapt and refine their strategies by learning from continuously updated data, leading to more accurate predictions (Bhowmick & Hazarika, 2018). Moreover, in domains that rely on large datasets, such as medicine and astrophysics, where ML is used to predict patient outcomes or analyse vast

amounts of astronomical data, identifying patterns that would otherwise go unnoticed (Ball & Brunner, 2010).

While ML has seen widespread success in industries such as finance, e-commerce, and healthcare, its adoption in fields like civil protection, territorial management, and disaster response has been slower. However, this trend is beginning to change, with increasing use of ML in these domains in recent years (Liu et al. 2023). It's important to note that the success of ML applications in any field, especially landslide susceptibility, depends heavily on the quality and availability of data. In this regard, ML models are often more reliant on high-quality data than traditional data-driven models, as the learning process is data-intensive and model performance is directly tied to the dataset's completeness and accuracy.

ML methods have already proven to perform better than statistical-based models (Kavzoglu et al. 2019), and within the ML models, more advanced deep learning models continue to improve the performance compared to "classical" ML techniques (Azarafza et al. 2021).

Despite their remarkable performance and ability to capture highly complex, non-linear relationships, there are several drawbacks of machine learning (ML) models that require attention. One primary concern is overfitting, which occurs when a model essentially "memorizes" the relationships between certain factors and the outcome, rather than generalizing from the data (Ghojogh & Crowley 2019). This overfitting results in a model that performs well on the training dataset but is less reliable when applied to new or unseen data. Given that landslide susceptibility mapping aims to predict the potential occurrence of landslides outside the inventory, an overfitted model will fail to provide useful information for those areas. In practice, this means the model will only predict high susceptibility (close to 1) within areas that already have landslides, which defeats the purpose of forecasting potential hazards in the whole territory. To prevent overfitting, the dataset is typically divided into training and testing sets. The testing set is used to evaluate the model's ability to generalize. Additional techniques like dropout (randomly ignoring some neurons during training) or adjusting the learning rate (controlling how much the model changes with each iteration) are commonly used to reduce the model's learning capacity and prevent overfitting.

Another key issue is the imbalance of data, which is particularly relevant for landslide datasets where there are often far more non-landslide (0) samples than landslide (1) samples. In these situations, the model might focus too much on predicting the non-landslide class, as it dominates the dataset, and neglect the factors that control landslide occurrence. This imbalance can result in a model that poorly represents the true landslide conditions. To address this, various sampling techniques are used to balance the dataset, such as undersampling the non-landslide areas or oversampling the landslide areas to ensure equal representation (e.g. SMOTE or random sampling by Chawla et al 2002, positive-unlabeled bagging and buffer control sampling by Gu et al. 2024, hybrid sampling based on Mahalanobis distance by Nam et al. 2024). Alternatively, class weights can be applied to give greater importance to landslide occurrences during training, forcing the model to pay more attention to the minority class.

Lastly, a significant challenge in ML models is interpretability. Unlike traditional statistical models, where the relationships between landslide occurrence and predictors are explicit (e.g.,

through formulas or graphs), ML models often operate as "black boxes." This lack of transparency is particularly problematic when trying to interpret the geomorphological processes driving landslides. To address this issue, Explainable Artificial Intelligence (XAI) has emerged as a solution (Figure 1.10). Initially proposed by Ribeiro et al. (2016) and then generalized by Gunning & Aha (2019). XAI techniques compute the contribution of individual factors to the model's predictions, providing insight into how and why a decision was made. Since its introduction, the implementation of XAI has grown exponentially, and to a lesser amount also in landslide modelling domain (Dahal et al. 2023, Chen & Fan 2024) XAI methods, such as SHapley Additive exPlanations (SHAP) (Lundberg & Lee 2017), Local Interpretable Model-agnostic Explanations (LIME) (Ribeiro et al. 2016), and Integrated Gradients (IG) (Sundararajan et al. 2017), rely on the core idea of computing gradients, which represent the rate of change of the model's output with respect to each input feature. These gradients are used to assess the influence of each feature on the model's prediction, providing interpretable explanation scores. Essentially, by calculating how much the output changes as the input features vary, XAI methods reveal which factors contribute most to the model's decisions. While this area is not the primary focus of this thesis, we discussed it further in Chapter 4.

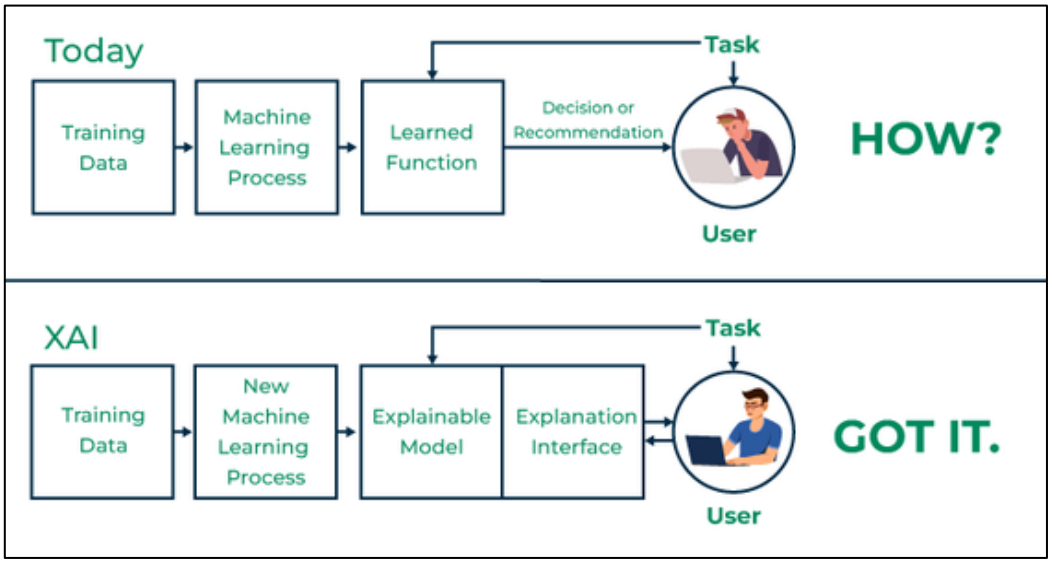


Figure 1.10. Explainable Artificial Intelligence (XAI) conceptual example (modified from Geeks for geeks).

Listed below are the most commonly used ML models in landslide susceptibility context, introducing them with a brief description highlighting their advantages and limitations.

Machine learning Logistic Regression (ML-LR)

Logistic Regression with a Machine Learning approach (ML-LR) is focused on optimizing prediction accuracy rather than directly capture the relationships between features and outcomes like traditional LR. While it shares the same mathematical framework as traditional logistic regression, ML-LR is applied in a more data-driven, performance-oriented context. Key characteristics include its ability to handle complex, non-linear relationships through regularization (L1/L2) to prevent overfitting, allowing for better generalization. The LR coefficients will still be estimated using maximum likelihood, but in the ML context,

regularization is often applied to control the complexity of the model. Adding, cross-validation and hyperparameter tuning in the training phase it will maximize predictive performance, better than statistical LR (Azarafza et al. 2021). In conclusion, unlike traditional logistic regression, ML-LR emphasizes prediction and generalization over interpretability, making it ideal for handling large, high-dimensional datasets.

Random Forest (RF)

Random Forest is an ensemble learning method that builds a "forest" of decision trees. Each tree is trained on a random sample of the data and uses a subset of the predictors. The final prediction is made by aggregating the outputs from all trees (e.g., majority vote for classification). This approach is robust to overfitting, handles non-linear relationships well, and works effectively with both continuous and categorical variables. However, RF has some limitations. Like most machine learning models, its interpretability is limited, making it difficult to understand which specific features are most important in the decision-making process. Additionally, RF is memory intensive; since it creates multiple decision trees, it may require significant memory resources, especially when working with large datasets or high-dimensional features. Thus, computational cost and training time can be substantial, particularly with a high number of trees. Finally, RF may struggle with imbalanced datasets (e.g., when there are more non-landslide than landslide pixels), and the model might be biased toward predicting the majority class (non-landslide), unless specific balancing techniques are applied (Lu et al., 2025).

Generalised Additive Model with gradient boost (GAMBoost)

GAMBoost combines the flexibility of Generalised Additive Models (GAMs) with the power of boosting. GAMs allow for non-linear relationships between each predictor and the response by fitting smooth functions, rather than assuming a fixed shape (Lou et al. 2013). Boosting enhances model accuracy by iteratively fitting the residuals of previous models. In landslide susceptibility modelling, GAMBoost enables the modelling of complex, smooth terrain-landslide relationships, while improving predictive performance through adaptive learning. This approach is particularly useful when the goal is both prediction and interpretability, as it allows easy interpretation using partial dependence plots that visualize the effect of each predictor on the outcome, including interaction effects between variables, unlike many other machine learning models. However, GAMBoost does have some limitations. It can become memory-intensive when working with large datasets or a high number of trees, similar to RF. The computational cost and training time can be substantial as boosting iterates through multiple models. While the model offers strong interpretability, the model complexity can increase as more trees are added, making it harder to fully understand the interaction of features in some cases.

Support Vector Machine (SVM)

Support Vector Machine (SVM) is a supervised learning algorithm that seeks to find the optimal boundary (or hyperplane) that separates landslide and non-landslide areas. It works well even in high-dimensional spaces and can model non-linear relationships using kernel functions. In

landslide susceptibility mapping, SVM can distinguish subtle patterns in the data, especially when the boundary between stable and unstable areas is not easily defined. SVM has proven to be effective with small dataset compare with other ML methods (Huang & Zhao 2018). However, it can be computationally demanding with high numbers of features, as it requires computing distances between all data points. Another point is that the performances of SVM heavily depends on kernel choice, regularization parameters, and the kernel parameters Optimizing these parameters and their optimization requires careful tuning and cross-validation, which can be time-consuming. SVM, like other ML acts like a "black box", thus not interpretable, plus is sensitive to class imbalance. Although SVM can achieve high accuracy, it may underperform when dealing with noisy data or when the decision boundary is not clearly defined. In such cases, more flexible models like Random Forests or Gradient Boosting can outperform SVM (Huang & Zhao 2018).

Extreme Gradient Boost (XGBoost)

XGBoost is a highly efficient and scalable gradient boosting algorithm. It constructs decision trees sequentially, each one correcting the errors of the previous, while incorporating regularisation to reduce overfitting. Compared to other boosting methods, it includes optimised routines for speed and accuracy with the improvement of the performance against other boosting methods (Sahin 2022). In landslide applications, XGBoost is prized for its high performance, especially when large datasets with many predictors are involved. Interpretability is still a weak point without the integration of XAI post training methods. XGBoost is suitable and generally better than RF and SVM for high-dimensional dataset, remaining relatively faster and less sensitive to imbalanced dataset (Sahin 2023).

Neural Network (NN)

NNs are inspired by the structure of the human brain. They consist of layers of interconnected "neurons" that process input data and progressively learn complex relationships. The simplest form of a neural network is the feedforward network, where data flows unidirectionally from input to output through one or more hidden layers. Each neuron applies a non-linear activation function to its inputs, with popular choices including the Rectified Linear Unit (ReLU) for its ability to mitigate the vanishing gradient problem (Nair and Hinton, 2010).

In landslide susceptibility, NNs can model highly non-linear interactions among multiple conditioning factors. NNs excel at capturing complex non-linear relationships, making them ideal for landslide susceptibility. Another aspect of NNs is that are extremely flexible working with tabula data or images without affecting their powerful predictive performances. Interpretability is again like a "black box", aligning with other ML models. They were designed to be trained on a huge dataset; thus, they generally require large datasets to perform well and they struggle with small datasets. NNs are highly prone to overfitting, especially when the model is too complex relative to the data size and variety. Regularization techniques like dropout, early stopping, and L2 regularization are often necessary to prevent this issue. Like others ML models, NNs suffer from an imbalanced dataset, without careful sampling techniques or class weighting, they might perform poorly in predicting the minority class (landslides). Lastly, NNs could be computationally expensive, that can require significant

hardware resources. In comparison to other models like RF and SVM, NNs can provide better predictive performance (Tien Bui et al. 2016), but they require careful tuning and sufficient data to avoid pitfalls like overfitting and bias from imbalanced datasets (Dou et al. 2015).

Convolutional Neural Network (CNN)

CNNs are a type of neural network designed to process grid-like data such as images or raster layers. They use convolutional layers that scan the input for patterns (e.g., textures, edges) and learn spatial hierarchies of features. The training process of a CNN involves a feedforward phase, where input data passes through layers, including convolutional layers that apply learnable filters (kernels) to detect local patterns like edges and textures.

In the context of landslide susceptibility mapping, CNNs can directly capture spatial dependencies and terrain features from rasterized input predictors, making them especially suitable for this type of analysis (Hölbling et al. 2017). Recently, the application of CNNs in landslide susceptibility modelling has expanded from binary classification (landslide or not) to probability prediction (Zhang et al. 2024), with enhanced regularization, pooling strategies, and kernel size/sampling mechanisms (Wang et al. 2025). CNNs have been shown to outperform traditional machine learning models (Azarafza et al. 2021), and when combined with real ground deformation data, they can further improve landslide susceptibility mapping (Shi et al. 2025). However, CNNs are memory-intensive, have limited transferability across regions, and require careful generalization to avoid overfitting. They also struggle with imbalanced datasets, but this can be mitigated using data augmentation, class weights, and focal loss functions.

Transformer Neural Network (TNN)

Originally developed for natural language processing, Transformers are now being explored in geospatial modelling due to their ability to handle sequential or spatial dependencies using attention mechanisms. Unlike Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs), transformers can process entire data sequences at once due to their self-attention mechanism. The transformer architecture consists of an encoder and decoder, each made up of identical layers that incorporate self-attention and fully connected feed-forward networks. The self-attention mechanism is the heart of transformers. It analyzes relationships between different parts of an input sequence using three components: Query (Q), Key (K), and Value (V). This approach allows the model to determine the importance of each element in the sequence, regardless of its position.

In landslide susceptibility, TNNs can potentially integrate temporal rainfall sequences and spatial predictors simultaneously, making them suitable for both susceptibility and time-dependent probability forecasting, as demonstrate in chapter 4 of this thesis. However, their use in this domain is still emerging and typically requires substantial computational resources together with strong anti-overfitting techniques like max-pooling and properly imbalanced data treatment.

From the computer vision domain, a similar architecture called the Vision Transformer (ViT) is applied, where image patches are treated as tokens. This allows effective visual data

processing and can be applied "like a CNN" directly from raster-based predictor grids (Zhao et al. 2024, Chen et al. 2024).

1.3.3 Models' evaluations

The quality of a model is determined by its ability to accurately represent the original ground truth during the prediction phase. In practice, the evaluation is typically performed by comparing the model's predicted probability values (ranging from 0 to 1) with the actual landslide presence/absence in the original landslide inventory, which serves as the target during training. The literature offers various methods to assess model performance. For a classification task such as susceptibility modelling, where the model classifies each mapping unit as either a landslide or non-landslide (with a probability between 0 and 1), all evaluation metrics rely on thresholding the output to obtain a binary classification for comparison with the ground truth. Here, distinctions emerge between multi-threshold metrics and those based on a single cutoff threshold. The choice of threshold is crucial for constructing the confusion matrix (built as illustrated in Figure 1.11), and its components: True Positive (TP), False Positive (FP), True Negative (TN), and False Negative (FN), form the basis for almost all evaluation metrics.

		Predicted condition	
	Total population = P + N	Positive (PP)	Negative (PN)
Actual condition	Positive (P)	True positive (TP)	False negative (FN)
	Negative (N)	False positive (FP)	True negative (TN)

Figure 1.11. Conventional confusion matrix (2x2) to build the 4 components of TP, FP, FN, TN.

Single threshold metrics are problematic because they require the classification of objects into two classes (e.g., stable/unstable) by defining an a-priori cutoff susceptibility value, which is often not trivial (Frattini et al. 2010). In contrast, Receiver Operating Characteristic (ROC) curves, with the corresponding Area Under the Curve (AUC), are cutoff-independent and can be effectively used to visualize and compare the performance of models. Other possible metrics include Cost curves (Drummond & Holte 2000) and the Success-Rate curve. The Cost curves take into account costs and a-priori probabilities, making them suitable for evaluating landslide susceptibility model performance from a practical standpoint. However, Cost curves require assigning weights to each mapping unit, based on the aspect we wish to emphasize. For example, we can choose the area of slope units as weights, which are not applicable to grid-based models (Frattini et al., 2010) or focus on proximity to elements at risk. This process introduces arbitrary selections, leading to a lack of standardization, which is essential for landslide susceptibility applications (Reichenbach et al. 2018). Similarly, the Success-Rate curve presents some inconsistencies (both practical and theoretical) for grid-cell models (Frattini et al. 2010).

Given this scenario, AUC has become the benchmark for clear evaluation, although it has its drawbacks. AUC can appear overly positive and unrepresentative of the actual model performance, especially with highly imbalanced datasets, where the number of True Negatives (TN) is much greater than the other elements of the confusion matrix. This results in a False Positive Rate (FPR) that is systematically low, regardless of the actual number of True Positives (TP). In other words, AUC can be very high, close to 1, because it falsely suggests that the model is excellent at separating classes, simply due to the high number of TN, while there could be a significant number of False Positives (FP) or False Negatives (FN).

Therefore, it is recommended to use several evaluation metrics to complement the information provided by the AUC. However, other metrics, as mentioned earlier, require a single cut-off threshold, which may be unrepresentative or restrictive. Nevertheless, they can complement the AUC by providing a more focused evaluation of the predicted positive mapping units, particularly in scenarios with imbalanced datasets. The cut-off threshold is typically set at 0.5 of the resulting probability. This a priori selection assumes that the model is designed to discriminate in a binary context (0 for no landslide, 1 for landslide), and thus 0.5 is often considered the best cut-off. However, this assumption can lead to misleading interpretations (Berrar 2019). Despite this, single threshold metrics are commonly used and provide helpful comparisons between the results of different studies. Among these, Accuracy (ACC), Recall (REC), F-beta score, and Matthew's Correlation Coefficient (MCC) are commonly applied, each with its advantages and limitations depending on the final goal. For example, Accuracy (ACC) may often appear high in imbalanced datasets but fails to capture the model's performance in detecting actual landslides, a shortcoming that Recall (REC) addresses. F1 and F2 scores are particularly useful for understanding trade-offs between different kinds of errors: while F1 treats false positives and false negatives equally, F2 focuses more on reducing false negatives, making it more suitable for contexts where missing a landslide is more problematic than predicting false alarms. However, F-scores do not take True Negatives (TN) into account (Powers 2015), a limitation that MCC addresses, making it particularly useful in imbalanced data situations (Chicco and Jurman 2020). Moreover, Cao et al. (2020) proposed an MCC-F1 curve, combining the MCC and F1 scores with the multi-threshold advantage. However, it has not yet gained practical use as a reliable evaluation metric, making it difficult to use for comparison.

Finally, other evaluation metrics can be applied in a regression framework but are suitable for probabilistic outputs. One such metric is the Continuous Ranked Probability Score (CRPS), which is a generalized version of the Brier Score. For binary outcomes, the CRPS simplifies into the Brier Score, representing the mean squared error between predicted probability and actual binary outcome. CRPS is a valuable evaluation technique for models predicting continuous or ordinal outcomes, such as landslide volume, runout distance, or whatever landslide hazard magnitude, especially when the model predicts a full probability distribution, as shown by Opitz et al. (2022) and Dahal et al. (2024).

1.3.4 Landslide Susceptibility reclassification

The ultimate goal of landslide susceptibility maps is to assist stakeholders, municipalities, and construction companies in enhancing territorial information, in combination with the landslide inventory and monitoring techniques (e.g., satellite or site-specific ground data). However, communicating the most susceptible areas to decision-makers remains complex and non-standardized, leading to misunderstandings based on the study area and the specific goals of the project (Fell et al. 2008). This issue arises because decision-makers are typically not engineers or geologists, making it essential to present a clear and reliable classification of susceptibility, probability, or hazard. Rather than using a continuous range from 0 to 1, the data should be presented in clearly delineated classes (Reichenbach et al. 2018).

Various methods for reclassifying the probability in n classes exist in the literature, most of which are based on statistical distributions or arbitrary intervals. As shown in Figure 1.12, the differences in how reclassification methods visualize the same results are striking. For example, reclassifying the case study results from Chapter 2 of this thesis reveals significant differences across methods. Each approach emphasizes different aspects of the data distribution: Normal Distribution focuses on the mean and standard deviation; Natural Breaks aims to minimize variance within clusters in the data; percentile-based methods, though commonly used, are highly dependent on the chosen cutoff values (e.g., 25th, 50th, and 75th percentiles); Geometrical intervals create classes with increasing widths, particularly useful for skewed data distributions; and arbitrary selection relies on the minimum and maximum values expected from the model (0-1).

Selecting a reclassified map as the final product for stakeholders and decision-makers is highly subjective and, therefore, can be unreliable. Different reclassification methods can lead to vastly different decisions. For example, one map might encourage decision-makers to impose restrictions in one area, while another could suggest entirely different actions.

So, which reclassification method is the best? There is no absolute answer, as no single approach guarantees that the resulting classes will accurately represent the real variability of susceptibility across different landscapes. The suitability of each method depends on both the site and the model output. Each reclassification method emphasizes different aspects of the statistical distribution, and since model results can vary in distribution, it is best to apply specific reclassification methods tailored to each case. Therefore, there is no universal or standardized method. This concern will be explored further in Chapter 2 of this thesis.

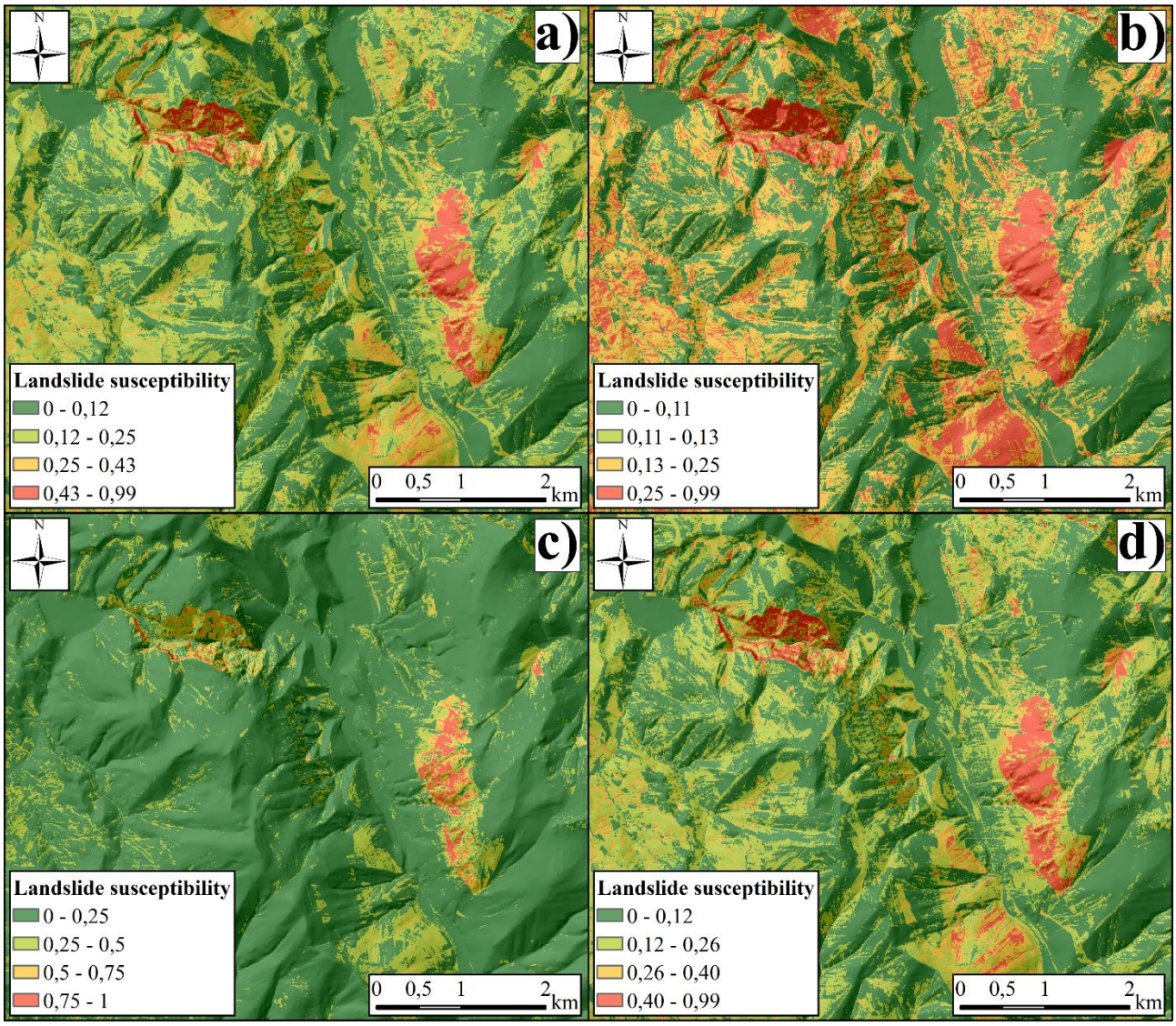


Figure 1.12. Same landslide susceptibility [0-1] reclassified in four classes using four different methods: ***a)*** *Natural Breaks (Jenks),* ***b)*** *Geometrical intervals,* ***c)*** *Arbitrary constant interval of 0.25,* ***d)*** *Normal distribution (mean ± standard deviation).*

The lack of standardization in reclassification was also noted by Fleuchaus et al. (2021) when comparing maps from different authors of the same area. Martinello et al. (2021) provide insights into how different mapping units can return varying classes using the same model, with Slope Units showing more consistency across the area, leading to more homogeneous and easier-to-interpret prediction maps. Cantarino et al. (2019) conducted a critical analysis of classification systems using various statistical methods. However, in cases where statistical performance is similar, they needed to introduce interpretation strategies to ensure a balanced approach to classification.

1.4 Landslide susceptibility in real world applications

Landslide susceptibility, initially designed to assist in territorial planning, has increasingly become a field focused on developing and testing higher-performance models, often losing sight of its original goal (Reichenbach et al., 2018). Despite the significant advances in

susceptibility and probability modeling, the actual application of landslide susceptibility, probability, or hazard maps remains rare among governments, civil protection agencies, and private companies. The following are some examples of real-world applications documented in the literature.

1.4.1 Land use and territorial planning

Governments' geological surveys and civil protection agencies often struggle to proactively use landslide susceptibility maps due to their bureaucratic structures and the complexity of communicating results and ensuring clear exposure and validation. Despite this, useful examples can be found in the literature.

The USGS has been notably progressive in incorporating landslide susceptibility maps into their programs. Although the models used may not always be the most advanced, Burns et al. (2016) produced a detailed and rigorous landslide susceptibility report for Oregon State, specifically to assist in regional planning and to identify areas needing more detailed landslide mapping. They emphasized that the product is for informational purposes and should not replace site-specific investigations by qualified practitioners. Similarly, Crawford et al. (2021) combined a bagged-trees and logistic regression approach to produce a landslide inventory and susceptibility map for Kentucky. A comprehensive report was also created by Hughes & Schulz (2020) for Puerto Rico using the frequency-ratio method. This study involved both academic institutions and government agencies in Puerto Rico, along with the USGS, aiming for practical use in planning, development, and emergency management. The resulting map is publicly available as georeferenced files that can be used in GIS platforms.

In Asia, Liu et al. (2024) studied landslide susceptibility in the Greater Xi'an Planning Area, China, using the Analytical Hierarchy Process (AHP). Their findings highlight that the Loess Plateau's topography and landslide mechanisms are similar to those of other cities, such as Lanzhou and Taiyuan, making the study relevant for future urban development in the region. In Bangladesh, Fuad et al. (2024) applied the Frequency Ratio model, focusing on the influence of human activities on slope instability. Their findings are crucial for policymakers and land-use planners, underscoring the importance of proactive measures to mitigate landslide risks and protect communities.

In Europe, Portugal has integrated landslide susceptibility maps into Municipal Emergency Plans, combining them with maps of exposed elements to assess exposure to landslides. Oliveira et al. (2017) analyzed the Loures municipality, located north of Lisbon, using a statistical bivariate Information Value method. They treated rotational slides, deep-seated translational slides, and shallow translational slides as separate inventory targets, which is crucial for an practical useful maps. They identified 34 classes of exposed elements from a civil protection and emergency management perspective. In Italy, the Autorità di Bacino, responsible for land-use management, sometimes uses landslide susceptibility maps in territorial decision-making. However, these maps are often neglected in practice, either because they are outdated or the class distinctions are unclear, as shown in Figure 1.13 by Roccati et al. (2021). The study also demonstrates the necessity and how to integrate new susceptibility maps with older ones.

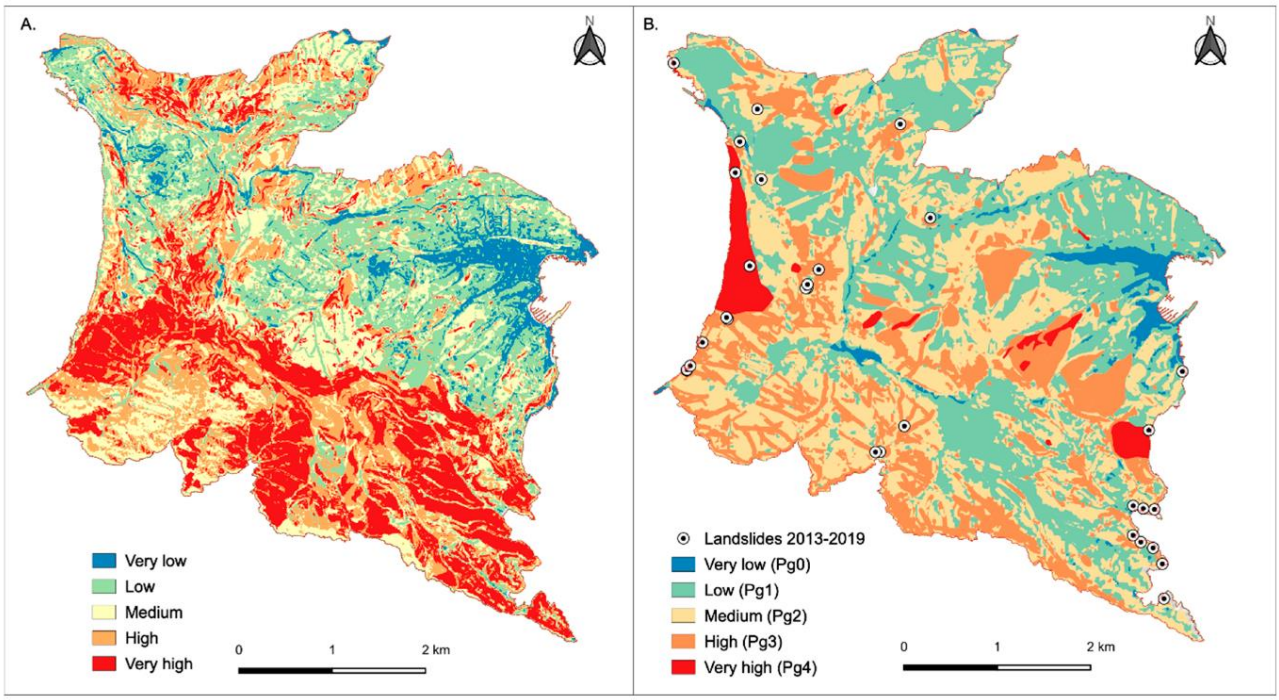


*Figure 1.13. Portofino (Liguria, GE) susceptibility maps, a comparison between: (**A**) the proposed landslide susceptibility map by Roccati et al. (2021), and (**B**) the landslide susceptibility zonation currently adopted by the regional land and risk management plan. Black dots represent rainfall induced landslides which occurred after 2013 (Roccati et al. 2021).*

1.4.2 Linear infrastructure (roads and railway)

Moving from general national and municipal territorial planning to specific infrastructure planning, particularly for linear infrastructure, the literature highlights that infrastructure, especially linear ones, can be more vulnerable to landslides due to their extensive coverage. Railways, in particular, are at a heightened risk compared to roads, primarily because of the larger curvature radius required for railways and the rigidity of both the construction materials and trains (Arslan et al. 2014). The risk is further compounded by the reduced capacity of trains to stop quickly, unlike road vehicles (Schafer et al. 2008). Furthermore, expanding the railway network plays a critical role in mitigating climate change, as rail transport is more environmentally friendly than road transport (Chester & Horvath 2010). In this context, climate change poses an increasing threat to the structural integrity of rail networks, making them more susceptible to the impacts of landslides (Schlögl & Matulla 2018).

Landslide susceptibility maps could be helpful during the feasibility phase of linear infrastructure route design. Several studies have explored this application, using varying methods, so following a stater of the art analysis is conducted.

- **Quinn et al. (2010):** The study was conducted in the Canadian Rockies, focusing on a regional scale along a railway corridor. It involved analyzing the landslide susceptibility of large areas prone to landslides affecting linear infrastructure. The authors applied the Weights of Evidence (WoE) method to large retrogressive landslide types. To relate the susceptibility results to the railway infrastructure, the authors overlaid the susceptibility map with the railway corridor, and considered the areas within 500 m, 1 km, and 2 km

from the infrastructure for further analysis. They then used air photo surveys to focus attention on areas with higher susceptibility. The innovative aspect of this work lies in its ability to reduce the area requiring further detailed investigation, narrowing it down to just 2% of the railway corridor based on susceptibility predictions. This targeted approach is efficient, making it easier for decision-makers to focus resources where they are most needed.

- **Bednarik et al. (2010):** The research focuses on a 50 km stretch of the Kraľovany–Liptovský Mikuláš railway in the Slovak Republic, applying a bivariate statistical approach utilizing the entropy index to assess landslide susceptibility. This method calculates the weight of each input parameter based on its contribution to landslide occurrence. The study did not differentiate between specific types of landslides. The focus was specifically on the railway network, analyzing the susceptibility of the land adjacent to the railway line.
- **Das et al. (2010):** The research was conducted along a road section of National Highway 108 in the northern Himalayas, India. The authors compared logistic regression (LR), a GIS-based quantitative modeling approach, with a rock mass classification system (SSPC) for landslide susceptibility mapping. The LR model used various geo-environmental factors (e.g., slope, aspect, lithology, land cover), while the SSPC method was based on rock mass parameters, such as intact rock strength, cohesion, friction angle, and the orientation of slope discontinuities. The study did not differentiate between specific landslide types but aimed to assess general landslide susceptibility in the area. The study concluded that both the logistic regression and SSPC methods are valuable for landslide susceptibility mapping. However, the SSPC method outperformed the logistic regression model in this specific case. This was attributed to the SSPC system's objective approach, which is based on rock mass parameters and includes a three-step process to classify slope rock mass. The SSPC method also allows for better adjustment based on discontinuity patterns and weathering, making it more effective for hill cut road sections. The authors concluded that geotechnical methods like SSPC can be more reliable for landslide susceptibility mapping, especially for infrastructure like highway road corridors, and should be considered in hazard and risk management programs.
- **Eker & Aydin (2016):** The research was conducted in the Gölyaka and Kardüz Forest Districts within the Gölyaka Forest Directorate, located in Düzce Province, Turkey. The authors employed the Logistic Regression (LR) method, a statistical approach integrated with GIS, to assess landslide susceptibility, without separating between different landslide types. The focus was specifically on forest road network overlaid with the province. The analysis revealed that approximately 1.3 km of roads are located within very high susceptibility areas, and 5.1 km are within high susceptibility areas. The study concluded that the LR method is effective for landslide susceptibility mapping in forested areas.
- **Martinović et al. (2016):** The research focused on the Athlone division of the Irish Rail network, specifically analyzing a 50 km stretch of railway, including 709 embankments

and 449 cuttings, with 74 rock cuttings, in whole Ireland country. The authors employed LR to assess landslide susceptibility along the railway network directly. The model used various geo-environmental factors, such as slope, vegetation, soil type, and annual rainfall, to predict areas vulnerable to landslides. The study focused on shallow translational landslides, which are the most common type affecting the railway infrastructure. Unlike traditional methods that overlay susceptibility maps with infrastructure corridors, the authors modelled the railway corridor directly, taking into account landslides that affected the railway itself, rather than relying on broader susceptibility areas. The study demonstrated that simple asset databases and visual survey data can effectively support landslide susceptibility assessments for railway networks. This methodology provides a proactive approach for identifying critical areas at risk, compared to the traditional reactive approach. The model achieved high accuracy, with an AUC value of 0.9 for the validation dataset.

- **Pellicani et al. (2017):** The authors employed a heuristic-bivariate statistical model that combined various conditioning factors (such as lithology, slope, soil type, rainfall, and land use) to create a landslide susceptibility map in the province of Matera (Basilicata region, Italy). This model was applied for the road network, specifically analysing provincial roads in the region, which spans approximately 1324 km and is a key element of regional connectivity. The landslide inventory used for model calibration and validation was derived from two sources: the River Basin Authority of Basilicata's Plan for the Hydrogeological Asset (PAI) and a field survey conducted along the provincial road network, identifying approximately 1280 landslides. The study not only provided a landslide susceptibility map but also extended the analysis to risk evaluation, incorporating both exposure and vulnerability of the road network. The authors used a matrix evaluation method to assess the risk of roads to landslides summarized in Figure 1.14. Susceptibility was evaluated based on terrain characteristics, while hazard was assessed combing susceptibility with landslide intensity (velocity and area) derived from the landslide inventory, based on the landslide type. Consequence was evaluated based on road vulnerability (previous damages and repairs) and exposure (traffic levels). This shift from susceptibility to risk analysis allowed the study to provide a more comprehensive understanding of the potential impacts on infrastructure, and to identify critical areas where mitigation measures should be prioritized. However, it is still restricted to the existing network, given the necessary information regarding the network, and is not properly designed for the study of new roads and railways. Overall, this approach is innovative because it combines hazard, vulnerability, and exposure in a GIS-based framework, making it a valuable tool for infrastructure planning and disaster risk management.

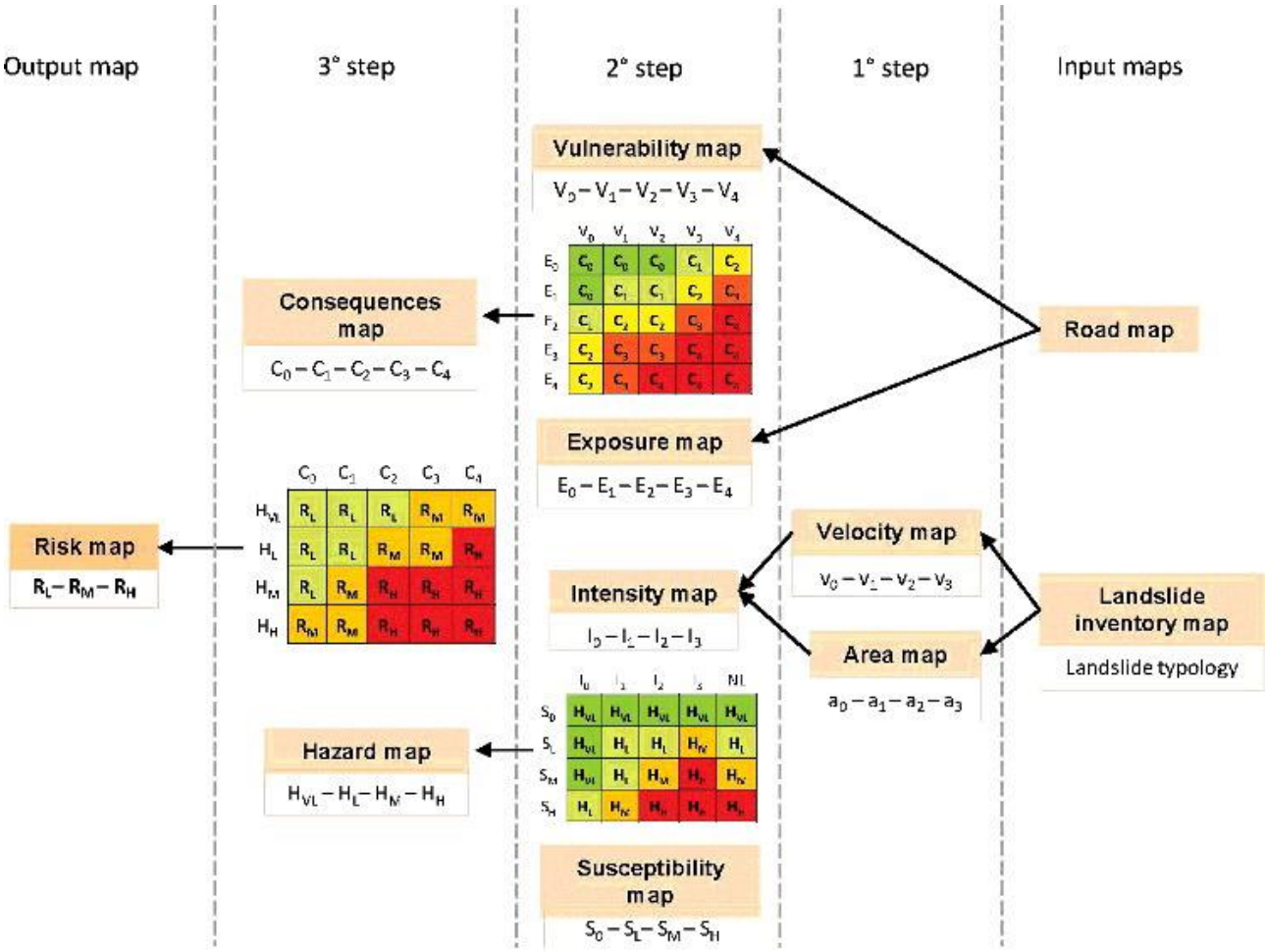


Figure 1.14. Procedure for assessing and mapping the landslide risk along roads corridos. Pellicani et al. (2017).

- **Liu et al. (2018):** The research focused on the Chinese railway system, encompassing both existing and planned railways across the country, which is pioneering. The study aimed to assess the susceptibility of these railways to rainfall-induced multi-hazards, including landslides, debris flows, and flooding, thus distinguishing between macro phenomena but not the landslide type apart from debris flows. The authors employed a machine learning approach, specifically the Random Forest (RF) algorithm, to analyze historical disaster events triggered between 1980 and 1998. The model was trained directly across the railway network, rather than post-intersecting susceptibility maps with infrastructure corridors. This was done by extracting the features of 1 x 1 km grid with the railway network before the analysis. Train a ML, like RF, with the dataset only from the railway network was possible thank to really long national Chinese railway system, providing enough data for the training. Although the model will be blind to what is some condition outside, perhaps being unreliable to predict new railway corridor susceptibility in unseen data. Given the multi-hazard approach, in particular the floods, the authors utilized the summer season daily rainfall, the rainband, and the climate zone, revealing their importance in influencing the results. Interesting, the authors compare the results with the China's climate change regionalization for future expectations. The

study achieved high predictive performance with accuracy > 93% for all the multi-hazard types in test-datasets. Overall, this approach enabled the identification of hotspot zones along the railways that require significant attention and preventative action from a national point of view, without being helpful in the design of specific individual works given the coarse resolution.

- **Bordoni et al. (2018):** The research was conducted in a 14 km² catchment located in the northeastern Oltrepò Pavese region of northern Italy. The authors employed a data-driven model based on the Generalized Additive Model (GAM) to identify road sectors susceptible to shallow landslides, specifically triggered in slopes upstream of the infrastructure. The analysis was applied to the road network, including provincial and municipal routes within the study area. Notably, the study utilized a 1-meter resolution Digital Elevation Model (DEM), plus the model incorporated an index of sediment connectivity (IC), calculated using both linear and non-linear approaches to evaluate the potential impact of sediment mobilization from upstream slope (Figure 1.15). The IC is defined as the potential connection between hillslopes and features that act as targets or storage areas (sinks) for mobilized sediments (e.g., channels, basin outlet, lakes or road networks). The authors further examined the possible changes in shallow landslide susceptibility given 3 different scenarios of land use, highlighting the importance of land management in mitigating landslide impacts.

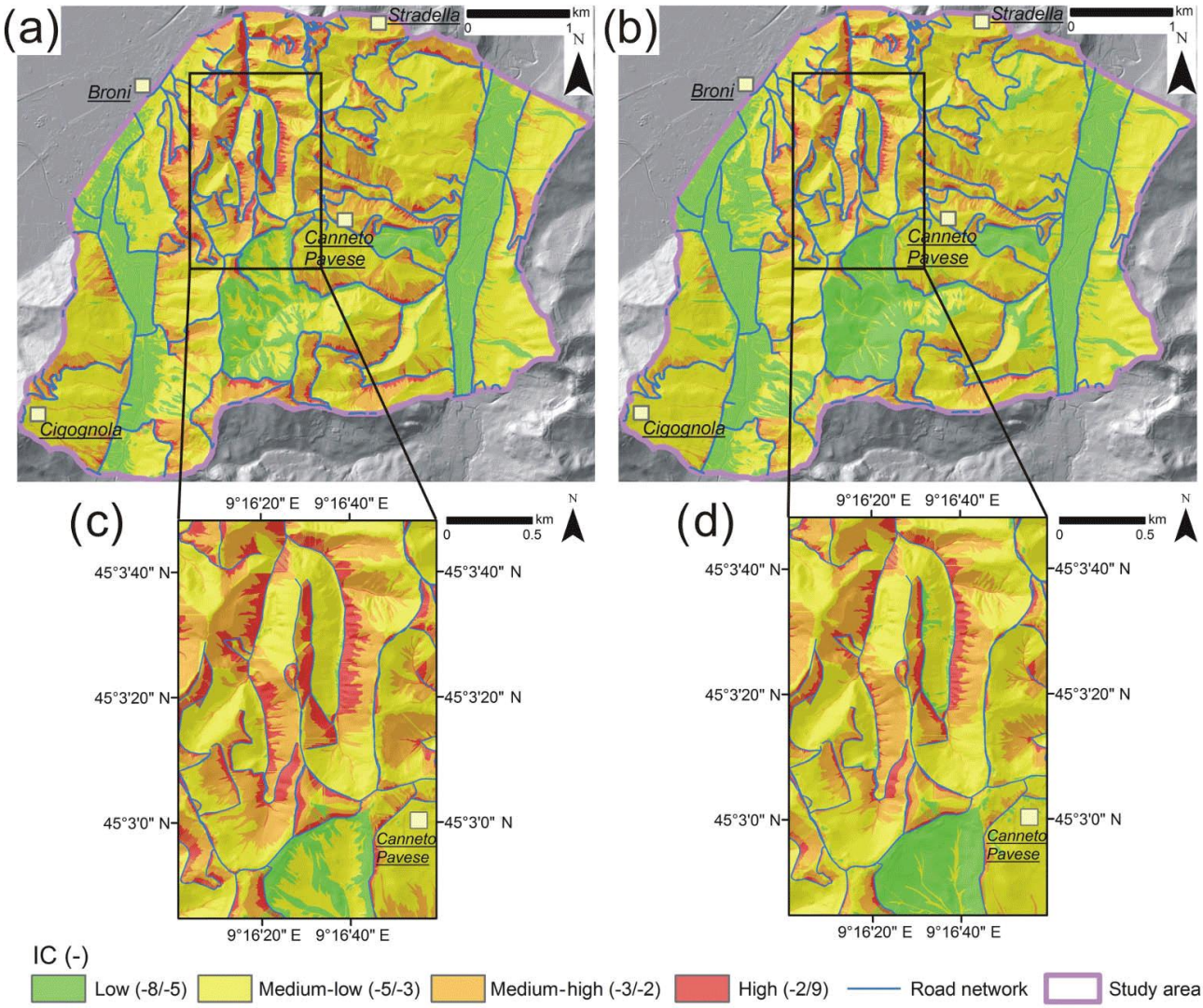


*Figure 1.15. Actual (2015) IC maps corresponding to the linear calculation of the index since linear **(a)** and to the non-linear calculation of the index since **(b)**. A detail of the northern sector of the study areas is reported for linear **(c)** and non-linear **(d)** maps. Bordoni et al. (2018).*

- **Meneses et al. (2019):** The study area is the Zêzere watershed, located in central Portugal. This area encompasses a diverse range of terrains, including mountainous regions and valleys. The authors employed the Information Value (IV) method to assess landslide susceptibility. Two models were developed: one utilizing detailed Land Use and Land Cover data from the Portuguese Land Cover Map (COS) and another using more generalized data from CORINE Land Cover (CLC). Both models incorporated a set of fixed independent layers as landslide predisposing factors, including slope angle, slope aspect, slope curvature, slope-over-area ratio, soil, and lithology. The analysis was focused solely on shallow landslide on the entire study area with 10 meter resolution, then intersected with road network to identify areas with varying degrees of susceptibility. The study highlighted the significant impact of land cover data properties on landslide susceptibility mapping. The model using the more detailed COS data produced better results in terms of accuracy and spatial resolution, leading to more precise identification of high-risk areas along the road network.

- **Pasang & Kubíček (2020):** The research focused on an 81 km stretch of the Asian Highway (AH48) in Bhutan, extending approximately 5 km on either side of the road. The study focused on shallow landslides affecting the road network, with a particular emphasis on those triggered by road construction activities. The aim of the study was also to compare three statistical methods, Information value (IV), Weight of Evidence (WoE), and logistic regression (LR), using various environmental factors, and then intersect with the highway. For that specific area, the authors found that all three statistical methods were capable of producing reliable landslide susceptibility maps, with the IV method showing the best predictive performance. The findings underscore the importance of selecting appropriate modelling techniques for landslide susceptibility assessments.
- **Sur et al. (2020):** The research focused on the Kalsi–Chakrata road corridor in the Lesser Himalayas of India, an area prone to frequent landslides, especially during the monsoon season. The focus was on the shallow landslides, which are the prevalent in the area. The study employed the Fuzzy Analytic Hierarchy Process (Fuzzy AHP), integrating it with geospatial technology to assess landslide susceptibility. This approach combines the subjective judgment of experts with fuzzy logic to handle uncertainty in the evaluation of various landslide conditioning factors. The authors utilized good-resolution (2.5 -10 m) remote sensing images and validated landslide inventory data to map and classify various conditioning factors. The model demonstrated an accuracy of 86.52%. The innovative aspect of this study lies in the application of the Fuzzy AHP technique, which effectively integrates expert judgment and geospatial data to assess landslide susceptibility in the region.
- **Sujatha (2021):** The research focused on a 6 km stretch of hill road severely affected by various types of landslides along the Kodaikkanal-Palani traffic corridor in the Western Ghats of India. The area is affected mainly by translational slides, rock falls, debris flows, checked and mapped by the authors. The study aims to predict all the landslide types in one model without distinction between them. The authors employed the Analytical Network Process (ANP), a heuristic model capable of accommodating the interdependencies among various conditioning factors influencing landslides. The model was directly applied to the narrow corridor of the analysed road. The study demonstrated that the ANP model effectively provide a comprehensive assessment of landslide susceptibility
- **Wang et al. (2022):** The study focused on the Ya'an-Linzhi section of the Sichuan-Tibet Railway in China, with a buffer of 25 km around the railway. Interesting, the study utilized a geospatial database generated from 587 landslide events identified through Interferometric Synthetic Aperture Radar (InSAR) stacking technology and field geological hazard surveys. So not past event but actual movement, perhaps less cautious and comprehensive but more accurate for nowadays environments. The authors choose two deep learning algorithms for landslide susceptibility mapping: Convolutional Neural Network (CNN) and Deep Neural Network (DNN). These models were chosen for their ability to handle complex spatial patterns and interactions among various

landslide-influencing factors. The results indicated that the CNN model (AUC = 0.88) slightly outperformed the DNN model (AUC = 0.84) in both training and testing phases. The study identified that slope, elevation, and rainfall were the primary factors influencing landslide occurrence. The innovative aspect of this research lies in the application of deep learning models, specifically CNN and DNN, and the use of InSAR to identify the target for training the models.

- **Dou et al. (2023):** The research analysed the national, provincial and country roads (432.479 km) of Youxi County, Central Fujian Province, China. The model was designed to assess susceptibility to all types of landslides, without distinguishing between specific mechanisms. The authors implemented 3 ML methods: SVM, RF, and Artificial Neural network (ANN), to model landslide susceptibility; then, the prediction accuracy of the models was compared using five-fold cross-validation, statistical accuracy indexes, and AUC values. Furthermore, they combined the three models in an ensemble approach, where each model's predictions are weighted and then averaged to minimize prediction errors. To achieve that, the authors choose to implement a global minimum variance portfolio (GMVP) method, which use the Lagrange multiplier method to find the best weights (sum to 1) that minimize errors. The combined model's validation results revealed that it outperforms the single model in terms of mean accuracy and robustness, as well as having a significantly higher mean AUC (reported in Figure 1.16).

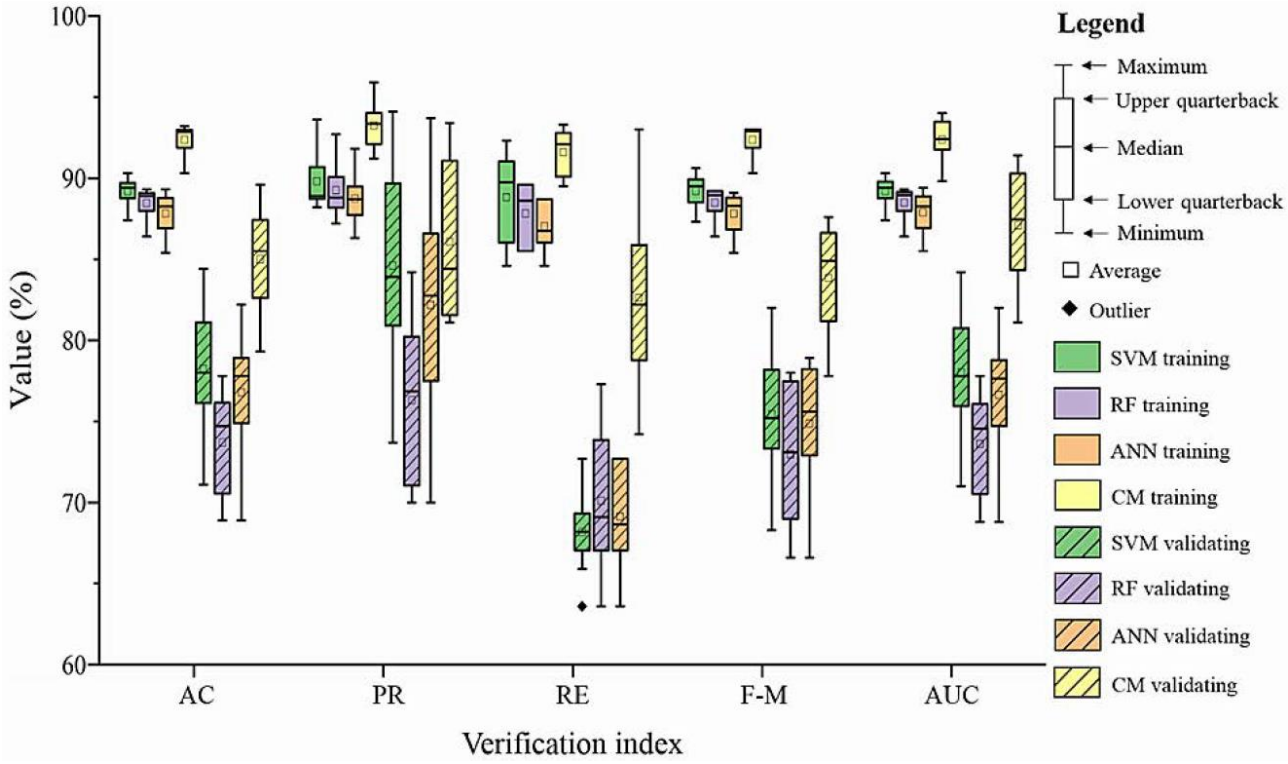


Figure 1.16. The box plot of model accuracy comparison. Notes: SVM = Support Vector Machine; RF = Random Forests; ANN = Artificial Neural Networks; CM =Combined Model; and AC=Accuracy, PR= precision, RE= Recall, F-M = F-score, AUC = Area under the curve. Dou et al. (2023).

- **Marchesini et al. (2024):** The study focuses on Italy's national railway network, encompassing approximately 17,000 km of tracks. The study specifically targets rapid flow-like landslides, including debris flows and mudflows, which can travel significant distances and cause extensive damage to infrastructure. The research introduces an innovative multi-phase methodology: the identification of the source area and the runout modelling. The authors used a Topographic Threshold (TT) approach, which use using slope gradient and drainage area, to evaluate the relative probability of each DEM cell (10x10 m) to be a source area. Then they employ the r.randomwalk model to simulate landslide runout paths. Finally, railway segments of 1 km are assessed for their vulnerability to flow-like landslide, considering factors like proximity to identified source areas and potential runout zones. The methodology relies solely on existing landslide inventories and DEMs, making it a cost-effective tool for infrastructure management.
- **Jiang et al. (2024):** The research was conducted in three different study area in China, where 3 segments of railway is studied using a buffer of 10 km wide. The model was designed to assess susceptibility to all types of landslides, with no distinction. The authors imply four models: LR, SVM, Long-Short term memory (LSTM), and Linformer. The last one was innovatively used, it is a "simplified" version of a transformer which avoids quadratic operations by reducing the complexity of sequence length in training from $O(n^2)$ to $O(n)$. Furthermore, the authors introduce the side-sampling, highlighting the importance of sampling strategy. The proposed side-sampling focuses on a single research cell and extends outward to a fixed window, thereby collecting information within a broader area and further mapping this information to the location of the subject cell (Figure 1.17). In this study they used a grid-cell database of 30 m resolution. The study demonstrates the effectiveness in geospatial data analysis of Linformer, plus the proposed side-sampling outperforms traditional models in predicting landslide susceptibility.

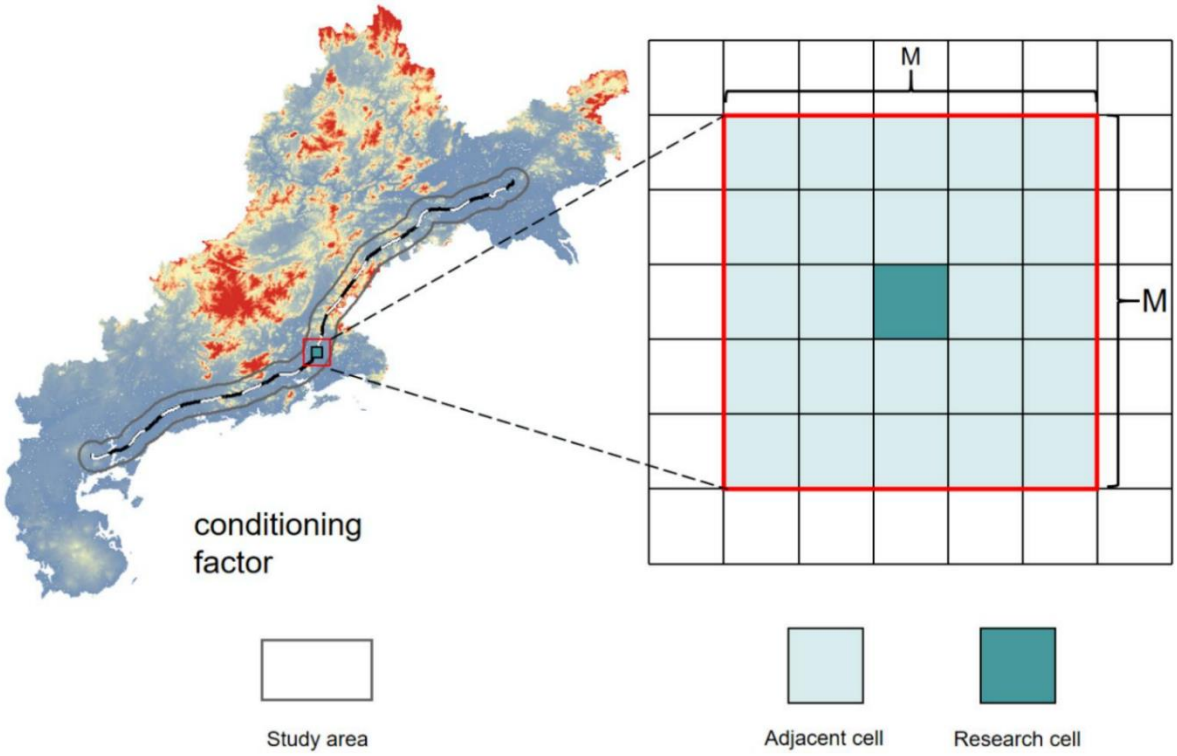


Figure 1.17. Diagrammatic sketch of the side-sampling method. Jiang et al. (2024).

- **Zhou M. et al. (2024):** The study focuses on the road network of six counties adjacent to Sichuan and Yunnan Provinces in China. The authors did not separate the landslide types from the inventory. The study is comprehensive for considering also the vulnerability of the road network. The authors apply an Analytic Hierarchy Process (AHP) to determine the landslide susceptibility given six influencing factors (a in Figure 1.18) while the Entropy Weight Method (EWM) was utilized to establish objective (road) weights. The amalgamation of these diverse weighting methods was then optimized following the principle of least squares, resulting in the derivation of combined weights. For the vulnerability the authors choose the following six influencing factors: Road cost, Transport output value, Highway anti-disaster ability index, Road network density, Gross Domestic Product (GDP), Population Density (b in Figure 1.18). Finally, they applied the landslide risk formula to integrate susceptibility and vulnerability, offering a quantitative approach to risk analysis to roads and linear infrastructure (c in Figure 1.18).

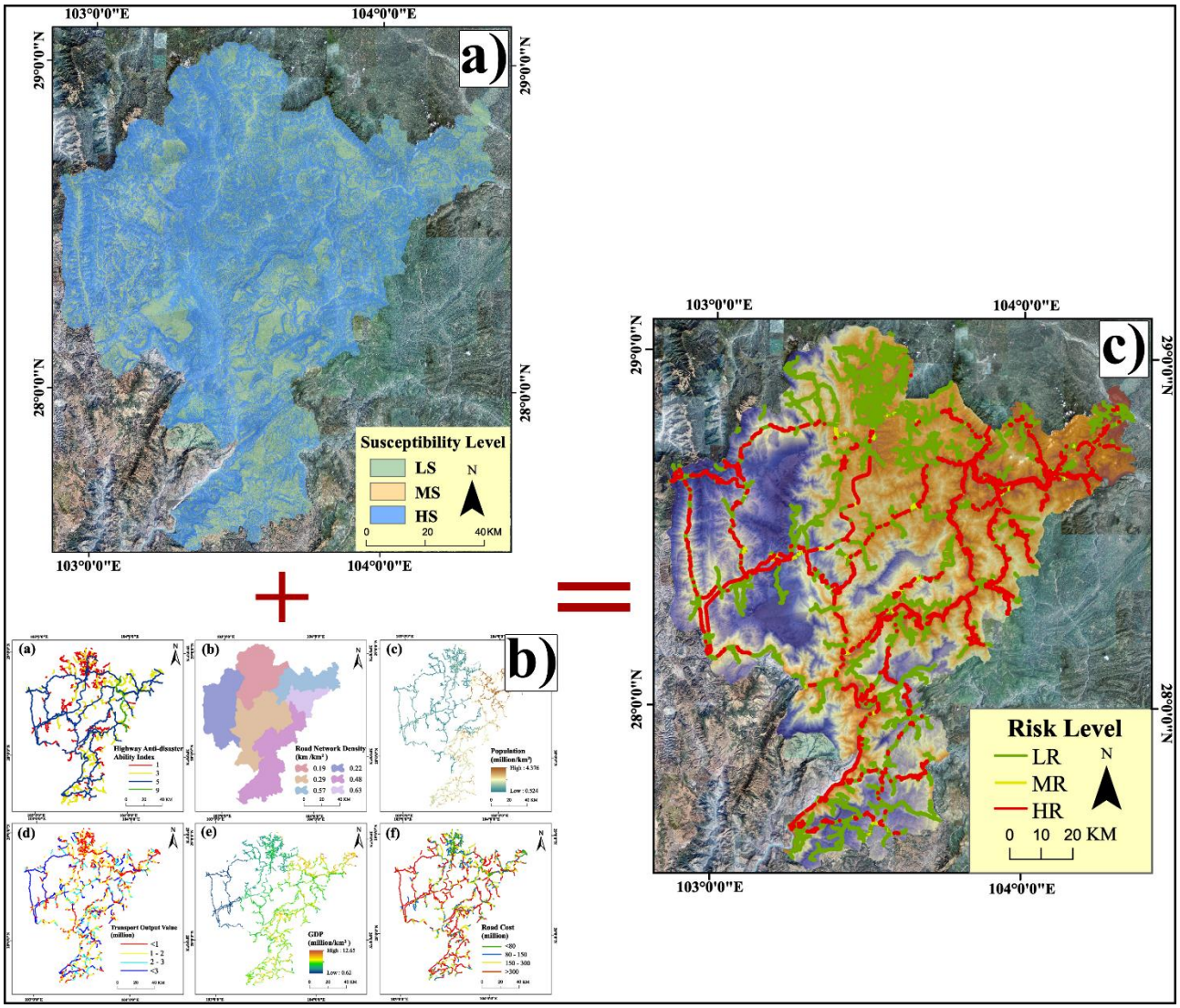


*Figure 1.18. The three products: **a)** Susceptibility map, which mainly includes three levels: low, moderate and high; **b)** Vulnerability related factors: (a) highway anti-disaster ability index, (b) road network density, (c) population, (d) transport output value, (e) GDP, (f) road cost; **c)** Risk map, which mainly includes three levels: low, moderate and high. Zhou M. et al. (2024) (modified).*

- **Singh et al. (2024):** The study area encompasses the under construction Imphal-Jiribam railway line, in the Tamenglong and Senapati districts, Manipur, India. The study employs four efficient statistical methods for landslide susceptibility mapping: Frequency Ratio (FR), Information Value (IoV), Weight of Evidence (WoE), Weighted Linear Combination (WLC). The authors did not consider separately different landslide types. Landslide susceptibility maps generated using the statistical methods were overlaid with the railway corridor to assess the potential impact of landslides on the infrastructure. The results of IoV obtained higher AUC and precision values for testing datasets, while higher values for recall, OA, F-score, and BA were obtained for the WLC method. The FR model had a lower value for all the statistical measures, yet all the models performed well in landslide susceptibility mapping which in turn was used in

predicting landslides along the Imphal-Jiribam railway line, which is under construction, thus in advance analysis to prevent issues.

- **Li et al. (2025):** The study area is located in Menyuan–Shandanmachang (China), where is present a section of the Lanzhou–Xinjiang high-speed railway. A large number of landslides introduced by earthquakes like the landslide used in the training phase, triggered by the 2016 Menyuan Ms6.4 Earthquake. The authors applied three ML models: LR, SVM, and RF with a 7:3 split method. Then, the trained models were used to predict the landslide susceptibility for the 2016 Menyuan Ms6.4 Earthquake and the 2022 Menyuan Ms6.9 Earthquake. Interestingly, the prediction results were contrasted by landslide samples and the inversion results of InSAR surface deformation to evaluate the performance of the ML models. The landslide warning degree of RF is higher than that of the other two models, but it performs poorly in the detailed depiction of low-seismicity areas. Although the prediction accuracy of the SVM model is slightly lower than that of RF, it has higher stability and generalization ability aligning with the slope deformation. The authors also evaluate the influence of the predators to the landslide susceptibility, making it clearer. The research applies landslide susceptibility models specifically to a railway corridor, demonstrating the practical application of these models in infrastructure planning and risk management. Especially, the authors also plot the resulting susceptibility and InSAR deformations along the railway corridor (Figure 1.19), this visualization could be helpful in practical applications pointing out critical section rapidly.

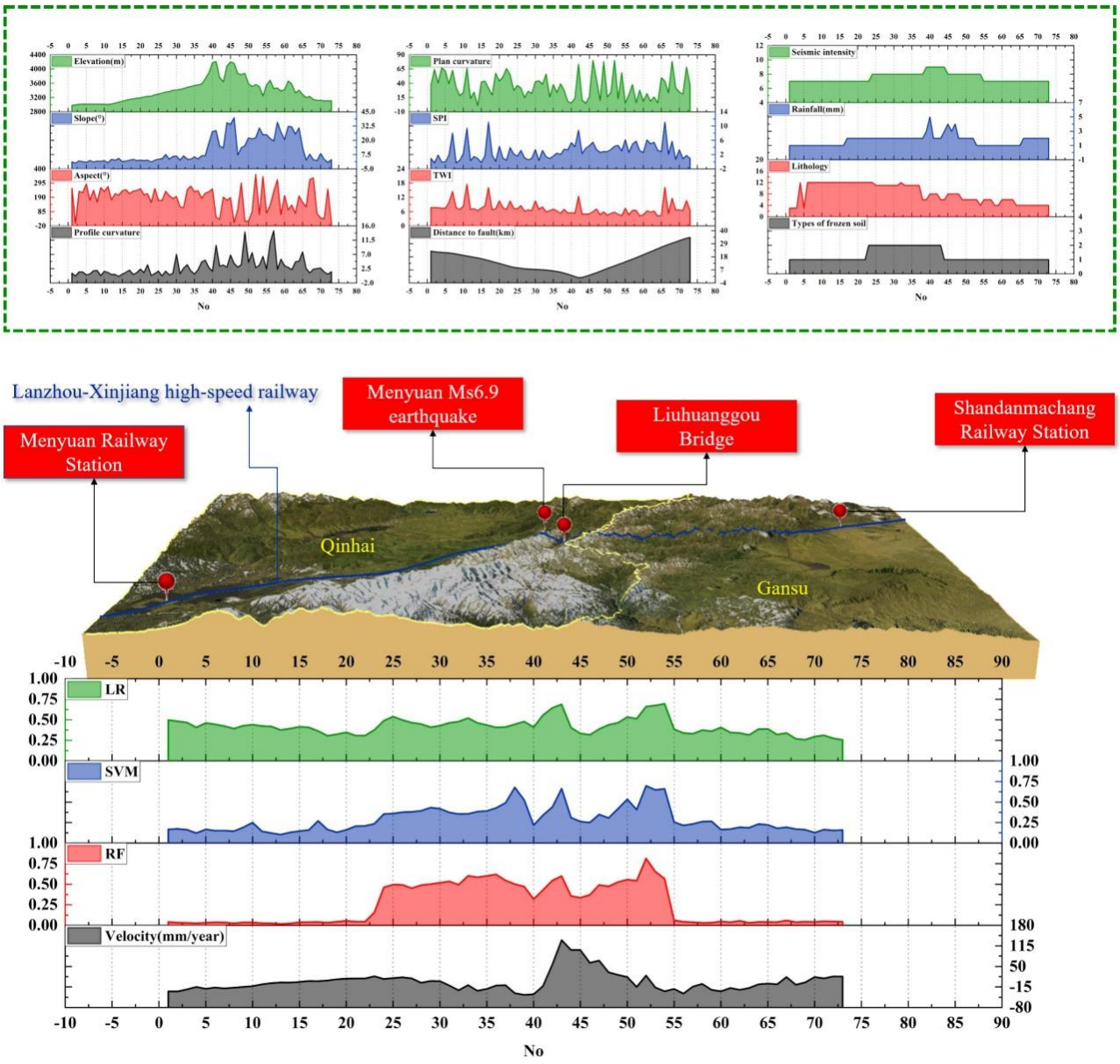


Figure 1.19. The changes of 12 influencing factors, the landslide susceptibility prediction results of 3 models and the surface deformation results based on InSAR along the Lanzhou-Xinjiang high-speed railway (Menyuan Railway Station - Shandanmachang Railway Station). Li et al. (2025).

1.5 Research Questions and Outline

In the previous sections, we conducted a comprehensive review of the existing literature, identifying key challenges and advancements in the field. While a significant portion of the literature has focused on enhancing model accuracy and alignment with the landslide inventory (ground truth), leading to the development of highly effective and promising predictive models, several challenges remain. Despite advances in machine learning techniques and the growing availability of landslide inventories, applying these methods in real-world scenarios remains difficult. This is particularly true in regions where geological, meteorological, and topographical conditions are highly variable and where a consistent, universally accepted outcome is needed.

The primary objective of this PhD research is to conduct a comprehensive landslide susceptibility analysis and quantitative hazard assessment, tailored to the design of linear infrastructure during the planning phase. Given the fragility of railway infrastructure and its potential impact on the surrounding environment, it is essential to understand landslide occurrences within a specific region. For broad assessments, regional-scale approaches using slope units are highly effective, while site-specific evaluations for detailed planning require fine pixel-based methods. This research seeks to address the gap in producing reliable and practical landslide susceptibility maps that are directly applicable to real-world scenarios. To achieve this, five core research concerns have been identified, which form the foundation of the research questions.

How can we conduct a comprehensive susceptibility analysis in an area prone to different types of landslides?

The simplification of assuming just one landslide type in the analysis, or treating different types as a single category, can lead to misleading results and a loss of valuable information. Each landslide type is associated with distinct failure mechanisms, which are triggered and propagated under unique conditions, governed by specific conditioning factors. Furthermore, each type requires its own tailored mitigation strategy. Treating them uniformly neglects these differences, potentially leading to ineffective risk management. By focusing limited resources on the most susceptible areas for each landslide type, the overall efficiency of hazard mitigation can be significantly improved. This issue was addressed in Chapters 2, 3, and 4, where we developed landslide-type specific susceptibility and probability maps.

Do we need detailed and classified landslide inventories for susceptibility analysis? How can we produce them?

As highlighted in section 1.2.3, the quality and completeness of the landslide inventory used to train the model play a critical role in the accuracy of the resulting susceptibility map. For practical applications, it is essential to carefully assess the characteristics of the inventory. If the inventory is incomplete, the map may be overly cautious and biased, whereas a broader, less detailed inventory may lead to an overly conservative model. Thus, it is crucial to have a reliable and well-classified landslide inventory for accurate predictions. Chapter 3 addresses this need by providing a detailed landslide mapping of the dramatic May 2023 event in Emilia Romagna, Italy. Then, in Chapter 4, this detailed inventory is used as the ground truth for the analysis.

How can we effectively reclassify the landslide susceptibly results to communicate with non-technical final decision makers?

The need for a clear and accessible reclassification output, which can be easily understood by non-technical decision-makers, was thoroughly discussed in section 1.3.4. Furthermore, during the design phase of a railway route, it is crucial to obtain reliable susceptibility classes to guide decisions on where to plan the railway. In Chapter 2, a methodology is proposed that utilizes several reclassification methods (ensemble method) to enhance the reliability of the resulting susceptibility classes.

How can preparatory factors (antecedent rainfall) be successfully integrated into susceptibility analysis with specific or multiple landslide events?

Integrating preparatory factors into the analysis is one of the most challenging aspects. While predisposing factors are generally static and easily available over time (spanning seasons and centuries), triggering factors are typically instantaneous and more directly linked to the triggered landslide. Therefore, the question is: how do preparatory factors influence landslide occurrence leading up to the triggering event? In Chapter 4, the relationship between antecedent rainfall and two major landslide events (separated by just 14 days) in Emilia-Romagna, Italy, is explored using a Transformer Neural Network (TNN). In the literature, is common to compress the continuous rainfall time series into single scalar values for specific aggregation periods antecedent the main triggered event. Although, this could be limited to the chosen period and require the exact day of triggering information. The method proposed in Chapter 4 overcome this problem taking the full time series (encompassing antecedent and triggering rainfall), capable of founding the relationship between preparatory and triggering factors with the landslide occurrences.

In developing countries with limited data on landslides, how can we conduct a reliable analysis of debris flow risk?

In developing countries, the documentation of past landslides is often lacking, while many regions are increasingly affected by heavy tropical rainfall, exacerbated by climate change. As a result, there is a significant gap in territorial knowledge regarding past recurrent landslide events. Additionally, the tropical environment, with its fast-growing vegetation, makes it challenging to apply automated mapping techniques to build inventories of past events. In this context, the use of geomorphological and topographical empirical equations offers a practical solution for estimating the susceptibility of debris flow-type landslides. The methodologies discussed in Chapter 5 provide an efficient, cost-effective approach to assessing debris flow risk to infrastructure in developing countries, requiring only a digital elevation model and satellite-derived land cover information.

1.6 References Chapter 1

Ado, M., Amitab, K., Maji, A. K., Jasińska, E., Gono, R., Leonowicz, Z., & Jasiński, M. (2022). Landslide susceptibility mapping using machine learning: A literature survey. Remote Sensing, 14(13), 3029. https://doi.org/10.3390/rs14133029

Agterberg, F. P. (1992). Combining indicator patterns in weights of evidence modeling for resource evaluation. Nonrenewable Resources, 1(1), 39–50. https://doi.org/10.1007/BF01782111

Aleotti, P., & Chowdhury, R. (1999). Landslide hazard assessment: Summary review and new perspectives. Bulletin of Engineering Geology and the Environment, 58(1), 21–44. https://doi.org/10.1007/s100640050066

Alvioli, M., Marchesini, I., Reichenbach, P., Rossi, M., Ardizzone, F., Fiorucci, F., & Guzzetti, F. (2016). Automatic delineation of geomorphological slope units with r.slopeunits v1.0 and their optimization for landslide susceptibility modeling. Geoscientific Model Development, 9, 3975–3991. https://doi.org/10.5194/gmd-9-3975-2016

Alvioli, M., Loche, M., Jacobs, L., Grohmann, C. H., Abraham, M. T., Gupta, K., Satyam, N., Scaringi, G., Bornaetxea, T., Rossi, M., Marchesini, I., Lombardo, L., Moreno, M., Steger, S., Camera, C. A. S., Bajni, G., Samodra, G., Wahyudi, E. E., Susyanto, N., Sinčić, M., Gazibara, S. B., Sirbu, F., Torizin, J., Schüßler, N., Mirus, B. B., Woodard, J. B., Aguilera, H., & Rivera-Rivera, J. S. (2024). A benchmark dataset and workflow for landslide susceptibility zonation. Earth-Science Reviews, 258, 104927. https://doi.org/10.1016/j.earscirev.2024.104927

Ahmed, M., Tanyaș, H., Huser, R., Dahal, A., Titti, G., Borgatti, L., Francioni, M., & Lombardo, L. (2023). Dynamic rainfall-induced landslide susceptibility: A step towards a unified forecasting system. International Journal of Applied Earth Observation and Geoinformation, 125, 103593. https://doi.org/10.1016/j.jag.2023.103593

Amato, G., Eisank, C., Castro-Camilo, D., & Lombardo, L. (2019). Accounting for covariate distributions in slope-unit-based landslide susceptibility models: A case study in the alpine environment. Engineering Geology, 260, 105237. https://doi.org/10.1016/j.enggeo.2019.105237

Ardizzone, F., Bucci, F., Cardinali, M., Fiorucci, F., Pisano, L., Santangelo, M., & Zumpano, V. (2023). Geomorphological landslide inventory map of the Daunia Apennines, southern Italy. Earth System Science Data, 15, 753–767. https://doi.org/10.5194/essd-15-753-2023

Arslan, A., Tari, E., Ziatdinov, R., & Nabiyev, R. I. (2014). Transition curve modeling with kinematical properties: Research on log-aesthetic curves. Computer-Aided Design and Applications, 11(5), 509–517. https://doi.org/10.1080/16864360.2014.902680

Atkinson, P. M., & Massari, R. (1998). Generalised linear modelling of susceptibility to landsliding in the central Apennines, Italy. Computers & Geosciences, 24(4), 373–385. https://doi.org/10.1016/S0098-3004(97)00117-9

Ayalew, L., & Yamagishi, H. (2005). The application of GIS-based logistic regression for landslide susceptibility mapping in the Kakuda-Yahiko Mountains, Central Japan. Geomorphology, 65(1-2), 15-31. https://doi.org/10.1016/j.geomorph.2004.06.010

Azarafza, M., Azarafza, M., Akgün, H., Atkinson, P. M., & Derakhshani, R. (2021). Deep learning-based landslide susceptibility mapping. Scientific Reports, 11(1), Article 24112. https://doi.org/10.1038/s41598-021-03585-1

Batar, A. K., & Watanabe, T. (2021). Landslide susceptibility mapping and assessment using geospatial platforms and weights of evidence (WoE) method in the Indian Himalayan Region: Recent developments, gaps, and future directions. ISPRS International Journal of Geo-Information, 10(3), Article 114. https://doi.org/10.3390/ijgi10030114

Baum, R. L., Savage, W. Z., & Godt, J. W. (2008). TRIGRS: A Fortran program for transient rainfall infiltration and grid-based regional slope-stability analysis, version 2.0 (US Geological Survey Open-File Report 2008-1159, 75 p.). Reston, VA, USA: U.S. Geological Survey. https://doi.org/10.3133/ofr20081159

Bednarik, M., Magulová, B., Matys, M., & Marschalko, M. (2010). Landslide susceptibility assessment of the Kraľovany–Liptovský Mikuláš railway case study. Physics and Chemistry of the Earth, Parts A/B/C, 35(3–5), 162–171. https://doi.org/10.1016/j.pce.2009.12.002

Berrar, D. (2019). Performance measures for binary classification. https://doi.org/10.1016/B978-0-12-809633-8.20351-8

Berti, M., Ciccarese, G., Fanti, R., Lapietra, I., Parise, M., & Liso, I. S. (2025b). Deliverable title: Rationale for selecting and scale-dependent weighing of predisposing factors.

Bhuyan, K., Rana, K., Ferrer, J. V., Cotton, F., Ozturk, U., Catani, F., & Malik, N. (2024). Landslide topology uncovers failure movements. Nature Communications, 15(1), 2633. https://doi.org/10.1038/s41467-024-46741-7

Bonham-Carter, G. (1994). Geographic information systems for geoscientists: Modelling with GIS (No. 13). Elsevier.

Bordoni, M., Persichillo, M. G., Meisina, C., Crema, S., Cavalli, M., Bartelletti, C., Galanti, Y., Barsanti, M., Giannecchini, R., & D'Amato Avanzi, G. (2018). Estimation of the susceptibility of a road network to shallow landslides with the integration of the sediment connectivity. Natural Hazards and Earth System Sciences, 18(6), 1735–1758. https://doi.org/10.5194/nhess-18-1735-2018

Bordoni, M., Vivaldi, V., Lucchelli, L., Ciabatta, L., Brocca, L., Galve, J. P., & Meisina, C. (2021). Development of a data-driven model for spatial and temporal shallow landslide probability of occurrence at catchment scale. Landslides, 18, 1209–1229. https://doi.org/10.1007/s10346-020-01592-3

Bornaetxea, T., Blais-Stevens, A., Miller, B., & Marchesini, I. (2025). Combination of statistical and conceptual approaches for debris-flow susceptibility modelling at a regional scale, British Columbia, Canada. Catena, 256, 109044. https://doi.org/10.1016/j.catena.2025.109044

Brabb, E. E., Pampeyan, E. H., & Bonilla, M. G. (1972). Landslide susceptibility in San Mateo County, California (No. 360). U.S. Geological Survey. https://doi.org/10.3133/mf360

Brenning, A. (2008). Statistical geocomputing combining R and SAGA: The example of landslide susceptibility analysis with generalized additive models. Hamburger Beiträge zur Physischen Geographie und Landschaftsökologie, 19(23–32), 410.

Broothaerts, N., Kissi, E., Poesen, J., Van Rompaey, A., Getahun, K., Van Ranst, E., & Diels, J. (2012). Spatial patterns, causes and consequences of landslides in the Gilgel Gibe catchment, SW Ethiopia. Catena, 97, 127–136. https://doi.org/10.1016/j.catena.2012.05.011

Budimir, M. E. A., Atkinson, P. M., & Lewis, H. G. (2015). A systematic review of landslide probability mapping using logistic regression. Landslides, 12, 419-436. https://doi.org/10.1007/s10346-014-0550-5

Burns, W. J., Mickelson, K. A., & Madin, I. (2016). Landslide susceptibility overview map of Oregon. Oregon Department of Geology and Mineral Industries.

Camilo, D. C., Lombardo, L., Mai, P. M., Dou, J., & Huser, R. (2017). Handling high predictor dimensionality in slope-unit-based landslide susceptibility models through LASSO-penalized Generalized Linear Model. Environmental Modelling & Software, 97, 145-156. https://doi.org/10.1016/j.envsoft.2017.08.003

Cao, C., Chicco, D., & Hoffman, M. M. (2020). The MCC-F1 curve: a performance evaluation technique for binary classification. arXiv preprint arXiv:2006.11278. https://doi.org/10.48550/arXiv.2006.11278

Carrara, A. (1983). Multivariate models for landslide hazard evaluation. Journal of the International Association for Mathematical Geology, 15, 403-426. https://doi.org/10.1007/BF01031290

Carrara, A., Cardinali, M., Detti, R., Guzzetti, F., Pasqui, V., & Reichenbach, P. (1991). GIS techniques and statistical models in evaluating landslide hazard. Earth surface processes and landforms, 16(5), 427-445. https://doi.org/10.1002/esp.3290160505

Chang, Z., Catani, F., Huang, F., Liu, G., Meena, S. R., Huang, J., & Zhou, C. (2023). Landslide susceptibility prediction using slope unit-based machine learning models considering the heterogeneity of conditioning factors. Journal of Rock Mechanics and Geotechnical Engineering, 15(5), 1127-1143. https://doi.org/10.1016/j.jrmge.2022.07.009

Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: synthetic minority over-sampling technique. Journal of artificial intelligence research, 16, 321-357. https://doi.org/10.1613/jair.953

Chen, C., & Fan, L. (2024). Interpretability of Statistical, Machine Learning, and Deep Learning Models for Landslide Susceptibility Mapping in Three Gorges Reservoir Area. arXiv preprint arXiv:2405.11762. https://doi.org/10.48550/arXiv.2405.11762

Chen, T., Wang, Q., Zhao, Z., Liu, G., Dou, J., & Plaza, A. (2024). LCFSTE: Landslide conditioning factors and swin transformer ensemble for landslide susceptibility assessment. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. https://doi.org/10.1109/JSTARS.2024.3373029

Chester, M., & Horvath, A. (2010). Life-cycle assessment of high-speed rail: The case of California. Environmental Research Letters, 5(1), 014003. https://doi.org/10.1088/1748-9326/5/1/014003

Chicco, D., & Jurman, G. (2020). The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation. BMC Genomics, 21, 6. https://doi.org/10.1186/s12864-019-6413-7

Chung, C. J. F., & Fabbri, A. G. (1999). Probabilistic prediction models for landslide hazard mapping. Photogrammetric engineering and remote sensing, 65(12), 1389-1399.

Clerici, A., Perego, S., Tellini, C., & Vescovi, P. (2006). A GIS-based automated procedure for landslide susceptibility mapping by the conditional analysis method: the Baganza valley case

study (Italian Northern Apennines). Environmental Geology, 50(7), 941-961. https://doi.org/10.1007/s00254-006-0264-7

Crawford, M. M., Dortch, J. M., Koch, H. J., Killen, A. A., Zhu, J., Zhu, Y., ... & Haneberg, W. C. (2021). Using landslide-inventory mapping for a combined bagged-trees and logistic-regression approach to determining landslide susceptibility in eastern Kentucky, USA. Quarterly Journal of Engineering Geology and Hydrogeology, 54(4), qjegh2020-177. https://doi.org/10.1144/qjegh2020-177

Dahal, A., & Lombardo, L. (2023). Explainable artificial intelligence in geoscience: A glimpse into the future of landslide susceptibility modeling. Computers & geosciences, 176, 105364. https://doi.org/10.1016/j.cageo.2023.105364

Dahal, A., Huser, R., & Lombardo, L. (2024). At the junction between deep learning and statistics of extremes: formalizing the landslide hazard definition. Journal of Geophysical Research: Machine Learning and Computation, 1(3), e2024JH000164. https://doi.org/10.1029/2024JH000164

Dal Seno, N., Evangelista, D., Piccolomini, E., & Berti, M. (2024). Comparative analysis of conventional and machine learning techniques for rainfall threshold evaluation under complex geological conditions. Landslides, 21(12), 2893-2911. https://doi.org/10.1007/s10346-024-02336-3

Das, I., Sahoo, S., van Westen, C., Stein, A., & Hack, R. (2010). Landslide susceptibility assessment using logistic regression and its comparison with a rock mass classification system, along a road section in the northern Himalayas (India). Geomorphology, 114(4), 627-637. https://doi.org/10.1016/j.geomorph.2009.09.023

Donat, M. G., Lowry, A. L., Alexander, L. V., O'Gorman, P. A., & Maher, N. (2016). More extreme precipitation in the world's dry and wet regions. Nature Climate Change, 6(5), 508-513. https://doi.org/10.1038/nclimate2941

Dou, J., Yamagishi, H., Pourghasemi, H. R., Yunus, A. P., Song, X., Xu, Y., & Zhu, Z. (2015). An integrated artificial neural network model for the landslide susceptibility assessment of Osado Island, Japan. Natural Hazards, 78, 1749-1776. https://doi.org/10.1007/s11069-015-1799-2

Dou, H. Q., Huang, S. Y., Jian, W. B., & Wang, H. (2023). Landslide susceptibility mapping of mountain roads based on machine learning combined model. Journal of Mountain Science, 20(5), 1232-1248. https://doi.org/10.1007/s11629-022-7657-2

Drummond, C., & Holte, R. C. (2000, August). Explicitly representing expected cost: An alternative to ROC representation. In Proceedings of the sixth ACM SIGKDD international conference on Knowledge discovery and data mining (pp. 198-207).

Eker, R., & Aydin, A. (2016). Landslide susceptibility assessment of forest roads. European Journal of Forest Engineering, 2(2), 54-60.

Fang, Z., Tanyas, H., Gorum, T., Dahal, A., Wang, Y., & Lombardo, L. (2023). Speech-recognition in landslide predictive modelling: A case for a next generation early warning system. Environmental Modelling & Software, 170, 105833. https://doi.org/10.1016/j.envsoft.2023.105833

Fell, R., Corominas, J., Bonnard, C., Cascini, L., Leroi, E., Savage, W. Z., & JTC-1 Joint Technical Committee on Landslides and Engineered Slopes. (2008). Guidelines for landslide

susceptibility, hazard and risk zoning for land use planning. Engineering geology, 102(3-4), 85-98. https://doi.org/10.1016/j.enggeo.2008.03.022

Fleuchaus, P., Blum, P., Wilde, M., Terhorst, B., & Butscher, C. (2021). Retrospective evaluation of landslide susceptibility maps and review of validation practice. Environmental Earth Sciences, 80, 1-15. https://doi.org/10.1007/s12665-021-09770-9

Frattini, P., Crosta, G., & Carrara, A. (2010). Techniques for evaluating the performance of landslide susceptibility models. Engineering geology, 111(1-4), 62-72. https://doi.org/10.1016/j.enggeo.2009.12.004

Fuad, N., Meandad, J., Haque, A., Sultana, R., Anwar, S. B., & Sultana, S. (2024). Landslide vulnerability analysis using frequency ratio (FR) model: a study on Bandarban district, Bangladesh. arXiv preprint arXiv:2407.20239. https://doi.org/10.48550/arXiv.2407.20239

Gaidzik, K., & Ramírez-Herrera, M. T. (2021). The importance of input data on landslide susceptibility mapping. Scientific reports, 11(1), 19334. https://doi.org/10.1038/s41598-021-98830-y

Galli, M., Ardizzone, F., Cardinali, M., Guzzetti, F., & Reichenbach, P. (2008). Comparing landslide inventory maps. Geomorphology, 94(3-4), 268-289. https://doi.org/10.1016/j.geomorph.2006.09.023

Gariano, S. L., & Rianna, G. (2025). How will the projected climate change influence rainfall-induced landslides in Europe? A review of modelling approaches. Landslides, 1-17. https://doi.org/10.1007/s10346-025-02550-7

Ghojogh, B., & Crowley, M. (2019). The theory behind overfitting, cross validation, regularization, bagging, and boosting: tutorial. arXiv preprint arXiv:1905.12787. https://doi.org/10.48550/arXiv.1905.12787

Goodchild, M. (2010). Twenty years of progress: GIScience in 2010. Journal of spatial information science, (1), 3-20. 10.5311/JOSIS.2010.1.2. https://doi.org/10.5311/JOSIS.2010.1.2

Gu, T., Duan, P., Wang, M., Li, J., & Zhang, Y. (2024). Effects of non-landslide sampling strategies on machine learning models in landslide susceptibility mapping. Scientific reports, 14(1), 7201. https://doi.org/10.1038/s41598-024-57964-5

Gunning, D., & Aha, D. (2019). DARPA's explainable artificial intelligence (XAI) program. AI magazine, 40(2), 44-58. https://doi.org/10.1609/aimag.v40i2.2850

Guo, Z., Tian, B., Zhu, Y., He, J., & Zhang, T. (2024). How do the landslide and non-landslide sampling strategies impact landslide susceptibility assessment?—A catchment-scale case study from China. Journal of Rock Mechanics and Geotechnical Engineering, 16(3), 877-894. https://doi.org/10.1016/j.jrmge.2023.07.026

Guzzetti, F., Carrara, A., Cardinali, M., & Reichenbach, P. (1999). Landslide hazard evaluation: a review of current techniques and their application in a multi-scale study, Central Italy. Geomorphology, 31(1-4), 181-216. https://doi.org/10.1016/S0169-555X(99)00078-1

Guzzetti, F., Reichenbach, P., Cardinali, M., Galli, M., & Ardizzone, F. (2005). Probabilistic landslide hazard assessment at the basin scale. Geomorphology, 72(1-4), 272-299. https://doi.org/10.1016/j.geomorph.2005.06.002

Guzzetti, F., Mondini, A. C., Cardinali, M., Fiorucci, F., Santangelo, M., & Chang, K. T. (2012). Landslide inventory maps: New tools for an old problem. Earth-Science Reviews, 112(1-2), 42-66. https://doi.org/10.1016/j.earscirev.2012.02.001

Hastie, T., & Tibshirani, R. (1986). Generalized additive models. Statistical science, 1(3), 297-310.

He, J., Tanyas, H., Dahal, A., Huang, D., & Lombardo, L. (2025). Spatial prediction of InSAR-derived hillslope velocities via deep learning. Bulletin of Engineering Geology and the Environment, 84(3), 1-14. https://doi.org/10.1007/s10064-025-04161-x

Hölbling, D., Eisank, C., Albrecht, F., Vecchiotti, F., Friedl, B., Weinke, E., & Kociu, A. (2017). Comparing manual and semi-automated landslide mapping based on optical satellite images from different sensors. Geosciences, 7(2), 37. https://doi.org/10.3390/geosciences7020037

Huabin, W., Gangjun, L., Weiya, X., & Gonghui, W. (2005). GIS-based landslide hazard assessment: an overview. Progress in Physical geography, 29(4), 548-567. https://doi.org/10.1191/0309133305pp462

Huang, Y., & Zhao, L. (2018). Review on landslide susceptibility mapping using support vector machines. Catena, 165, 520-529. https://doi.org/10.1016/j.catena.2018.03.003

Hughes, K. S., & Schulz, W. (2020). Map depicting susceptibility to landslides triggered by intense rainfall, Puerto Rico (No. 2020-1022). US Geological Survey. https://doi.org/10.3133/ofr20201022

Hussin, H. Y., Zumpano, V., Reichenbach, P., Sterlacchini, S., Micu, M., van Westen, C., & Bălteanu, D. (2016). Different landslide sampling strategies in a grid-based bi-variate statistical susceptibility model. Geomorphology, 253, 508-523. https://doi.org/10.1016/j.geomorph.2015.10.030

Jacobs, L., Kervyn, M., Reichenbach, P., Rossi, M., Marchesini, I., Alvioli, M., & Dewitte, O. (2020). Regional susceptibility assessments with heterogeneous landslide information: Slope unit-vs. pixel-based approach. Geomorphology, 356, 107084. https://doi.org/10.1016/j.geomorph.2020.107084

Jebur, M. N., Pradhan, B., & Tehrany, M. S. (2014). Optimization of landslide conditioning factors using very high-resolution airborne laser scanning (LiDAR) data at catchment scale. Remote Sensing of Environment, 152, 150-165. https://doi.org/10.1016/j.rse.2014.05.013

Jiang, N., Li, Y., Han, Z., Yang, J., Fu, B., Li, J., & Li, C. (2024). A side-sampling based Linformer model for landslide susceptibility assessment: a case study of the railways in China. Geomatics, Natural Hazards and Risk, 15(1), 2354507. https://doi.org/10.1080/19475705.2024.2354507

Kavzoglu, T., Sahin, E. K., & Colkesen, I. (2015). Selecting optimal conditioning factors in shallow translational landslide susceptibility mapping using genetic algorithm. Engineering Geology, 192, 101-112. https://doi.org/10.1016/j.enggeo.2015.04.004

Kavzoglu, T., Colkesen, I., & Sahin, E. K. (2019). Machine learning techniques in landslide susceptibility mapping: a survey and a case study. Landslides: Theory, practice and modelling, 283-301. https://doi.org/10.1007/978-3-319-77377-3_13

Kikstra, J. S., Nicholls, Z. R., Smith, C. J., Lewis, J., Lamboll, R. D., Byers, E., ... & Riahi, K. (2022). The IPCC Sixth Assessment Report WGIII climate assessment of mitigation pathways: from emissions to global temperatures. Geoscientific Model Development, 15(24), 9075-9109. https://doi.org/10.5194/gmd-15-9075-2022

Kim, H., Lee, J. H., Park, H. J., & Heo, J. H. (2021). Assessment of temporal probability for rainfall-induced landslides based on nonstationary extreme value analysis. Engineering Geology, 294, 106372. https://doi.org/10.1016/j.enggeo.2021.106372

Kritikos, T., & Davies, T. (2015). Assessment of rainfall-generated shallow landslide/debris-flow susceptibility and runout using a GIS-based approach: application to western Southern Alps of New Zealand. Landslides, 12, 1051-1075. https://doi.org/10.1007/s10346-014-0533-6

Lee, S., & Choi, J. (2004). Landslide susceptibility mapping using GIS and the weight-of-evidence model. International Journal of Geographical Information Science, 18(8), 789-814. https://doi.org/10.1080/13658810410001702003

Lee, S., & Sambath, T. (2006). Landslide susceptibility mapping in the Damrei Romel area, Cambodia using frequency ratio and logistic regression models. Environmental Geology, 50, 847-855. https://doi.org/10.1007/s00254-006-0256-7

Lei, X., Liu, H., Chen, Z., Li, S., Chen, H., Zeng, S., ... & Picco, L. (2025). Investigating the landslide susceptibility assessment methods for multi-scale slope units based on SDGSAT-1 and Graph Neural Networks. International Journal of Digital Earth, 18(1), 2468913. https://doi.org/10.1080/17538947.2025.2468913

Li, L., Lan, H., Guo, C., Zhang, Y., Li, Q., & Wu, Y. (2017). A modified frequency ratio method for landslide susceptibility assessment. Landslides, 14, 727-741. https://doi.org/10.1007/s10346-016-0771-x

Li, J., Zhou, Z., & Ma, W. (2025). Assessment of landslide susceptibility along the Lanzhou-xinjiang high-speed railway: A case study of Menyuan-Shandanmachang. Transportation Geotechnics, 50, 101473. https://doi.org/10.1016/j.trgeo.2024.101473

Liao, M., Wen, H., & Yang, L. (2022). Identifying the essential conditioning factors of landslide susceptibility models under different grid resolutions using hybrid machine learning: a case of Wushan and Wuxi counties, China. Catena, 217, 106428. https://doi.org/10.1016/j.catena.2022.106428

Lima, P., Steger, S., Glade, T., & Mergili, M. (2023). Conventional data-driven landslide susceptibility models may only tell us half of the story: potential underestimation of landslide impact areas depending on the modeling design. Geomorphology, 430(108638), 10-1016. https://doi.org/10.1016/j.geomorph.2023.108638

Liu, K., Wang, M., Cao, Y., Zhu, W., & Yang, G. (2018). Susceptibility of existing and planned Chinese railway systems subjected to rainfall-induced multi-hazards. https://doi.org/10.1016/j.tra.2018.08.030

Liu, S., Wang, L., Zhang, W., He, Y., & Pijush, S. (2023). A comprehensive review of machine learning-based methods in landslide susceptibility mapping. Geological Journal, 58(6), 2283-2301. https://doi.org/10.1002/gj.4666

Liu, X., Shao, S., & Shao, S. (2024). Landslide susceptibility zonation using the analytical hierarchy process (AHP) in the Great Xi'an Region, China. Scientific reports, 14(1), 2941. https://doi.org/10.1038/s41598-024-53630-y

Liu, X., Shao, S., Zhang, C., & Shao, S. (2025). Effect of different mapping units, spatial resolutions, and machine learning algorithms on landslide susceptibility mapping at the township scale. Environmental Earth Sciences, 84(5), 138. https://doi.org/10.1007/s12665-025-12148-w

Loche, M., Alvioli, M., Marchesini, I., Bakka, H., & Lombardo, L. (2022). Landslide susceptibility maps of Italy: Lesson learnt from dealing with multiple landslide types and the uneven spatial distribution of the national inventory. Earth-Science Reviews, 232, 104125. https://doi.org/10.1016/j.earscirev.2022.104125

Lombardo, L., & Tanyas, H. (2020). Chrono-validation of near-real-time landslide susceptibility models via plug-in statistical simulations. Engineering geology, 278, 105818. https://doi.org/10.1016/j.enggeo.2020.105818

Lou, Y., Caruana, R., Gehrke, J., & Hooker, G. (2013). Accurate intelligible models with pairwise interactions. In Proceedings of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (pp. 623–631). https://doi.org/10.1145/2487575.2487579

Lu, J., He, Y., Zhang, L., Zhang, Q., Gao, B., Chen, H., & Fang, Y. (2024). Ensemble learning landslide susceptibility assessment with optimized non-landslide samples selection. Geomatics, Natural Hazards and Risk, 15(1), 2378176. https://doi.org/10.1080/19475705.2024.2378176

Lu, Z. Y., Liu, G. Y., Zhao, X. D., Sun, K., Chen, Y. S., Song, Z. H., ... & Yang, M. S. (2025). Landslide susceptibility assessment based on an interpretable coupled FR-RF model: A case study of Longyan City, Fujian Province, Southeast China. China Geology, 8(2), 281-294. https://doi.org/10.31035/cg2024123

Ma, P., Chen, L., Yu, C., Zhu, Q., Ding, Y., Wu, Z., Li, H., Tian, C., & Fan, X. (2025). Dynamic landslide susceptibility mapping over last three decades to uncover variations in landslide causation in subtropical urban mountainous areas. Remote Sensing of Environment, 326, 114800. https://doi.org/10.1016/j.rse.2025.114800

Ma, S., Shao, X., & Xu, C. (2023). Landslide susceptibility mapping in terms of the slope-unit or raster-unit, which is better?. Journal of Earth Science, 34(2), 386-397. https://doi.org/10.1007/s12583-021-1407-1

Marchesini, I., Althuwaynee, O., Santangelo, M., Alvioli, M., Cardinali, M., Mergili, M., Reichenbach, P., Peruccacci, S., Balducci, V., Agostino, I., Esposito, R., & Rossi, M. (2024). National-scale assessment of railways exposure to rapid flow-like landslides. Engineering Geology, 332, 107474. https://doi.org/10.1016/j.enggeo.2024.107474

Martinello, C., Cappadonia, C., Conoscenti, C., Agnesi, V., & Rotigliano, E. (2021). Optimal slope units partitioning in landslide susceptibility mapping. Journal of Maps, 17(3), 152-162. https://doi.org/10.1080/17445647.2020.1805807

Martinović, K., Gavin, K., & Reale, C. (2016). Development of a landslide susceptibility assessment for a rail network. Engineering Geology, 215, 1-9. https://doi.org/10.1016/j.enggeo.2016.10.011

Meena, S. R., Puliero, S., Bhuyan, K., Floris, M., & Catani, F. (2022). Assessing the importance of conditioning factor selection in landslide susceptibility for the province of Belluno (region of Veneto, northeastern Italy). Natural hazards and earth system sciences, 22(4), 1395-1417. https://doi.org/10.5194/nhess-22-1395-2022

Meneses, B. M., Pereira, S., & Reis, E. (2019). Effects of different land use and land cover data on the landslide susceptibility zonation of road networks. Natural Hazards and Earth System Sciences, 19(3), 471-487. https://doi.org/10.5194/nhess-19-471-2019

Mergili, M., Schwarz, L., & Kociu, A. (2019). Combining release and runout in statistical landslide susceptibility modeling. Landslides, 16, 2151-2165. https://doi.org/10.1007/s10346-019-01222-7

Montgomery, D. R., & Dietrich, W. E. (1994). A physically based model for the topographic control on shallow landsliding. Water resources research, 30(4), 1153-1171. https://doi.org/10.1029/93WR02979

Moreno, M., Steger, S., Tanyas, H., & Lombardo, L. (2023). Modeling the area of co-seismic landslides via data-driven models: The Kaikōura example. Engineering Geology, 320, 107121. https://doi.org/10.1016/j.enggeo.2023.107121

Myhre, G., Alterskjær, K., Stjern, C. W., Hodnebrog, Ø., Marelle, L., Samset, B. H., ... & Stohl, A. (2019). Frequency of extreme precipitation increases extensively with event rareness under global warming. Scientific reports, 9(1), 16063. https://doi.org/10.1038/s41598-019-52277-4

Nair, V., & Hinton, G.E. (2010). Rectified linear units improve restricted boltzmann machines. Proceedings of the 27th International Conference on Machine Learning (ICML-10), 807-814.

Nam, K., Kim, J., & Chae, B. G. (2024). Exploring class imbalance with under-sampling, over-sampling, and hybrid sampling based on Mahalanobis distance for landslide susceptibility assessment: a case study of the 2018 Iburi earthquake induced landslides in Hokkaido, Japan. Geosciences Journal, 28(1), 71-94. https://doi.org/10.1007/s12303-023-0033-6

Nava, L., Carraro, E., Reyes-Carmona, C., Puliero, S., Bhuyan, K., Rosi, A., Monserrat, O., Floris, M., Meena, S.R., Galve, J.P., & Catani, F. (2023). Landslide displacement forecasting using deep learning and monitoring data across selected sites. Landslides, 20(10), 2111-2129. https://doi.org/10.1007/s10346-023-02104-9

Neuland, H. (1976). A prediction model of landslips. Catena, 3(2), 215-230. https://doi.org/10.1016/0341-8162(76)90011-4

Oliveira, S. C., Zêzere, J. L., Guillard-Gonçalves, C., Garcia, R. A., & Pereira, S. (2017). Integration of landslide susceptibility maps for land use planning and civil protection emergency management. In Advancing Culture of Living with Landslides: Volume 1 ISDR-ICL Sendai Partnerships 2015-2025 (pp. 543-553). Springer International Publishing. https://doi.org/10.1007/978-3-319-59469-9_49

Opitz, T., Bakka, H., Huser, R., & Lombardo, L. (2022). High-resolution Bayesian mapping of landslide hazard with unobserved trigger event. The Annals of Applied Statistics, 16(3), 1653-1675. https://doi.org/10.1214/21-AOAS1561

Ozdemir, A., & Altural, T. (2013). A comparative study of frequency ratio, weights of evidence and logistic regression methods for landslide susceptibility mapping: Sultan Mountains, SW Turkey. Journal of Asian Earth Sciences, 64, 180-197. https://doi.org/10.1016/j.jseaes.2012.12.014

Parker, R. N., Hales, T. C., Mudd, S. M., Grieve, S. W., & Constantine, J. A. (2016). Colluvium supply in humid regions limits the frequency of storm-triggered landslides. Scientific Reports, 6(1), 34438. https://doi.org/10.1038/srep34438

Pasang, S., & Kubíček, P. (2020). Landslide susceptibility mapping using statistical methods along the Asian Highway, Bhutan. Geosciences, 10(11), 430. https://doi.org/10.3390/geosciences10110430

Patanè, G., Bortolotti, T., Yordanov, V., Biagi, L. G. A., Brovelli, M. A., Truong, X. Q., & Vantini, S. (2025). An interpretable and transferable model for shallow landslides detachment combining spatial Poisson point processes and generalized additive models. Stochastic Environmental Research and Risk Assessment, 39(4), 1723-1740. https://doi.org/10.1007/s00477-025-02943-z

Pellicani, R., Argentiero, I., & Spilotro, G. (2017). GIS-based predictive models for regional-scale landslide susceptibility assessment and risk mapping along road corridors. Geomatics, Natural Hazards and Risk, 8(2), 1012-1033. https://doi.org/10.1080/19475705.2017.1292411

Powers, D. M. (2020). Evaluation: From precision, recall and F-measure to ROC, informedness, markedness and correlation. arXiv preprint arXiv:2010.16061. https://arxiv.org/abs/2010.16061

Pradhan, B., & Lee, S. (2010). Delineation of landslide hazard areas on Penang Island, Malaysia, by using frequency ratio, logistic regression, and artificial neural network models. Environmental Earth Sciences, 60, 1037-1054. https://doi.org/10.1007/s12665-009-0245-8

Quinn, P. E., Hutchinson, D. J., Diederichs, M. S., & Rowe, R. K. (2010). Regional-scale landslide susceptibility mapping using the weights of evidence method: An example applied to linear infrastructure. Canadian Geotechnical Journal, 47(8), 905–927. https://doi.org/10.1139/T10-019

Radbruch-Hall, D.H. The systematic evaluation of landslide incidence and susceptibility in the United States. Bulletin of the International Association of Engineering Geology **16**, 82–86 (1977). https://doi.org/10.1007/BF02591456

Regmi, N. R., Giardino, J. R., McDonald, E. V., & Vitek, J. D. (2014). A comparison of logistic regression-based models of susceptibility to landslides in western Colorado, USA. Landslides, 11, 247-262. https://doi.org/10.1007/s10346-012-0380-2

Reichenbach, P., Rossi, M., Malamud, B. D., Mihir, M., & Guzzetti, F. (2018). A review of statistically-based landslide susceptibility models. Earth-science reviews, 180, 60-91. https://doi.org/10.1016/j.earscirev.2018.03.001

Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). " Why should i trust you?" Explaining the predictions of any classifier. In Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining (pp. 1135-1144). https://doi.org/10.1145/2939672.2939778

Roccati, A., Paliaga, G., Luino, F., Faccini, F., & Turconi, L. (2021). GIS-based landslide susceptibility mapping for land use planning and risk assessment. Land, 10(2), 162. https://doi.org/10.3390/land10020162

Sahin, E. K. (2022). Comparative analysis of gradient boosting algorithms for landslide susceptibility mapping. Geocarto International, 37(9), 2441-2465. https://doi.org/10.1080/10106049.2020.1831623

Sahin, E. K. (2023). Implementation of free and open-source semi-automatic feature engineering tool in landslide susceptibility mapping using the machine-learning algorithms RF,

SVM, and XGBoost. Stochastic Environmental Research and Risk Assessment, 37(3), 1067-1092. https://doi.org/10.1007/s00477-022-02330-y

Santangelo, M., Marchesini, I., Bucci, F., Cardinali, M., Fiorucci, F., & Guzzetti, F. (2015). An approach to reduce mapping errors in the production of landslide inventory maps. Natural Hazards and Earth System Sciences, 15(9), 2111–2126. https://doi.org/10.5194/nhess-15-2111-2015

Santangelo, M., Althuwaynee, O., Alvioli, M., Bucci, F., Cardinali, M., Fiorucci, F., & Guzzetti, F. (2023). Inventory of landslides triggered by an extreme rainfall event in Marche-Umbria, Italy, on 15 September 2022. Scientific Data, 10, 427. https://doi.org/10.1038/s41597-023-02336-3

Schafer, D. H., & Barkan, C. P. L. (2008). A prediction model for broken rails and an analysis of their economic impact. In Proceedings of the American Railway Engineering and Maintenance-of-Way Association (AREMA) Annual Conference.

Schlögl, M., & Matulla, C. (2018). Potential future exposure of European land transport infrastructure to rainfall-induced landslides throughout the 21st century. Natural Hazards and Earth System Sciences, 18(4), 1121–1132. https://doi.org/10.5194/nhess-18-1121-2018

Segoni, S., Nocentini, N., Barbadori, F., Medici, C., Gatto, A., Rosi, A., & Casagli, N. (2025). A novel prototype national-scale landslide nowcasting system for Italy combining rainfall thresholds and risk indicators. Landslides, 1-26. https://doi.org/10.1007/s10346-024-02452-0

Shi, Q., Xu, X., Li, J., Deng, H., Zhang, Q., Tan, D., & He, Y. (2025). Spatiotemporal effect driven landslide susceptibility mapping at fine scales: a deep learning model based on multidimensional feature fusion and source data adaptation. Engineering Applications of Artificial Intelligence, 156, 110924. https://doi.org/10.1016/j.engappai.2025.110924

Singh, A., Ashuli, A., Dhiman, N., Dubey, C. S., & Shukla, D. P. (2024). Evaluating causative factors for landslide susceptibility along the Imphal-Jiribam railway corridor in the North-Eastern part of India using a GIS-based statistical approach. Environmental Science and Pollution Research, 31(41), 53767-53784. https://doi.org/10.1007/s11356-023-28966-z

Steger, S., Brenning, A., Bell, R., Petschko, H., & Glade, T. (2016). Exploring discrepancies between quantitative validation results and the geomorphic plausibility of statistical landslide susceptibility maps. Geomorphology, 262, 8-23. https://doi.org/10.1016/j.geomorph.2016.03.015

Steger, S., Scorpio, V., Comiti, F., & Cavalli, M. (2022a). Data-driven modelling of joint debris flow release susceptibility and connectivity. Earth Surface Processes and Landforms, 47(11), 2740-2764. https://doi.org/10.1002/esp.5421

Steger, S., Moreno, M., Crespi, A., Zellner, P. J., Gariano, S. L., Brunetti, M. T., Melillo, M., Peruccacci, S., Marra, F., Kohrs, R., Goetz, J., Mair, V., & Pittore, M. (2022b). Deciphering seasonal effects of triggering and preparatory precipitation for improved shallow landslide prediction using generalised additive mixed models. Natural Hazards and Earth System Sciences Discussions, 2022, 1-38. https://doi.org/10.5194/nhess-23-1483-2023

Steger, S., Moreno, M., Crespi, A., Gariano, S. L., Brunetti, M. T., Melillo, M., Peruccacci, S., De Vugt, L., Zieher, T., Rutzinger, M., Mair, V., & Pittore, M. (2024). Adopting the margin of stability for space–time landslide prediction–A data-driven approach for generating spatial dynamic thresholds. Geoscience Frontiers, 15(5), 101822. https://doi.org/10.1016/j.gsf.2024.101822

Sujatha, E. R. (2021). An integrated landslide susceptibility model to assess landslides along linear infrastructure for environmental management. Environmental Earth Sciences, 80(12), 447. https://doi.org/10.1007/s12665-021-09747-8

Sundararajan, M., Taly, A., & Yan, Q. (2017). Axiomatic attribution for deep networks. In D. Precup & Y. W. Teh (Eds.), Proceedings of the 34th International Conference on Machine Learning (Vol. 70, pp. 3319–3328).

Sur, U., Singh, P., & Meena, S. R. (2020). Landslide susceptibility assessment in a lesser Himalayan road corridor (India) applying fuzzy AHP technique and earth-observation data. Geomatics, Natural Hazards and Risk, 11(1), 2176–2209. https://doi.org/10.1080/19475705.2020.1836038

Tien Bui, D., Tuan, T. A., Klempe, H., Pradhan, B., & Revhaug, I. (2016). Spatial prediction models for shallow landslide hazards: a comparative assessment of the efficacy of support vector machines, artificial neural networks, kernel logistic regression, and logistic model tree. Landslides, 13, 361-378. https://doi.org/10.1007/s10346-015-0557-6

Trigila, A., Iadanza, C., & Spizzichino, D. (2010). Quality assessment of the Italian Landslide Inventory using GIS processing. Landslides, 7, 455–470. https://doi.org/10.1007/s10346-010-0213-0

Tsangaratos, P., Ilia, I., Hong, H., Chen, W., & Xu, C. (2017). Applying Information Theory and GIS-based quantitative methods to produce landslide susceptibility maps in Nancheng County, China. Landslides, 14, 1091-1111. https://doi.org/10.1007/s10346-016-0769-4

Van Westen, C. J., Rengers, N., & Soeters, R. (2003). Use of geomorphological information in indirect landslide susceptibility assessment. Natural hazards, 30, 399-419. https://doi.org/10.1023/B:NHAZ.0000007097.42735.9e

Van Westen, C. J., Van Asch, T. W., & Soeters, R. (2006). Landslide hazard and risk zonation-why is it still so difficult?. Bulletin of Engineering geology and the Environment, 65, 167-184. https://doi.org/10.1007/s10064-005-0023-0

Varnes, D. J. (1984). Landslide hazard zonation: a review of principles and practice (No. 3).

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., & Polosukhin, I. (2017). Attention is all you need. Advances in neural information processing systems, 30.

Wang, S., Zhuang, J., Mu, J., Zheng, J., Zhan, J., Wang, J., & Fu, Y. (2022). Evaluation of landslide susceptibility of the Ya'an–Linzhi section of the Sichuan–Tibet Railway based on deep learning. Environmental Earth Sciences, 81(9), 250. https://doi.org/10.1007/s12665-022-10375-z

Wang, Y., Wu, X., Zhou, K., Lin, G., Peng, B., & Fang, Z. (2025). Integrating a multi-dimensional deep convolutional neural network with optimized sample selection for landslide susceptibility assessment. Geo-spatial Information Science, 1-21. https://doi.org/10.1080/10095020.2024.2443483

Wood, S. N. (2017). Generalized additive models: an introduction with R. chapman and hall/CRC.

Woodard, J. B., Mirus, B. B., Crawford, M. M., Or, D., Leshchinsky, B. A., Allstadt, K. E., & Wood, N. J. (2023). Mapping landslide susceptibility over large regions with limited

data. Journal of Geophysical Research: Earth Surface, 128(5), e2022JF006810. https://doi.org/10.1029/2022JF006810

Yano, A., Shinohara, Y., Tsunetaka, H., Mizuno, H., & Kubota, T. (2019). Distribution of landslides caused by heavy rainfall events and an earthquake in northern Aso Volcano, Japan from 1955 to 2016. Geomorphology, 327, 533-541. https://doi.org/10.1016/j.geomorph.2018.11.024

Yu, L., Wang, Y., & Pradhan, B. (2024). Enhancing landslide susceptibility mapping incorporating landslide typology via stacking ensemble machine learning in Three Gorges Reservoir, China. Geoscience Frontiers, 15(4), 101802. https://doi.org/10.1016/j.gsf.2024.101802

Zêzere, J., Neves, M., Henriques, C. S., Garcia, R., Oliveira, S., & Piedade, A. (2009). Effects of landslide inventories uncertainty on landslide susceptibility modelling. Landslide processes: from geomorphologic mapping to dynamic modelling, 81-86.

Zhang, B., Tang, J., Huan, Y., Song, L., Shah, S. Y. A., & Wang, L. (2024). Multi-scale convolutional neural networks (CNNs) for landslide inventory mapping from remote sensing imagery and landslide susceptibility mapping (LSM). Geomatics, Natural Hazards and Risk, 15(1), 2383309. https://doi.org/10.1080/19475705.2024.2383309

Zhao, Z., Chen, T., Dou, J., Liu, G., & Plaza, A. (2024). Landslide susceptibility mapping considering landslide local-global features based on CNN and transformer. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. https://doi.org/10.1109/JSTARS.2024.3379350

Zhou, C., Cao, Y., Gan, L., Wang, Y., Motagh, M., Roessner, S., ... & Yin, K. (2024). A novel framework for landslide displacement prediction using MT-InSAR and machine learning techniques. Engineering Geology, 334, 107497. https://doi.org/10.1016/j.enggeo.2024.107497

Zhou, M., Yuan, M., Yang, G., & Mei, G. (2024). Risk analysis of road networks under the influence of landslides by considering landslide susceptibility and road vulnerability: A case study. Natural Hazards Research, 4(3), 387-400. https://doi.org/10.1016/j.nhres.2023.09.013

2. Chapter 2: Landslide susceptibility methodology for railway planning - a comparative analysis of statistical and machine learning methods in a case study of Marche region, Italy.

2.1 Preface

This chapter addresses the practical challenges of landslide susceptibility in a real railway design scenario. The presented work represents a key step in my PhD research, arising from the natural issues and uncertainties encountered during the real-world application of landslide susceptibility modelling, where the treatment of different landslide types, the reliability of the landslide inventory, and the reclassification problem emerge as critical points alongside the modelling process itself.

The focus is on developing an efficient methodology to estimate landslide susceptibility for the railway line planned to connect Fabriano and Castelpliano in the Marche region. The methodology involves separate analyses for each mapped landslide type in the study area, acknowledging that different types are driven by distinct processes and require specific mitigation strategies. An overall susceptibility map is then produced through an ensemble of reclassification methods, yielding a comprehensive, reliable, and easily interpretable map for decision-makers, who are often non-technical.

The chapter also presents a comparison of different approaches found in the literature, including a statistical method (WoE) and a machine-learning-based method (GAMBoost), highlighting their respective strengths and evaluating their performance along the proposed railway track.

This research underlines the necessity of a standardized methodology for landslide susceptibility modelling in railway planning. The methodology and results presented in this chapter provide a first step toward a more unified and reliable approach, which is further discussed in the concluding Chapter 6 of the thesis.

The chapter is based on the article published in Transportation Geotechnics journal:

Rani R.[1*], Sciarra M.[2], Rodani S.[2], Benedetti G.[2], Berti M.[1], 2025. Landslide susceptibility methodology for railway planning: A comparative analysis of statistical and machine learning methods in a case study of Marche region, Italy. https://doi.org/10.1016/j.trgeo.2025.101731

[1]Department of Biological, Geological, and Environmental Sciences (BiGea), University of Bologna, Via Zamboni 67, 40126, Bologna, Italy

[2]ITALFERR S.p.A., Gruppo Ferrovie dello Stato Italiane - Architecture, Environment & Territory Department - Geology Division, Via Galati 87, 00155, Roma Italy

2.2 Abstract

Landslides pose serious risks to infrastructure, particularly railways, due to their rigid construction and essential transport role. Susceptibility mapping is a valuable tool during the feasibility phase of railway projects, helping identify high-risk areas and inform mitigation strategies. However, effective application requires both reliable classification of landslide types and robust reclassification methods for clear communication with stakeholders. This study presents a comprehensive workflow for landslide susceptibility mapping, combining Weight of Evidence (WoE) and a Generalized Additive Model with boosting (GAMB). We generated separate susceptibility maps for five landslide types and evaluated them using AUROC metrics. The maps were merged into an overall susceptibility map using a complementary probability approach, which also allowed assessment of each type's sensitivity to the overall susceptibility. To improve threshold reliability, we implemented an ensemble reclassification method using six approaches and applied the statistical mode to define more objective class boundaries. Visualizations of susceptibility along the railway route and its adjacent sides were developed for practical application. The methodology was applied to a 22 km planned railway section in the Marche region (Italy). Results revealed high spatial variability: rockfall types showed the highest accuracy (AUC = 0.94 WoE, 0.98 GAMB), while slides performed poorest. GAMB consistently outperformed WoE in reliability and smoothness of results. Finally, a comparison with EGMS ground motion data showed no significant correlation ($R^2 \approx 0.1$), underscoring the temporal disconnect between long-term susceptibility and short-term ground deformation.

2.3 Introduction

Landslides rank among the most devastating natural hazards, capable of dramatically reshaping landscapes, destroying man-made structures, and endangering lives (Aleotti & Chowdhury 1999; Froude & Petley 2018). Identifying areas prone to landslides is crucial for human safety, environmental management, urban planning, and mitigating economic losses (Kavzoglu et al. 2014). A significant aspect of this identification process involves understanding the unique vulnerabilities of different infrastructures to landslide impacts, particularly those of railway systems (Liu et al. 2018; Marchesini et al. 2024).

Railways, unlike individual buildings, cover extensive areas, making them more vulnerable to landslides. This risk is heightened in comparison to roads due to the larger curvature radius required for railways and the rigidity of both the construction materials and trains (Arslan et al. 2014). These characteristics limit railways' ability to adapt to terrain shifts, with even minor track deformations potentially causing derailments, posing serious safety risks and service interruptions (Cannon et al. 2003). The risk is further intensified by trains' reduced capacity to stop quickly, compared to road vehicles (Schafer et al. 2008). Additionally, expanding the railway network plays a critical role in mitigating climate change, as rail transport is more environmentally friendly (Chester & Horvath 2010). In this context, climate change poses a growing threat to the structural integrity of rail networks (Schlögl & Matulla 2018).

Given these challenges, incorporating landslide susceptibility (LS) analysis during the feasibility phase of rail route design is crucial to prevent and mitigate disasters. Several studies have explored this application, using methods such as Weight of Evidence (WoE) (Quinn et al.

2010), bivariate statistical analysis (Bednarik et al. 2010), and a combination of information value, WoE, and logistic regression (Pasang & Kubíček 2020). Other approaches, like the Analytical Network Process (Sujatha 2021), the Fuzzy Analytic Hierarchy Process (Sur et al. 2020), and Random Forest and Support Vector Machine (Li et al. 2025) have also been applied for landslide analysis along railway lines. However, the limited number of practical studies in railway design emphasizes the need to integrate academic findings into real-world applications.

Two critical issues are often overlooked in the literature: (i) the complex relationship between the territory and railway infrastructure, and (ii) the challenging task of determining reliable landslide value thresholds for safe and unsafe locations by planners (Guzzetti et al. 1999). Regarding the first issue, extensive railway projects often encounter different geological domains, characterized by high landscape heterogeneity and different landslide types. Although some studies have started to address the role of different landslide types (Guo et al. 2025; Yu et al. 2024, Loche et al. 2022, Dou et al. 2020), this aspect remains largely underexplored, particularly in the context of long-term, non-event-based inventories. In these cases, the conventional approach of treating all landslide types uniformly proves inadequate given the diverse geological and geomorphological settings (Guzzetti et al. 1999). Moreover, each landslide type necessitates a unique mitigation strategy, and valuable information is lost when landslides are uniformly treated. Focusing limited landslide management resources on the most susceptible locations, which vary depending on the landslide type, is crucial (Zêzere 2002; Taalab et al. 2018). Additionally, the sampling strategy used to represent landslides (e.g., focusing on the triggering area, the entire polygon, or the centroid) can also influence modelling results, as recently examined by Guo et al. (2024), Lima et al. (2023) and Steger et al. (2016).

Furthermore, the subdivision of landslide susceptibility in hazard classes is often ambiguous, which can render the analysis ineffective (Reichenbach et al. 2018). While susceptibility models often present probabilities, practical applications demand straightforward categorizations, such as low, moderate, or high susceptibility. However, defining these categories is complex and can lead to varying interpretations and applications, with only a few authors raised this concern (Chung & Fabbri 2003, Ayalew & Yamagishi 2005; Alvioli et al. 2024; Steger et al. 2024). Cantarino et al. (2019) conducted a critical analysis of the results of classification systems, using several statistical methods. However, in the case of similar statistical performance, they also introduced interpretation strategies to maintain a balanced approach to classification.

In this study, we explore how landslide susceptibility methods can be effectively integrated into railway design, focusing on a test area located in the Marche region of Italy. More specifically, we compare two susceptibility models: the classical statistical-based Weight of Evidence (WoE; Bonham-Carter 1994) and the more advanced Generalized Additive Model with gradient boost (GAMB; Hastie & Tibshirani 1990, Lou et al. 2013). The models are first applied separately to five landslide types, using distinct target datasets for each type, obtaining five distinct susceptibility maps. Then we combine the single maps to derive an overall landslide susceptibility map which represents the likelihood of any landslide occurrence. The landslide probability values are finally classified using a new approach based on the application of six

different techniques in order to obtain a reliable representation of hazard classes. The results of the susceptibility models are finally compared with the ground movements provided by the European Ground Motion Service (EGMS), to investigate the potential correlation between actual ground movements and predicted landslide susceptibility.

2.4 Study area

The study area is located in the Marche region of central Italy (Figure 2.1). The area includes a significant portion of the Esino and Sentino Rivers basin, extending from Castelplanio (province of Ancona), which marks the downstream terminus of the rail project, and proceeding upstream up to the city of Fabriano (same province). Encompassing approximately 450 km², the study area was strategically selected with a wider boundary to ensure the acquisition of a comprehensive and diverse dataset for landslide inventory. Rather than focusing solely on areas adjacent to the railway, this broader area provides a more complete understanding of the types and distribution of landslides, which is crucial for the robustness of our data-driven analyses. The boundaries were defined by including the Esino River catchment up to an arbitrary point in the valley beyond the village of Castelplanio, thereby incorporating the entire railway within the natural limits of the basin.

2.4.1 Geological setting

The study area is characterized by two distinct geological domains (A and B), as illustrated in Figure 2.1. To the west, Domain A consists of older, hard rocks, which encompasses the "Marchean Ridge" primarily composed of massive and micritic limestone (blue colours in Figure 2.1), and the "Inner Marchean Basin" to the west, which features Oligocene to Miocene marly sediments (green colours in Figure 2.1), with occasional gypsum and clayey deposits in its innermost parts (Ori et al. 1991). In contrast, Domain B to the east is composed of younger, weaker rocks (red-grey colours in Figure 2.1), including clays interbedded with fine sandstone and limestone, alternating marly-arenaceous and pelitic layers (with local variations in ratio) originated from turbidite phenomena, coarser sandstone lithotypes, and, lastly, Messinian gypsum and marls (Cardellini et al. 2009). Domain B forms part of the peri-Adriatic Basin, characterized by arenaceous and pelitic sediments from the Lower-Middle Pliocene to the Lower Pleistocene (Centamore et al. 1979).

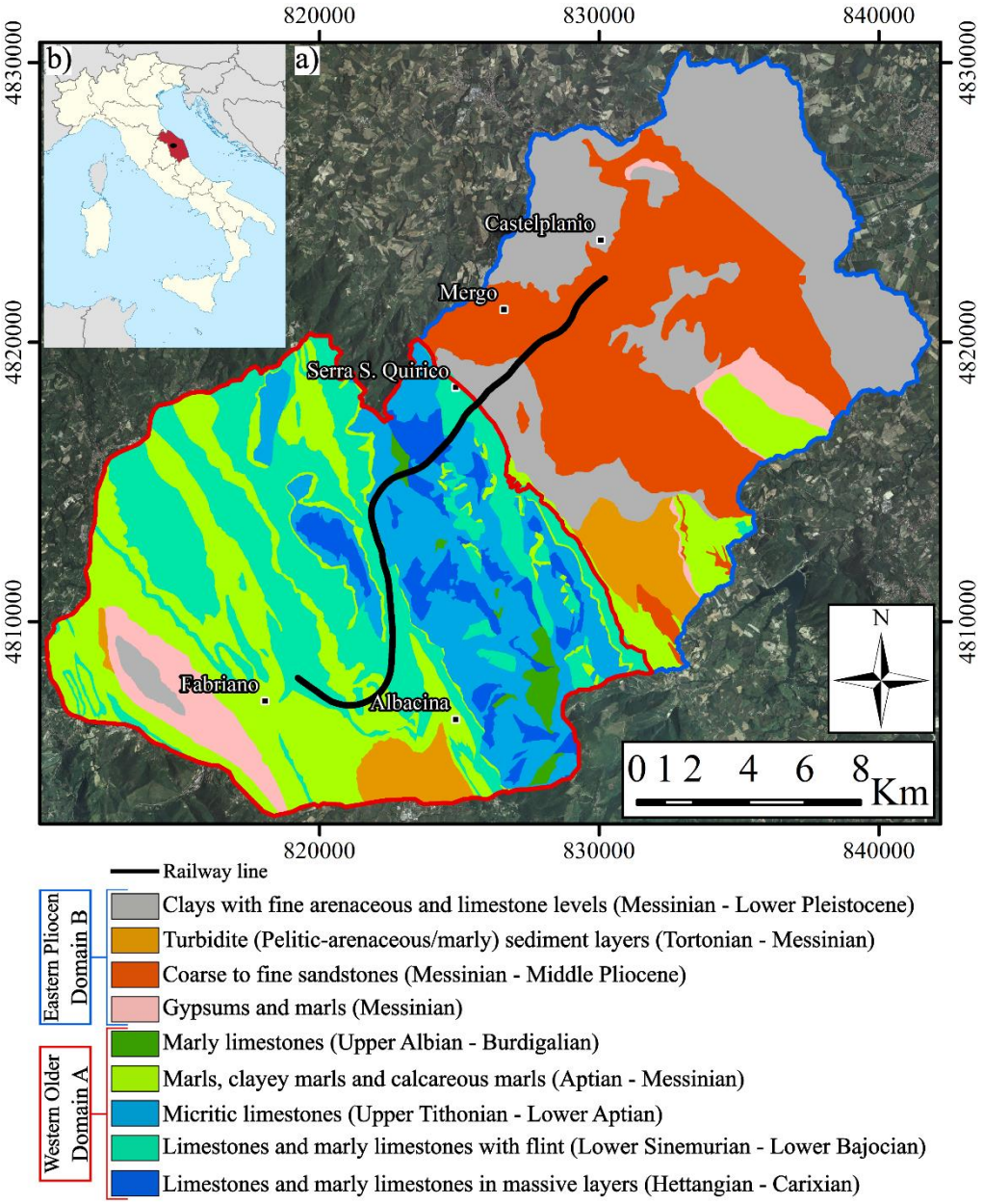


Figure 2.1. a) Geology map of the study area subdivided into domains A and B, with railway line and localities. b) Study area location (black dot) in the Marche region (red area) in Italy.

This diverse geological-structural setting results in a wide range of landforms. The Domain A reaches high elevations (around 1000 m a.s.l.) with steep slopes and sharp escarpments, particularly in areas dominated by massive limestone. In contrast, the domain B exhibits gentler slopes and lower elevations (300-500 m a.s.l.). The transition between these two geomorphological zones is abrupt, contributing to a varied landscape. These geological and geomorphological constraints significantly influence the distribution and types of landslides across the area (Guzzetti et al., 1996). Steeper slopes in hard rock areas are more prone to falls and topples (domain A), while complex and slow-moving landslides are more prevalent on the clay-rich substrates in the east (domain B). These differences in landslide types pose challenges for accurately assessing landslide susceptibility.

2.4.2 The railway line

The study was conducted in the context of a planned upgrade for approximately 21.5 km of the historic railway line between Fabriano and Castelplanio (Figure 2.1). The proposed project involves a new line that will run parallel to the old, discontinued track. The new railway line will serve as the main link between the Adriatic coast of the Marche region and the Apennine hinterland, where there are important industrial centers such as Fabriano. In certain sections, the project will include the construction of new natural tunnels and engineering structures, such as viaducts and artificial tunnels. The planned alignment could accommodate varying speeds, ranging from 70 km/h to a maximum of about 180 km/h, with a maximum longitudinal gradient of around 12‰. From an elevation perspective, the project design aims to minimise interference with existing road infrastructure, ensure hydraulic compatibility, and meet the necessary conditions for tunnel construction and associated safety measures.

2.5 Material and methods

The workflow followed to evaluate landslide susceptibility along the rail line is depicted in Figure 2.2 and described in detail in the following sections.

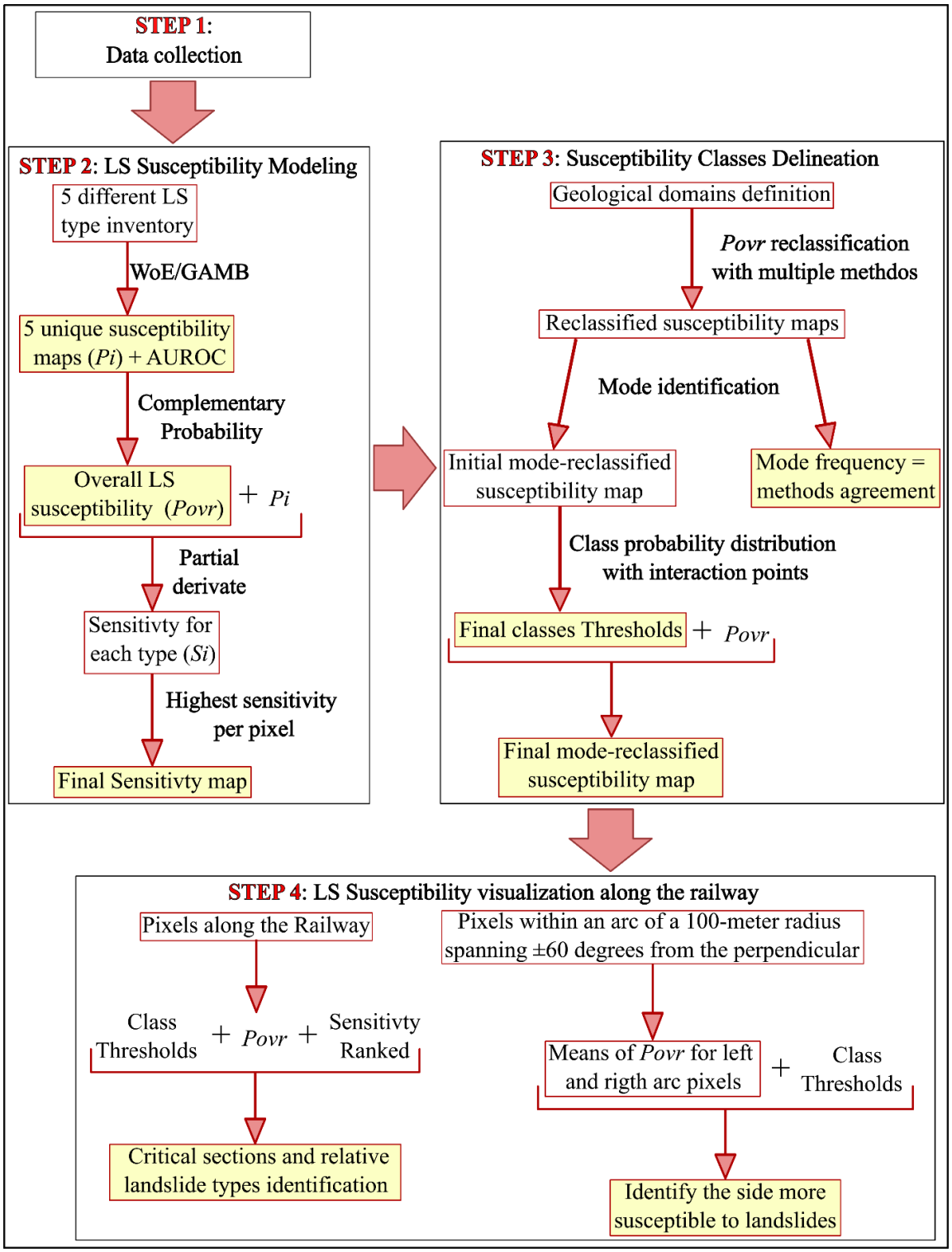


Figure 2.2. The full methodology workflow splitted into 4 steps.

2.5.1 Data collection (STEP 1)

The primary dataset for any susceptibility analysis is the landslide inventory map. In feasibility projects, this map is typically sourced from existing maps and eventually updated to include recent landslides. Our initial source was the IFFI (Italian Landslide Phenomena Inventory), a national database compiled by the National Higher Institute for Environmental Protection and Research (ISPRA; Trigila et al. 2010). The IFFI Project provides a comprehensive survey of

landslide occurrences across Italy, making it a crucial resource for assessing landslide hazards at the regional scale.

The IFFI Project provides a comprehensive survey of landslide occurrences across Italy, representing a key resource for assessing landslide hazards at the regional scale. Conducted primarily during the first decade of the 2000s, the IFFI inventory is updated annually to include major national events. The initial cartographic catalogue was created by integrating multiple data sources:

- Photointerpretation: a fundamental tool for systematic geomorphological surveys over large areas, enabling the identification and delineation of major landslide phenomena based on isohypses patterns.
- Historical and archival data analysis: used to reconstruct past landslide events and assess their evolution, recurrence, and intensity, although such detailed information is often limited.
- Field surveys: conducted to verify and supplement data acquired through photointerpretation and archival research.

Each landslide record in the IFFI inventory contains basic information, including location, type of movement, and activity status. In some cases, additional details are provided, such as morphometry, geology, lithology, land use, causes, and dates of past activations. Rarely, optional information on damages, investigations, or remediation measures is also included.

In our study area (Figure 2.3), the dataset follows this general pattern. Only two Fall/Topple events (from 1997) have detailed information, while the remaining landslide polygons were mapped using photointerpretation, field surveys, and local archives. The mapping scale was generally 1:10,000, with higher detail (1:5,000) near urbanized areas, providing an overall resolution suitable for regional and national-scale landslide studies.

As a result, the IFFI project includes a wide range of landslides formed under various environmental and climatic conditions, spanning an indefinite timespan. The extensive nature of the IFFI inventory means that using it as a target dataset for landslide susceptibility modelling could result in conservative or overly cautious susceptibility maps, as it encompasses all hypothetical landslide scenarios across a broad timescale. Given that the railway in our study is expected to operate for about 100 years, adopting a long-term conservative approach based on such a dataset is justified.

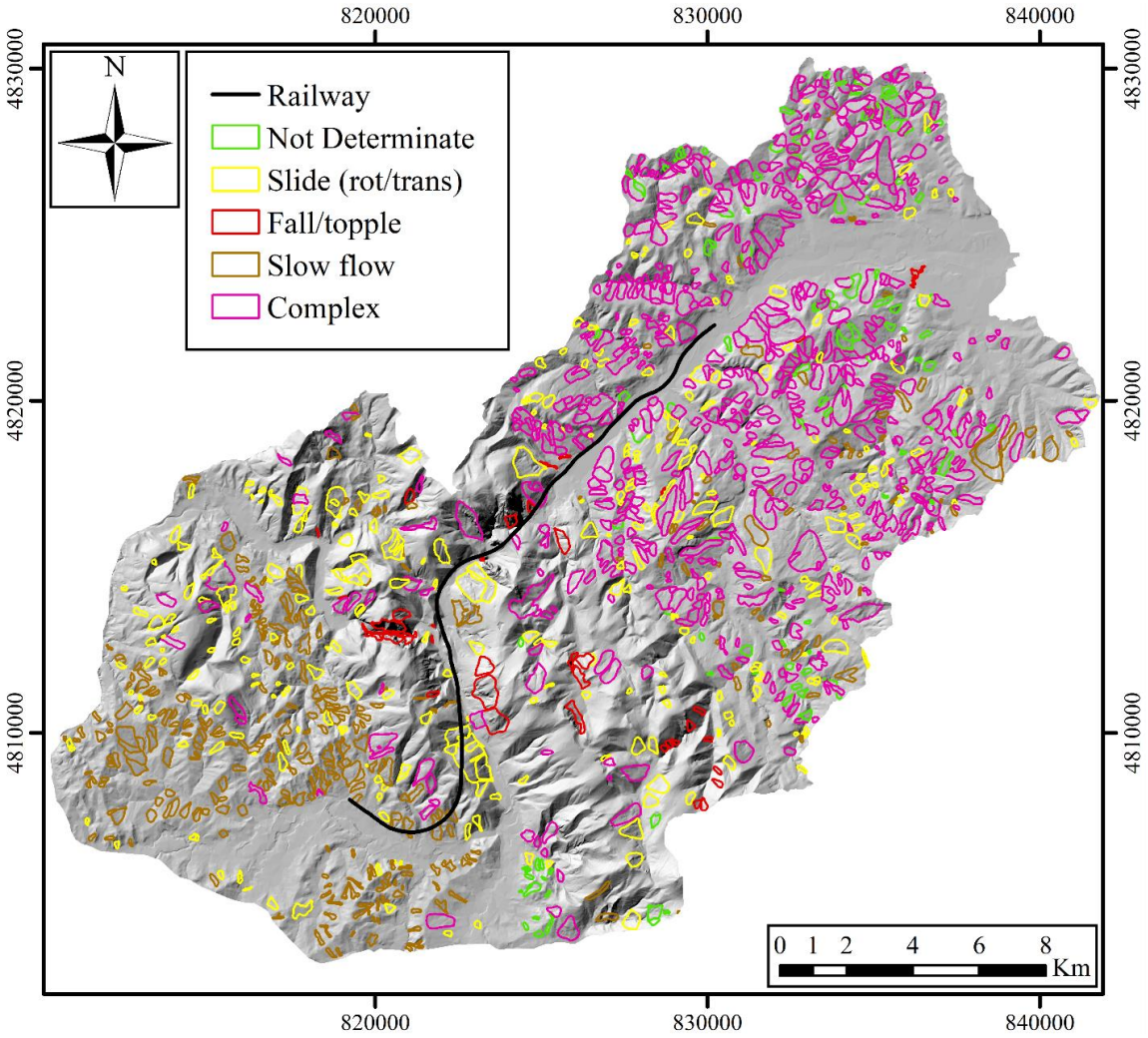


Figure 2.3. Landslide inventory from IFFI with hillshade effect.

For the classification and nomenclature of landslides, the IFFI project mainly based on Cruden and Varnes (1996), for detail see Trigila et al. (2010). In our study area, five landslide types are present: Slide (S), Fall/Topple (FT), Slow Flow (SF), Complex (C), and Not Determinate (ND). Each type has distinct spatial and geomorphological characteristics.

Slide (S): Slides involve soil or rock and can be either rotational or translational with no expressed distinction in the IFFI inventory. This type of movement is characterized by a distinct zone of weakness that separates the slide material from more stable underlying material. Slide type is characterized by variety of velocity range, but generally around from slow to moderate and rarely to rapid. In our study area, there are 278 mapped slides, covering 18.15 km² (4% of the area) with an average area of 0.066 km². Slides are distributed across different lithologies (percentage of landslide area in each lithology): Limestones and marly limestones with flint (28%), marls, clayey marls, and calcareous marls (23%), sandstones (19%), and micritic limestones (14%), with the remaining lithologies each below 10%. Slide type mean slope in our area is 19° with a standard deviation (St. Dev.) of 9°. Slides are widely dispersed across both geological domains, occurring from gentle to steep scarps, without evident clustering (Figure 2.3).

Fall/Topple (FT): Falls are rapid movements of rock masses that detach from steep slopes or cliffs, while topples involve forward rotation of rock units about a pivot point. FT landslides in our dataset are 34, covering 3.58 km² (0.8% of the area) with an average area of 0.1 km². They predominantly occur in Limestones and marly limestones in massive layers (63%) and micritic limestones (25%). These landslides are associated with the steepest slopes in the study area, with a mean slope of 37° (St. Dev. = 13°), reflecting their occurrence on steep scarps. Fall/Topple movements are rapid to extremely rapid, consistent with their physical nature.

Slow Flow (SF): Slow Flows, corresponding to Earthflows in Cruden and Varnes (1996), have an elongated, "hourglass" shape, often forming a depression or bowl at the head. SF landslides in our study area are 267, covering 15.22 km² (3.4% of the area), with an average area of 0.057 km². They are mainly located on the western side of domain A in marls, clayey marls, and calcareous marls (39%), with additional occurrences in the clay-rich portions of domain B. The mean slope for SF landslides is 13° (St. Dev. = 6°), indicating a preference for gentle to moderate slopes, with consistently slow movement velocities.

Complex (C): Complex type is described a combination in time and/or space of two or more principal types of movement. In our area, the complex type present usually are characterized by a gentle slope with multi-cambered and concave shapes portions, which identified the scarps or feet of different slip surfaces in space and time now no longer discernable between each other. Usually, in the study area, they typically combine Slide and Slow Flow movements, with a slow to very slow velocity, but huge volume. Complex landslides are especially common in domain B (in both Sandstone (39%) and clays (35%) lithology), where they cover a significant portion of the landscape. In total, there are 436 Complex landslides, covering 53.37 km² (11.8% of the area), with an average area of 0.122 km², being the larger landslide type of the study area.

Not Determinate (ND): This category includes landslides that lack clear diagnostic features for specific classification, perhaps due to aging processes or human activities. ND landslides are often similar to Complex landslides but lack definitive indicators of multiple material types or movement patterns, which can lead the mapper to attribute Complex as the movement type. In our study area, there are 115 ND landslides, covering a total area of 5.15 km² (1.15% of the study area), with an average landslide area of 0.045 km². They are mainly present in domain B in Sandstone (38%), clay (33%) and Pelitic-arenaceous and arenaceous-marly alternations (11%).

Large landslide inventories such as the IFFI are inherently inhomogeneous, often compiled by various mappers who bring different levels of expertise and judgment. Inconsistencies typically arise in the classification of landslide types and can also extend to variations in mapping accuracy. These limitations need to be taken into account during both the training and validation phases of analysis.

Eight predisposing factors (Figure 2.4) were considered based on their geological and geomorphological relevance, literature guidance, and the available data. We included geology, land cover, slope, and aspect as the basic and most significant factors (Budimir et al., 2015) because they represent the primary controls on landslide occurrence. Given that our inventory is mainly derived from photointerpretation, therefore sensitive to landform, we also

incorporated plan (horizontal) and profile (vertical) curvature to distinguish concave and convex portions of slopes, and we added the bedding plane–slope orientation relationship, which is mechanistically linked to translational slides on cataclinal slopes. The Topographic Wetness Index (TWI) was used as an integrated hydrological predictor because it combines slope and flow accumulation, so adding distance to streams or additional flow metrics would be largely redundant.

We intentionally excluded several predictors aiming to a baseline (long-term) susceptibility for a long-life infrastructure. Topographic Position Index (TPI) and related roughness measures were not used because their signal overlaps with that already captured by plan/profile curvature; distance to faults (and, analogously, seismic PGA proxies) was omitted due to scale mismatch and mapping uncertainty, with structural predisposition better represented by bedding–slope orientation and geology; NDVI and other spectral indices are temporally unstable (seasonal depending) and largely reflect land cover (already included); and soil maps/soil thickness were not used because of heterogeneous scale/quality and limited independent information beyond geology. We did not include distance to roads because our inventory chiefly represents long-term, natural landslides, and our goal is to characterize the background susceptibility of the terrain prior to any railway construction; proximity to existing roads would introduce anthropogenic bias not representative of natural predisposition. Land cover remains valuable, as it encodes the prevailing hillslope cover (including human presence) without forcing proximity effects.

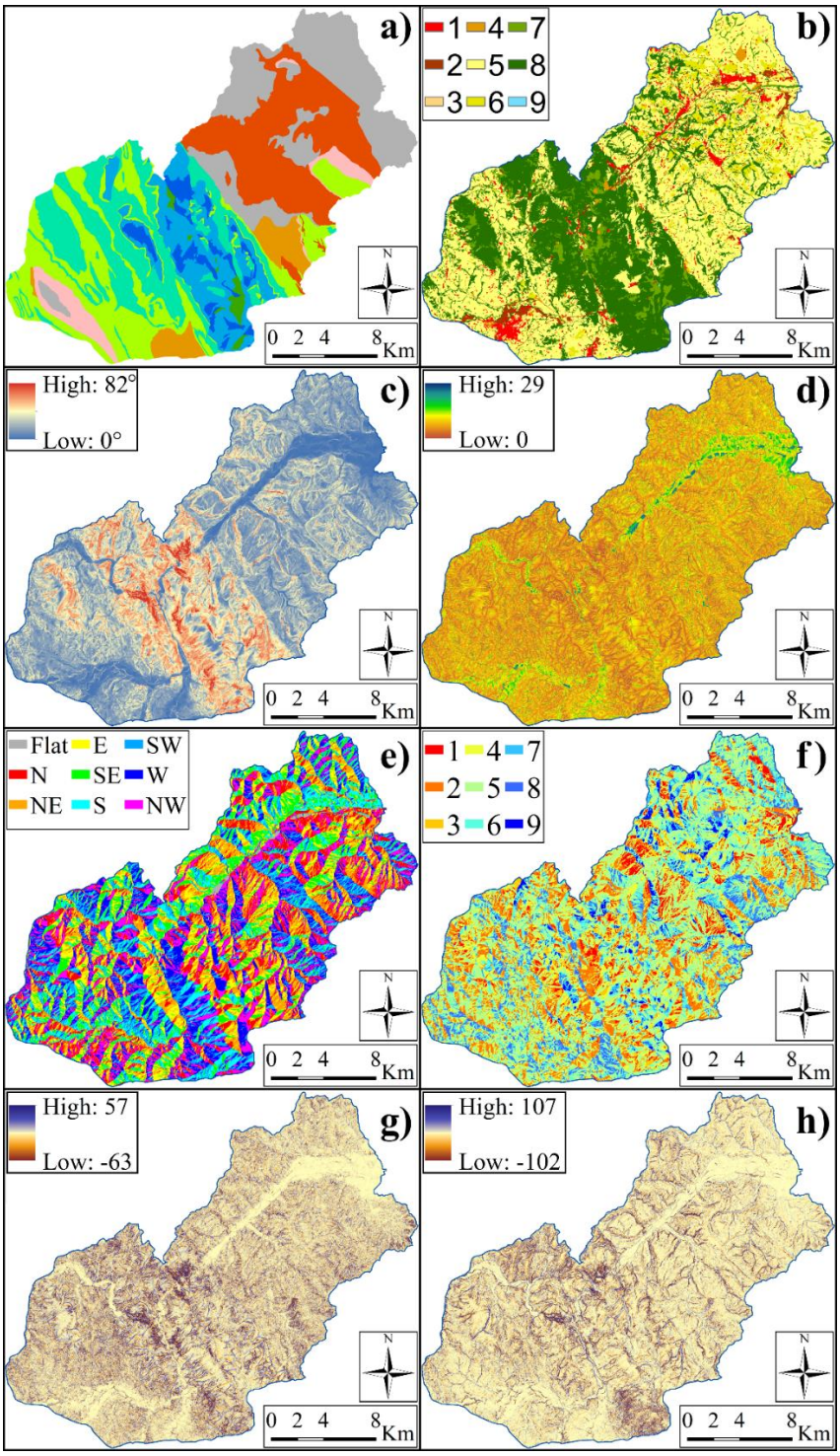


Figure 2.4. Used geomorphology layers as predictors. a) Geology: see legend Fig. 1. b) Land cover: Urbanised residential area (1), Urbanised industrial area (2), Urban green area (3), Excavating area (4), Heterogeneous arable agricultural land (5), permanent arable land (tall trees) (6), Meadows with sparse vegetation (7), Mixed wooded territories (8), Water bodies (9). c) Slope. d) Topographic Wetness Index (TWI). e) Aspect. f) Bedding planes-slope orientation relationships: Cataclinal horizontal (1), Cataclinal under-dip (2), Cataclinal over-dip (3), Oblique cataclinal (4), Orthoclinal (5), Oblique anaclinal (6), Anaclinal horizontal (7), Anaclinal under-dip (8), Anaclinal over-dip (9). g) Planar curvature. h) Profile curvature.

Bedrock geology was obtained from the regional geological map at a 1:10,000 scale, available on the regional website (https://www.regione.marche.it/Regione-Utile/Paesaggio-Territorio-Urbanistica/Cartografia/Repertorio/Cartageologicaregionale10000). From this fine resolution, regional, digitalized map, bedding plane measurement were also extracted. Land cover data were sourced from the national geoportal (https://geodati.gov.it/geoportale/) at the same 1:10,000 scale. This dataset was generated through automated classification of multispectral images (RGB and false colors) acquired in June-July 2007 with a Leica ADS40, orthorectified with 1-meter pixels (in accordance with 1:10000 scale). The minimum mappable unit is 3000 square meters, with 80% accuracy, and the legend originally adopted is the CORINE second level. The Digital Elevation Model (DEM) was obtained from the TINITALY project (https://tinitaly.pi.ingv.it/), originally published in 2007 and recently updated in 2023. The DEM, with a resolution of 10x10 meters, was constructed using contour lines and spot heights derived from Italian regional topographic maps, satellite-based global positioning system points, ground-based altimetry, and radar altimetry data (Tarquini et al. 2007). The 10-m resolution TINITALY is compared by Tarquini et al. (2012) with two, coarser-resolution, global-coverage, spaceborne-based Shuttle Radar Topographic Mission (SRTM) and Advanced Spaceborne Therma Emission ad Reflection Radiometer (ASTER) DEMs and with a high resolution LIDAR-derived DEM, showing an root mean square error in elevation (RMSEz) of 1.98 m against the latter, while the SRTM and ASTER DEMs against the LIDAR show a RMSEz of 7.78 and 9.36 m respectively.

Using this DEM, slope and aspect were computed in ArcGIS 10.8 (Slope, Aspect) with the planar method and default settings: a 3×3 neighbourhood and the Horn (1981) finite-difference algorithm; aspect is reported 0–360° (−1 for flats). Curvature was derived with the Curvature tool using the default 3×3 local surface fit, extracting both plan (planar) and profile curvature (positive = convex, negative = concave). The Topographic Wetness Index (TWI) was produced with the Hydrology tools using default D8 flow direction and Flow Accumulation (D8), combined with the slope grid per Beven & Kirkby (1979) (no custom parameters applied).

A special in-depth description is given of the Bedding planes-slope orientation relationship. The relationship between bedding planes and the orientation of a slope plays a crucial role in influencing landslide phenomena, especially for translational slide type. However, evaluating this relationship across a large study area is challenging (De Kemp 1998; Meentemeyer and Moody 2000; Gunther 2003; Grelle 2011; Santangelo et al. 2015). In particular, assuming spatial continuity of geologic bedding via interpolation of point measurements is prone to errors, especially in geo-structurally complex areas with numerous folds and faults. Sparse measurement points may fail to capture these complexities thus providing unrealistic results. To mitigate this problem, we interpolated the data using the method proposed by Meentemeyer & Moody (2000), which involves decomposing directional data into x and y components through sine and cosine functions, compute the interpolation of both and then reconverted these values into degrees, allowing for direct comparisons with the aspect. Various techniques were tested to assess suitable interpolation method for spatial point measurements in GIS environment. These included Diffusion Kernel with Barriers (geological faults as barriers),

Kernel with Barriers, SPline, IDW without Barriers, Natural Neighbor without Barriers, and two variations of Kriging without Barriers (Childs 2004). The evaluation revealed issues such as excessive averaging (Diffusion Kernel with Barriers, Kernel with Barriers), NoData output pixels (Kernel with Barriers, SPline), and challenges in maintaining original data values during interpolation (Kriging without Barriers). IDW and Natural Neighbor without Barriers showed more reasonable results, achieving a balance between individual data point significance and spatial continuity. Ultimately, the choice fell on the Natural Neighbor without Barriers because the IDW produced a ring of values around the measurement points (360° direction) in an attempt to accommodate differences between two close measurement points with very different values, increasing interpretation and often producing unrealistic effects. Finally, by comparing strata attitudes with slope aspect and angles, we defined nine categories of geometrical relationships (see Tab. 1), following the classification proposed by Peng et al. (2014).

2.5.2 Landslide susceptibility modelling (STEP 2)

2.5.2.1 Susceptibility models

Landslide susceptibility was calculated using two established methods: the Weight of Evidence (WoE) (Bonham-Carter 1994) and the Gradient boosted Generalized Additive Model (GAMB) (Hastie & Tibshirani 1990, Lou et al. 2013).

These approaches analyze the relationship between a binary dependent variable, in our case the presence or absence of a landslide, and various categorical or continuous independent variables that might influence this event, like slope, geology, and land cover. WoE was selected for its simplicity, flexibility, and clear demonstration of how factors contribute to landslide susceptibility, which is crucial for effective communication with stakeholders. On the other hand, GAMB excels at capturing complex, non-linear relationships between factors and variables, managing continuous factors without needing discretization, and goes beyond traditional linear models by incorporating additive structures where the total effect is the sum of the effects of individual predictors (James et al. 2013).

We have selected grid-cell units (10x10 meters) for our landslide susceptibility analysis, recognizing their high spatial resolution and alignment with current field practices, as noted by Reichenbach et al. (2018). This choice effectively facilitates detailed assessments at the decameter scale, which is crucial for the precision required in railway infrastructure planning. While slope unit-based analysis has merits in geomorphological terms and validation results (Jacobs et al. 2020; Ba et al. 2018), it generally classifies larger areas uniformly regarding hazard potential. This uniform classification may obscure specific high-hazard areas within a slope, a concern raised by Ba et al. (2018) and Martinello et al. (2021). Such precision is particularly critical in railway safety studies where localized site conditions significantly influence risk assessments. Despite the higher computational demands, employing grid-cell units is needed to achieve a comprehensive yet detailed analysis.

Unlike methods that distinguish between the source (or triggering) area and the deposit, we opted to treat each landslide as a whole. This decision reflects our goal of modelling the probability of landslide occurrence in its entirety, ensuring a more comprehensive and cautious

assessment. Furthermore, Lima et al. (2023) demonstrated that this integrated approach can outperform source-only strategies when not relying on physically based simulations, which would otherwise significantly increase computational demands (Yildiz et al. 2023).

Weight of Evidence (WoE)

WoE calculates the likelihood of an event (e.g., landslide occurrence) for each category of an independent variable using Bayes' theorem (Bonham-Carter 1994). Implementing WoE, involves a series of steps that begin with calculating the logit of the priori landslide probability $P(L)$:

$$\text{Logit}(L) = \log_e\left(\frac{P(L)}{1-P(L)}\right) \quad [2.1]$$

where N(L) is the number of pixels with landslide presence (L), N(T) is the number of total pixels, and P(L)=N(L)/N(T).

This is followed by computing positive weights (W+), which indicate how the factor is positively associated with the occurrence of the event, and negative weights (W-), which indicate how the factor is negatively associated with the occurrence of the event. These weights are computed as the ratio of conditional probabilities with respect to each class of each predictor (B):

$$W^+ = \log\left(\frac{P(B|L)}{P(B|\bar{L})}\right) \quad [2.2]$$

$$W^- = \log\left(\frac{P(\bar{B}|L)}{P(\bar{B}|\bar{L})}\right) \quad [2.3]$$

where the $\bar{L}$ and $\bar{B}$ represent the absent of landslide and predictor class respectively.

The contrast (CS) is then evaluated as the difference between W^+ and W^-, to indicate each category's contribution to instability. Positive weights allow to update the a priori logit to obtain the logit of the posteriori probability, according to Meyer et al. (2014):

$$\text{Logit}\left(L|B_{i\ (i=1\text{ to }n)}^{j\ (j=1\text{ to }k)}\right) = \text{Logit}(L) + \sum_{i=1}^{n}\sum_{j=1}^{k} W_{ij}^{+} \quad [2.4]$$

where B_i^j is the j^{th} class of the predictor i.

This logit is finally converted back into a probability, yielding the weighted landslide probability:

$$P(L|B) = \frac{\exp(\text{Logit}(L|B))}{1+\exp(\text{Logit}(L|B))} \quad [2.5]$$

Table 2.1 shows the division into specific classes of the conditioning factors, necessary for the WoE method. Geology and Land cover are by their nature pre-categorised so there is no need to be subdivide them (see Figure 2.1 and 2.4).

Predictors	Subdivision method	WoE Class		
Bedding planes-slope orientation relationship	Arbitrary classification following Peng et al. (2014)		Aspect-dip direction (°)	Slope-dip (°)
		1 - Cataclinal Horizontal	0 - 30	(-5) - 5
		2 - Cataclinal Under-Dip	0 - 30	< -5
		3 - Cataclinal Over-Dip	0 - 30	> 5
		4 - Oblique Cataclinal	30 - 60	-
		5 - Orthoclinal	60 - 120	-
		6 - Oblique Anaclinal	120 - 150	-
		7 - Anaclinal Horizontal	150 - 180	(-5) - 5
		8 - Anaclinal Under-Dip	150 - 180	< -5
		9 - Anaclinal Over-Dip	150 - 180	> 5
Slope	Arbitrary classification		Slope values interval (°)	
		1 (almost flat)	0 - 5	
		2	5 - 10	
		3	10 - 12.5	
		4	12.5 - 15	
		5	15 - 17.5	
		6	17.5 - 20	
		7	20 -22.5	
		8	22.5 - 25	
		9	25 - 27.5	
		10	27.5 - 30	
		11	30 - 35	
		12	35 - 40	
		13 (really steep)	> 40	
Aspect	Arbitrary classification		Aspect direction interval (°)	
		1 - North (N)	0 - 22.5	
		2 - North East (NE)	22.5 - 67.5	
		3 - East (E)	67.5 - 112.5	
		4 - South East (SE)	112.5 - 157.5	
		5 - South (S)	157.5 - 202.5	
		6 - South West (SW)	202.5 - 247.5	
		7 - West (W)	247.5 - 292.5	
		8 - North West (W)	292.5 - 337.5	
		1 - North (N)	337.5 - 360	
Topographic Wetness Index (TWI)	Quantiles		TWI values interval	
		1 (more surface runoff)	0 - 0.27	
		2	0.27 - 4.12	
		3	4.12 - 4.68	
		4	4.68 - 5.14	
		5	5.14 - 5.59	
		6	5.59 - 6.04	
		7	6.04 - 6.72	
		8	6.72 - 7.85	
		9 (more waterlogging)	7.85 - 29.13	
Planar curvature	Standard deviation (1/2)		Curvature values interval	
		1 (concave surface, valleys and gullies)	-63.72 - (-2.65)	
		2	-2.65 - (-2.21)	
		3	-2.21 - (-1.77)	
		4	-1.77 - (-1.33)	
		5	-1.33 - (-0.89)	

		6	-0.89 - (-0.45)
		7 (almost plain, slightly concave)	-0.45 - 0
		8 (almost plain, slightly convex)	0 - 0.43
		9	0.43 - 0.87
		10	0.87 - 1.31
		11	1.31 - 1.75
		12	1.75 - 2.19
		13 (convex surface, hill and ridges)	2.19 - 57.89
Profile curvature	Standard deviation (1/2)		Curvature values interval
		1 (convex surface, upward camber)	-102.62 - (-3.49)
		2	-3.49 - (-2.92)
		3	-2.92 - (-2.34)
		4	-2.34 - (-1.76)
		5	-1.76 - (-1.19)
		6	-1.19 - (-0.61)
		7 (almost plain, slightly convex)	-0.61 - 0
		8 (almost plain, slightly concave)	0 - 0.54
		9	0.54 - 1.11
		10	1.11 - 1.69
		11	1.69 - 2.27
		12 (concave surface, less steepness and downward camber)	2.27 - 107.91

Table 2.1. Detailed classes subdivision of predictors with continues values, used in the WoE.

Generalized Additive Model with tree-based boosting (GAMB)

Generalized Additive Models (GAMs) are advanced statistical models that extend the Generalized Linear Model by incorporating non-linear relationships through smooth functions, typically spline functions, for each predictor (Hastie & Tibshirani 1990). Unlike the WoE, GAMs seamlessly accommodates continuous variables and maintains interpretability, enabling an understanding of the impact of each predictor and its non-linear components. Bayesian GAMs have been applied to model landslide susceptibility (Goetz, Guthrie & Brenning 2011; Petschko et al. 2014; Pourghasemi & Rossi 2017), and Poissonian GAMs have been used for modelling spatial landslide counts (Lombardo et al. 2020). In our study, we used the fitcgam function implemented in the MATLAB environment for binary classification, which extends the traditional GAM by incorporating gradient boosting (Lou et al., 2013). The model can be represented as:

$$y_i = \beta_0 + \sum_{j=1}^{p} f_j(x_j) + \sum_{i,j\in\{1,2,\ldots,p\}} f_{ij}(x_i x_j) \qquad [2.6]$$

Where: y_i represents the probability of success (landslide occurrence). β_0 is the intercept or baseline value. $\sum_{j=1}^{p} f_j(x_j)$ is the sum of univariate functions, each capturing the main effect of a single predictor x_j on the response. These functions are not smooth functions in the classical sense, but ensemble functions built from boosted decision trees. $\sum_{i,j\in\{1,2,\ldots,p\}} f_{ij}(x_i x_j)$ represents the sum of bivariate interaction functions, capturing pairwise interactions between predictors

$x_i x_j$. These interaction terms are also built using gradient-boosted decision trees, allowing the model to capture complex nonlinear dependencies.

The boosting process in fitcgam follows an adaptive gradient boosting algorithm, which initially builds sets of univariate (predictor) trees to approximate the marginal effects of individual predictors on the response. It then adds interaction trees to capture the combined effects of pairs of predictors. During each boosting iteration, the function fits a tree to the residuals (the difference between the current model prediction and the true response), scales this tree using a learning rate, and updates the overall model accordingly. This iterative approach allows the model to adaptively fit both main effects and interactions while minimizing overfitting, as the boosting process is carefully tuned using the OptimizeHyperparameters option in MATLAB, which varies key parameters such as learning rates, maximum number of trees, and tree complexity to minimize cross-validation loss.

2.5.2.2 Modeling workflow

To apply the susceptibility models, we developed an Excel-based structure that handles the original input file and works as an interface for the MATLAB Script where the analysis (WoE and GAMB) is done. This structure allows the MATLAB code to remain unchanged, while input variables can be easily modified by simply updating the Excel file, resulting in a flexible and user-friendly methodology, aim of our work.

We decided to use the complete dataset for computing the weights and training the models because, from a practical user-designer point of view, the model must be deliberately site-specific and achieve the best performance for that specific area (best success rate) at the sacrifice of replicability (prediction rate).

To face the common issue in literature of neglecting variations among different landslide types (Reichenbach et al. 2018), we decided to implement the models using different landslide types as distinct target datasets. The landslide type's classification was kept from the IFFI (see par. 3.1 for details): Not Determinate (ND), Slide (S), Fall and topple (FT), Slow flow (SF), and Complex (C). In the first stage, we therefore obtain a landslide susceptibility map for each type of landslide. These individual maps are then combined to create a comprehensive susceptibility map for all landslides in the area.

The overall landslide susceptibility map represents the probability of failure for any landslide type, together with the sensitivity of overall susceptibility to each landslide type. This map was computed using a complementary probability approach (Rényi 2007), which is particularly effective when the occurrences are mutually exclusive. The methodology involved the following steps:

I. The overall landslide probability (P_{ovr}) for each pixel is computed by subtracting the product of the complementary probabilities (1 minus the probability of each landslide type P_i) from 1. Essentially, it combines the individual probabilities of all considered landslide types (n) to estimate the likelihood of any landslide event happening.

$$P_{ovr} = 1 - \prod_{i=1}^{n} (1 - P_i) \quad [2.7]$$

II. The contribution of each landslide to the overall probability, that is the influence of an individual landslide type (i) on the P_{ovr}, is determined by computing the partial derivate of P_{ovr} over each probability (P_i). Each sensitivity raster (S_i) is then normalized to percentages ensuring that they sum up to 100.

$$S_i = \frac{\partial P_{ovr}}{\partial P_i} \quad [2.8]$$

where i indicate the different landslide types: 1=ND; 2=S; 3=FT; 4=SF; 5=C.

III. To enhance the understanding of how different landslide types contribute to overall susceptibility, we assign to each pixel the landslide types with the major sensitivity in that pixel.

These categorizations and thresholds were derived through a process of visual interpretation in ArcMap. Our aim was to ensure that areas with a high confidence level in landslide susceptibility fully encompass known landslide type areas, while uncertain class is assigned to regions with evident stability, such as valleys and other stable zones. This approach allowed us to tailor the classification scheme to reflect real-world conditions in our study area. However, it is important to note that the chosen thresholds may require adjustment when applied to different geographical contexts.

Finally, we employed the classical statistical measure, AUROC (Area Under the Receiver Operating Characteristics) (Hanley & McNeil, 1982), to assess the performance of both individual landslide-type susceptibility maps and the overall maps for the WoE and GAMB models. This evaluation utilizes the IFFI inventory as the ground truth.

2.5.3 Class delineation (STEP 3)

One of the key challenges of susceptibility analysis is the definition of susceptibility classes. This definition is crucial for decision-making processes, such as categorizing areas as low, moderate, or high susceptibility, and for communicating the results to end-users like construction companies (Fell et al. 2008).

The literature proves to be fractionated and abundant with different methods (see Table 2.2) around this concept of defining thresholds, revealing a lack of a common standard (Reichenbach et al. 2018). For instance, Normal distribution emphasizes the mean and standard deviation; F-beta scoring try to minimize the number of false negatives (at the increasing of beta); AUC optimization identifies the set of thresholds that best balance the true positive and false positive rates; Natural Breaks is the best for minimize the variance between natural cluster in the distribution; and percentile-based methods (overall/landslide area), though common in literature, are highly dependent on the chosen cutoff values (e.g., 25th, 50th, 75th percentiles). However, none of these approaches alone guarantees that the resulting classes will accurately reflect the real variability of susceptibility across the landscape.

Method	References
No re-classification	Brenning (2005), Catani et al. (2013), Lombardo et al. (2015), Trigila (2015).
Percentile of whole area values	Carrara (1983), Ayalew & Yamagishi (2005), Goetz et al. (2011), Meyer et al. (2014), Steger et al. (2016), Lombardo & Mai (2018), Cantarino et al. (2019), Merghadi et al. (2020), Alvioli et al. (2024).
Percentile of observed landslide values	Van Westen et al. (2003), Guinau M. et al. (2007), Petschko et al. (2014), Sinčić et al. (2022).
Normal distribution (Mean plus/minus standard deviation)	Ayalew & Yamagishi (2005)
Equal intervals	Atkison & Massari (1998), Ayalew & Yamagishi (2005), Guzzetti et al. (2005), Simoni et al. (2008), Conoscenti et al. (2016), Alvioli et al. (2024), Dahal et al. (2024), Moreno et al. (2024).
Specific arbitrary interval	Guzzetti et al. (2006a), Frattini et al. (2010), Rossi et al. (2010), Yu et al. (2016), Lombardo et al (2020), Rosi et al. (2023), Bozzolan et al. (2023).
Binary classification followed by a spatial agreement between models or inventory	Ardizzone (2002), Carrara et al. (2008).
Natural breaks (Jenks)	Ayalew & Yamagishi (2005), Peng (2014), Sciarra et al. (2017), Chen et al. (2017), Cantarino et al. (2019), Alvioli et al. (2024).
Head/Tail breaks	Cantarino et al. (2019).
ROC based (AUC optimization, Youden index, Generalized Youden Index, Specific thresholds)	Cantarino et al. (2019), Martinello (2021), Martinello (2022), Titti et al. (2021), Titti et al. (2022), Steger et al. (2024).
Geometrical interval	Chen et al. (2018).
Effectiveness ratio	Chung & Fabbri (2003), Fang et al. (2024).

Table 2.2. List of proposed methods in literature to reclassify the landslide probability values.

To address this issue, we developed a novel methodology of ensembled methods for categorizing susceptibility, which is depicted in Figure 2.5. For each pixel with a given P_{ovr} for each landslide type separately, we applied six distinct reclassification techniques. Each technique employs different criteria to establish three thresholds that delineate four susceptibility classes: very low (1), low (2), moderate (3), and high (4). We opted for this 4 classes' terminology given the overcautious nature of the IFFI landslide inventory. The methods used are as follows:

1. **Overall percentiles** consider the full distribution of susceptibility values across all pixels, setting the 25^{th} percentile as the threshold between classes 1 and 2, the 50^{th} percentile between classes 2 and 3, and the 75^{th} percentile between classes 3 and 4
2. **Normal Distribution** calculates the mean and standard deviation of the entire dataset and places the three thresholds at one standard deviation below the mean susceptibility, the mean itself, and one standard deviation above the mean.

3. **Landslide percentiles** operate similarly to the first method but focus solely on the pixels within the landslide areas to derive thresholds tailored specifically for unstable slopes.
4. **Natural Breaks (Jenks)** finds the best arrangement in four classes aiming to minimize variance within classes while maximizing variance between them.
5. **AUC-optimized thresholds** are obtained by discretizing the ROC curve into 4 straightforward segments where the changing points are selected in order to maximize the AUC under the new segmented ROC, for optimal sensitivity and specificity.
6. **F-beta scores** (Van Rijsbergen, 1979) define thresholds using F-scores with different weights (β), to emphasize true positives (TP) and false negatives (FN) in order to minimize the latter. In the formula:

$$F_\beta = (1+\beta^2)\frac{(1+\beta^2)*TP}{(1+\beta^2)*TP+\beta^2*FN+FP} \quad [2.9]$$

The factors β utilized were 1, 2 and 3; where 3 represent the lowest threshold as it minimizes the FN at the expense of FP (more conservative approach), followed by F2 and F1. So, we compute the F1, F2, and F3 using different thresholds of P_{ovr}, and for the highest F1, F2, F3 we kept the relative threshold.

Each method provides a susceptibility class for the pixel, ranging from 1 (very low) to 4 (high). We then determine the pixel class by calculating the mode of these classifications, along with the level of agreement between methods. For example, as shown in Figure 2.5, a pixel with a susceptibility value of P_{ovr}=0.23 is classified differently depending on the method used: the first 3 methods and the 5th classify it in class 2 (low susceptibility), whereas method 4 categorizes it as class 1 (very low susceptibility), and the last method in class 3 (moderate susceptibility). The most common classification across the methods is class 2 (low susceptibility), occurring with a 67% consensus (four out of six methods assign it to class 2).

After this initial classification, the susceptibility values of the four classes are extracted to calculate the corresponding kernel density function. Ideally, density functions should be entirely separated without overlapping values, since they represent different degrees of susceptibility. However, especially close to the class boundaries, it may happen that pixels with higher susceptibility fall in the lower class. To address this issue, we redefined the thresholds between classes by identifying the intersection points between the kernel densities distributions. These updated thresholds are then applied to reclassify the original P_{ovr}, ensuring clearer delineation in boundary cases.

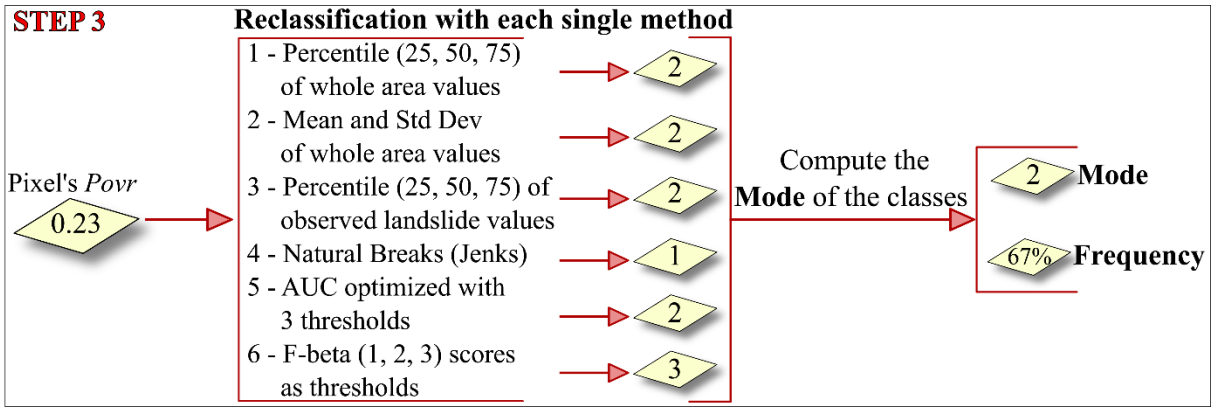


Figure 2.5. Example of reclassification method with the mode concept.

This approach was implemented independently for the two geological domains, A and B. As mentioned earlier, the landslide characteristics differ significantly between these areas: Domain A predominantly experiences rock falls and rapid rockslides, while Domain B is characterized by slow-moving earth slides and flows (Figure 2.3). Using the same numerical thresholds for such diverse landscapes could lead to inaccurate representations of the actual landslide hazard, as identical susceptibility classes might correspond to very different types of instability phenomena. Therefore, we set distinct thresholds for each domain based on their specific probability values, ensuring that our classification better reflects the hazard levels of the landscape.

2.5.4 Results visualization (STEP 4)

The final step in the analysis involves visualizing the results of the susceptibility models along the examined rail line. An effective visualization strategy should provide a clear and immediate representation of the critical sections along the railroad route, together with essential information about the type of landslide hazard. Additionally, it should consider the average susceptibility of areas adjacent to both the left and right sides of the route, in order to determine which side of the railway is more vulnerable to landslide.

To summarize all this information in a single plot, we first sampled the points (pixels) along the railway line and selected all pixels within a 100-meter radius. From these selected pixels, we further refined our focus to include only those within a 120° arc (±60 degrees from the perpendicular to the railway line) on each side, separately (see Figure 2.6). After identifying the relevant pixels on each side, we calculated the average susceptibility for these groups and then plotted these averages together to allow direct comparison between the left and right sides. This step provides clear insights into which side of the railway is more prone to landslides. The choice of a 120° arc with a 100-meter radius on each side of the railway was intentional, as it effectively focuses less on pixels closest to the railway, so reducing redundant information from immediately adjacent pixels. Indeed, these pixels often mirror the information already captured directly from the susceptibility plot along the rail line, so by emphasizing pixels further away, this approach offers a broader and clearer understanding of landslide susceptibility in the surrounding slope area.

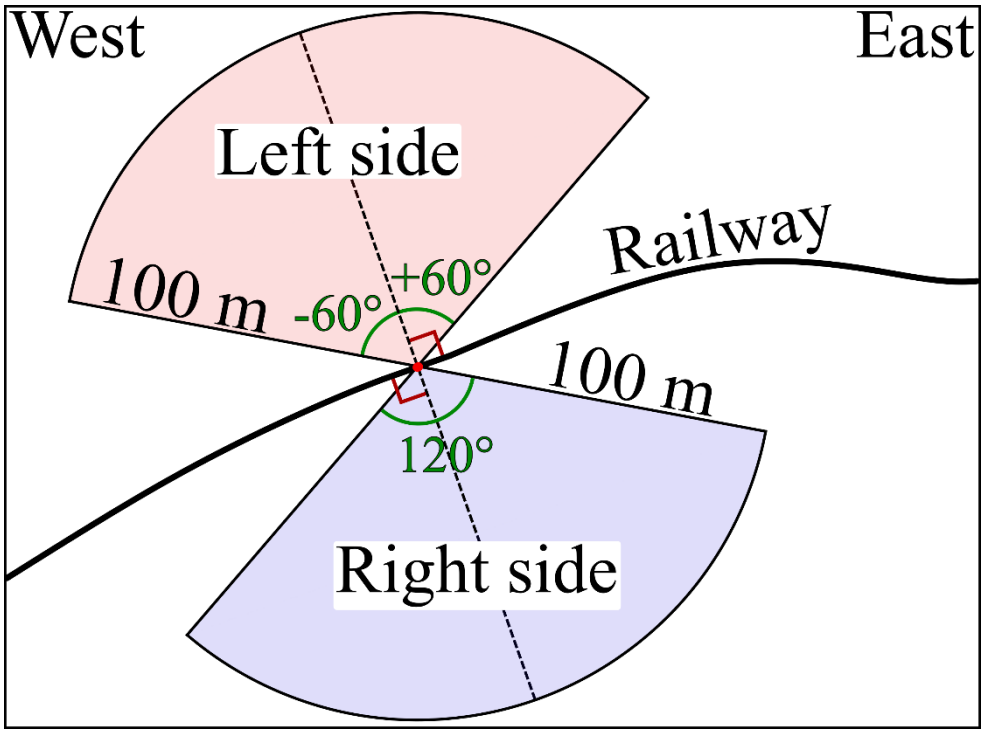


Figure 2.6. Representation of pixels selection (left and right side) method from each pixel on the line (red dot on railway).

2.6 Results

2.6.1 Landslide susceptibility maps

Figure 2.7 displays the landslide susceptibility maps generated using the Weight of Evidence (WoE) and Gradient Boosted Generalized Additive Model (GAMB) methods. These maps illustrate the overall susceptibility, calculated by the complementary probability of susceptibility for each individual landslide type, as detailed in section 2.5.2.2. Maps for the various landslide types are included in the supplementary material (S 2.1, S 2.2, S 2.3, S 2.4, and S 2.5). Figure 2.7 includes also the confusion matrix wheel (easier reading), build up using the 0.5 as binary thresholds, showing True Positives (TP), True Negatives (TN), False Positives (FP), and False Negatives (FN). Thus, TP indicates for example where susceptibility is higher than 0.5 and observed landslide is present. The maps in Figure 2.7 show the importance of varying geological conditions within the study area and highlight differences between the two models. Specifically, domain B exhibits greater susceptibility to slope instability, particularly in the GAMB model, whereas the WoE model tends to emphasize some areas in the carbonate ridge (domain A). Both models identify the main valleys with low susceptibility, while WoE points out also the smaller valley and gullies in domain B. Overall, GAMB appears more homogeneous inside the two respective domains; on the other hand, the WoE model reveals more heterogeneous and scattered susceptibility. The confusion matrix (wheel) underscores this first conclusion with GAMB achieving a slightly better identification of TN and especially TP pixels, slightly reducing the FN of WoE.

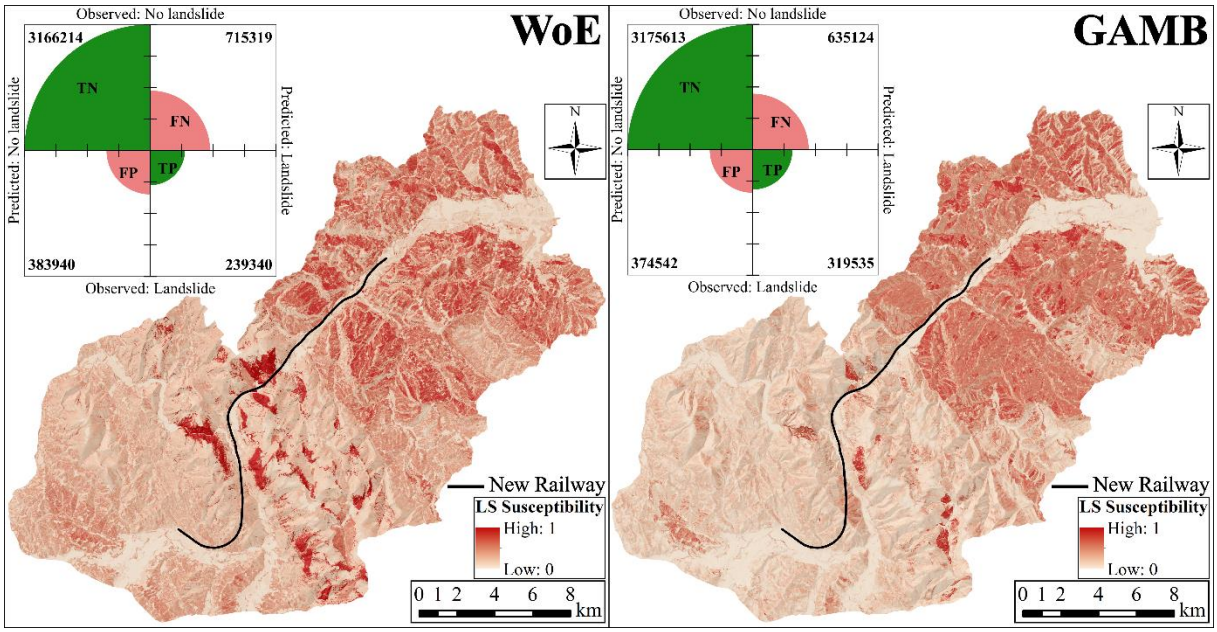


Figure 2.7. Overall landslide type susceptibility [0 – 1] for both methods with relative confusion matrix (wheel).

The sensitive maps in Figure 2.8 highlight the contribution of each landslide type to the overall susceptibility. This information is crucial for railway planning as it helps identify the more hazardous landslide types, such as rapid collapses and slides, over less hazardous ones like slow flows.

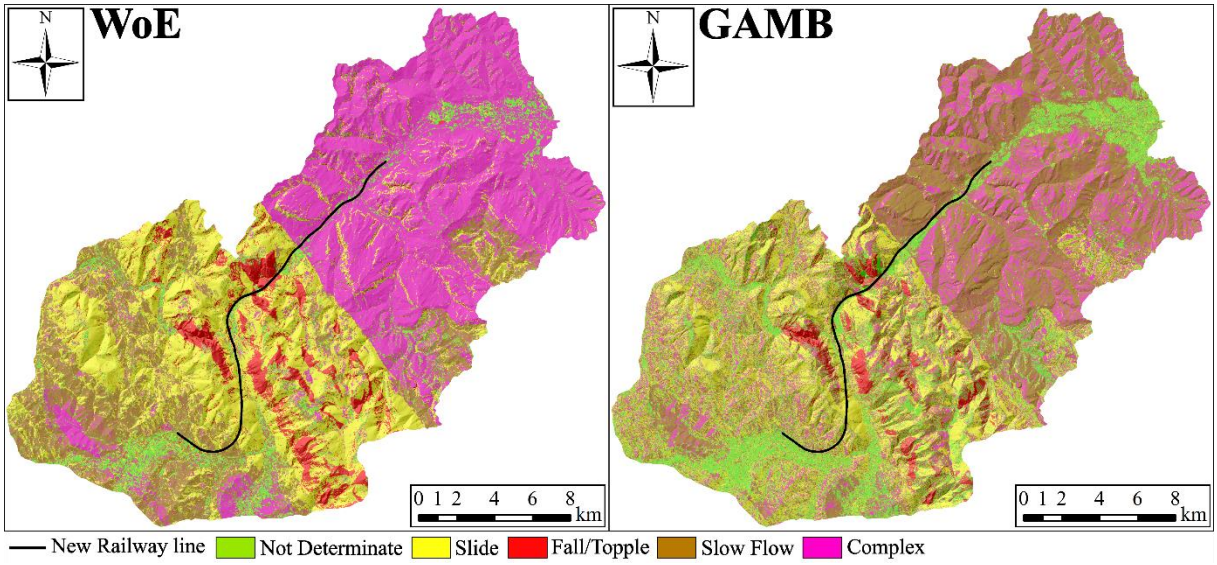


Figure 2.8. Sensitivity maps of landslide types for both methods.

The Slow-Flow (SF) and Complex (C) types appear to be the most challenging to distinguish due to similar geomorphological conditions, leading to divergent results between the models. The WoE model predominantly associates the hilly slope areas of domain B with the Complex landslides. In contrast, GAMB differentiates these areas into slopes affected by both Complex and Slow Flow types, mainly based on the aspect of the slope. Other differences are visible in the north-west portion of domain A. Here, like in domain B, GAMB tends to associate larger area to SF, instead WoE to C. In domain A, comparing with the landslide inventory (Figure

2.3), we notice that the GAMB is more accurate assigning the SF type presence, while in domain B, WoE is closer to the original IFFI classification assigning C type. Both models indicate that Slide (S) and Fall/Topple (FT) are the landslide types that control susceptibility in domain A, especially along the carbonate ridge. Here, the FT type shows a broader influence in the WoE model compared with the GAMB, covering wider.

In the valleys we notice that almost everywhere the predominant sensitivity of landslide type is Not determinate (ND) landslide. This behavior may be explained by two main factors. (i) In valley areas, the influence of other landslide types may be reduced because the models are better at recognizing the distinctive conditions associated with those types, and consequently better at excluding areas with opposite features, such as valley. In the ND case, unclear geomorphological features lead to less recognizable features, and it is more difficult to exclude opposite conditions too. (ii) The ND type is the most unbalanced in the dataset, with a very limited number of landslides. This imbalance may lead to a less robust model for the ND class, potentially exaggerating its sensitivity, particularly in zones with generally low susceptibility. The different behavior of the two models is further in-depth examinate in the discussion paragraph with the help of contrast and partial dependences presented in the supplementary material.

Each landslide type susceptibility model was evaluated using the ROC method and the corresponding AUC values, as shown in Table 2.3. The Fall/Topple type achieved the highest AUC values with both models (WoE and GAMB), indicating that these geomorphological features are the most easily recognizable in this area. Conversely, the Slide type resulted in the lowest AUC values for both models. Generally, the AUC values are within a normal/good range compared to other studies in the literature (Alvioli et al., 2024). When comparing the two models, the GAMB consistently outperformed the WoE model, with a notable difference (0.05 - 0.1) in all landslide type analyses. The overall susceptibility map, evaluated using the entire landslide inventory as ground truth, showed the poorest results. This outcome is expected because the overall evaluation assesses the combined effect of different geomorphological features on various landslide types, each influenced by different factors.

Model's AUROC	**Not determinate**	**Slide**	**Fall/Topple**	**Slow Flow**	**Complex**	**Overall**
WoE	0.785	0.699	0.941	0.769	0.763	0.685
GAMB	0.868	0.784	0.978	0.849	0.821	0.734

Table 2.3. AUC values from the ROC validation method.

2.6.2 Classes delineation

A Clear and univocal identification of susceptibility classes is not possible looking at the distributions of the landslide probability in different domain (P_{ovr}) shown in Figure 2.9. The WoE inside domain A presents a clear frequency distribution pick of P_{ovr} around 0.1-0.2, and then no other evident picks, if not at the highest values close to 1. GAMB in domain A, instead, display one clear pick of frequency at P_{ovr} close to 0, and then another one, wider, at 0.1. In domain B, the WoE shows a noisy distribution and almost flat from 0 to 0.5 of P_{ovr}. The GAMB

in domain B still shows the high pick of very low values, plus a less evident pick at around 0.5. In both domains, the GAMB model identifies most pixels outside potential landslide areas (e.g., flat valley areas) correctly, assigning them extremely low landslide probabilities. Conversely, the WoE model tends to flattener all values, failing to distinguish these extremely very low-probability pixels effectively. These results are common in susceptibility analysis, which often leads to uncertainty in the delineation of descriptive classes (see section 3.3).

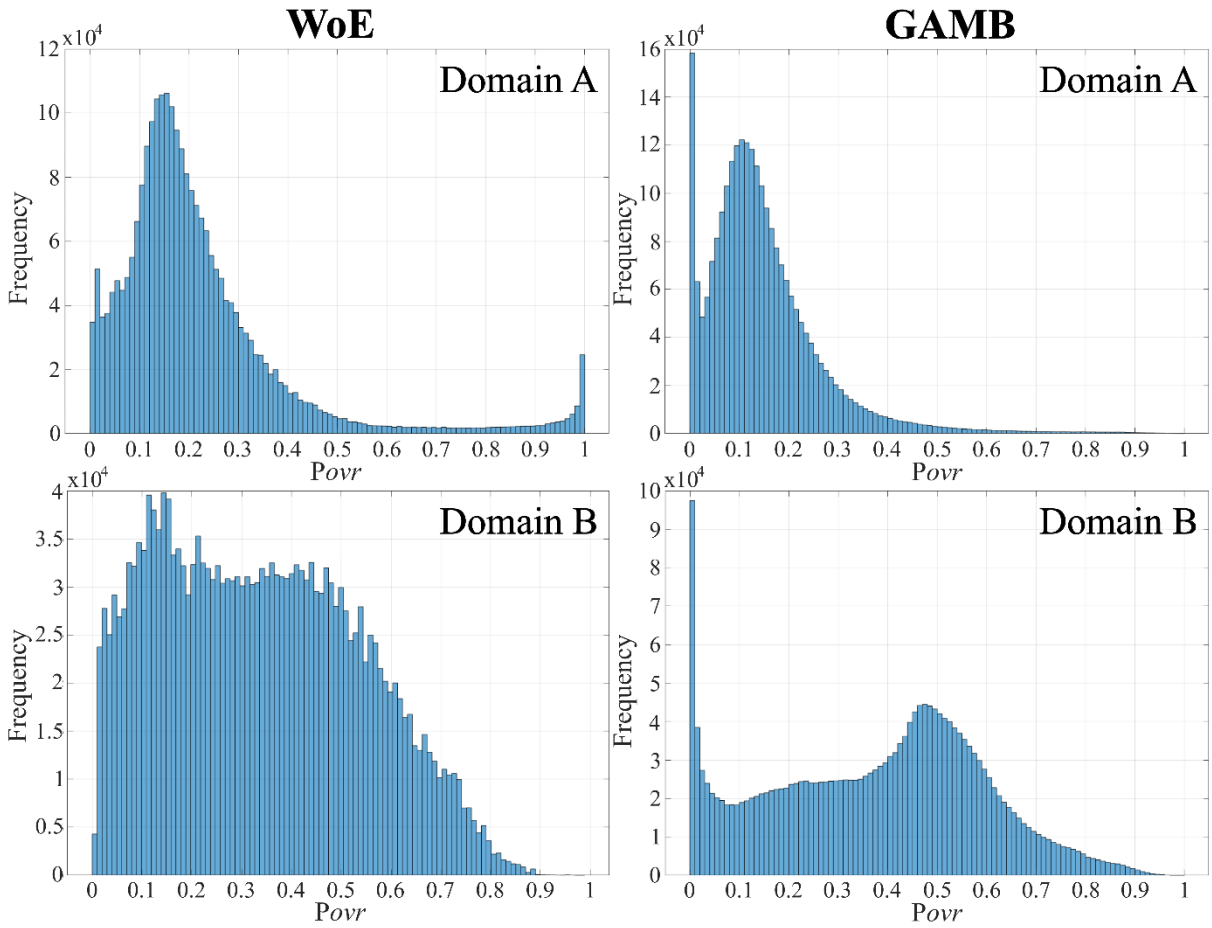


Figure 2.9. Frequency distribution of the overall probability per model over the domain A and B.

The application of the six classification methods outlined in section 3.3 aids in objectively defining the susceptibility. Table 2.4 presents the three threshold values of P_{ovr} for each method, which delineate the four susceptibility classes: very low, low, medium, and high. Generally, the Natural Breaks method is the least cautious, featuring relatively high thresholds that result in a narrower high-susceptibility class. Landslide percentiles provide also high thresholds, particularly the third threshold that separates medium to high susceptibility in WoE – B. On the contrary, the F-beta method appears the most cautious. The thresholds obtained from AUC optimization method are placing always in the mid-range compared to the other methods, apparently becoming the best one to avoid over optimistic or pessimistic classification in our study area.

Model - Domain	Thresholds	Overall percentiles	Normal Distribution	Landslide percentiles	F-beta	AUC optimized	Natural Breaks
WoE - A	Very low - Low	0,113	0,041	0,112	0,065	0,093	0,157
	Low - Moderate	0,182	0,221	0,176	0,111	0,180	0,304
	Moderate - High	0,272	0,401	0,327	0,180	0,322	0,630
WoE - B	Very low - Low	0,168	0,133	0,165	0,060	0,136	0,196
	Low - Moderate	0,322	0,339	0,309	0,117	0,292	0,383
	Moderate - High	0,487	0,545	0,561	0,245	0,432	0,576
GAMB - A	Very low - Low	0,074	0,030	0,079	0,044	0,060	0,102
	Low - Moderate	0,126	0,149	0,131	0,085	0,131	0,218
	Moderate - High	0,194	0,267	0,279	0,170	0,244	0,448
GAMB - B	Very low - Low	0,200	0,158	0,228	0,069	0,126	0,175
	Low - Moderate	0,403	0,374	0,414	0,121	0,361	0,384
	Moderate - High	0,533	0,590	0,597	0,343	0,525	0,595

Table 2.4. Resulting thresholds trio (from lower to higher) from each single reclassification method for the two models WoE and GAMB inside domain A and B. Color scale (green-yellow-red) from the lowest to the highest method's threshold (row) per each model – domain.

The final thresholds for the susceptibility classes are determined for each method and geological domain by plotting the kernel density distribution of the reclassified pixels (Figure 2.10). In these plots, various colors represent the frequency distribution of the probability P_{ovr} for pixels within a given class, as defined by the modal class across the six classification methods. The figure highlights the differences between the west (A) and east (B) domains, as well as between the two models (WoE and GAMB), aligning with the histograms in Figure 2.9. It can be observed that there is a lack of clear separation, as discussed in Section 3.3. In the WoE model, within domain A, the division between the very low and low classes is not distinctly separated in the distribution itself (small peak of green line inside the yellow one). Therefore, the lower intersection points between distributions are adopted to ensure a more cautious delineation between susceptibility classes.

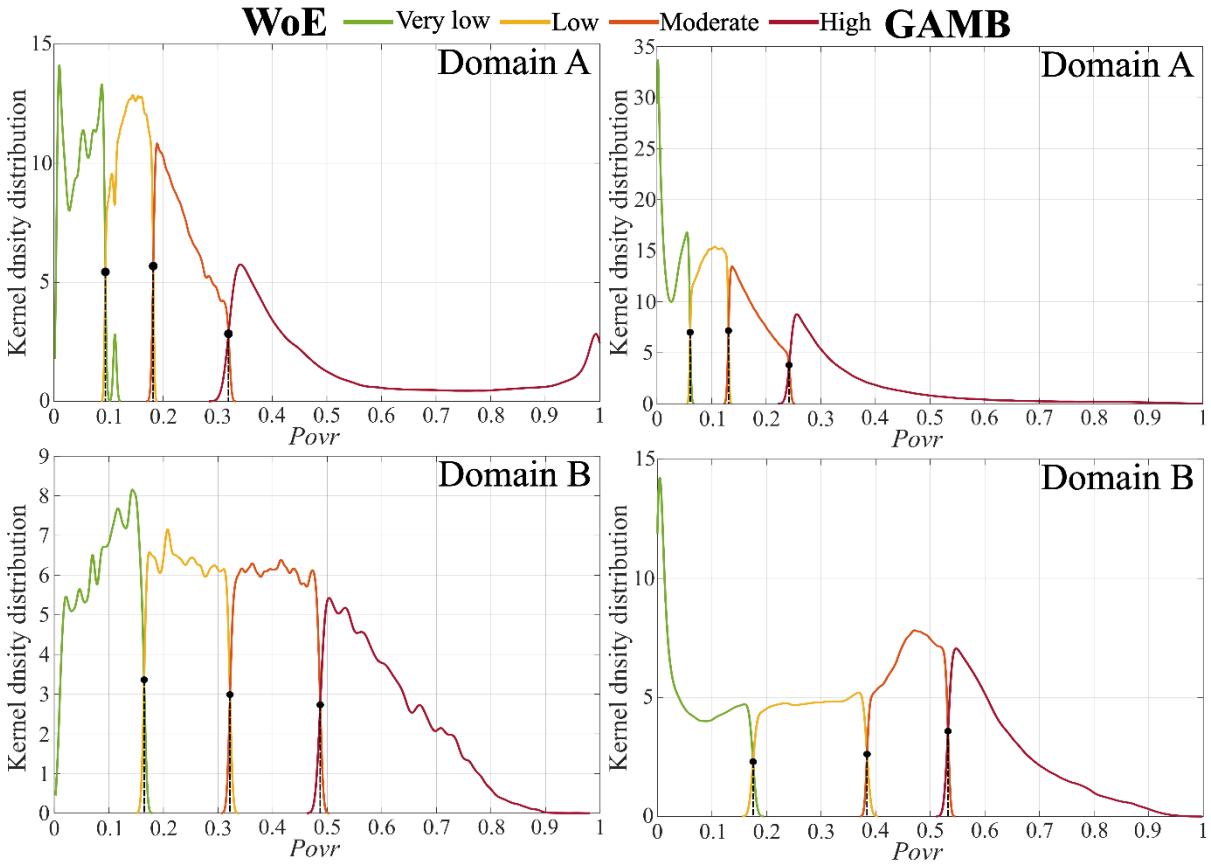


Figure 2.10. Density distribution of overall susceptibility (Povr) for each class per domain and method.

As a result of the thresholds delineation, we obtained the final susceptibility maps shown in Figure 2.11. The very low susceptibility class predominantly corresponds to valley areas, as well as some southern hilly sectors within domain B, where turbiditic and marl lithologies are prevalent (see Figure 2.1). The low and moderate classes cover extensive portions of domain A, particularly to the west of the carbonate ridge. There we can observe that both models (even if with different shapes) identify as highly susceptible area with gypsum and clay lithologies, affected by Slow flow/Complex landslide. The GAMB spots also a highly susceptible area in slightly to the north, not spotted by WoE.

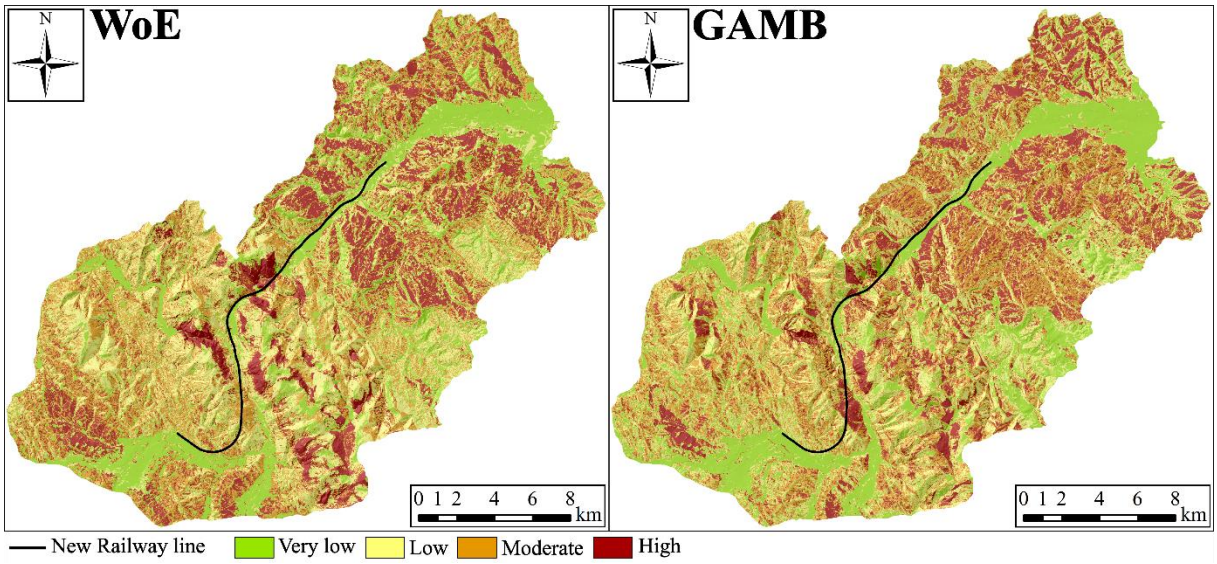


Figure 2.11. Final reclassified maps in four classes of overall susceptibility.

Along the carbonate ridge, WoE tends to classify broader areas as highly susceptible compared to GAMB, except for a specific slope crossed by the railway line, which is classified as high susceptibility by GAMB and only moderate by WoE. Notably, almost all areas set as highly susceptible by WoE in the ridge are related to Fall/Topple landslide type. Instead, the GAMB gives attention also to both Fall/Topple and Slide types. In domain B, the two models show both agreements and divergences. A significant difference appears immediately east of the ridge, on the slopes flanking the railway. Here, the GAMB model classifies the terrain as highly susceptible, while the WoE model assigns low to moderate susceptibility levels.

The frequency of the mode in the reclassification process indicates the level of agreement between each single reclassification method, providing valuable insights into areas with higher or lower consistency in the results, and assessing the reliability of the assigned class for each pixel. As shown in Figure 2.12, there is greater agreement in valley areas and regions with high susceptibility values, reflected by clear very low and high susceptibility levels, respectively. Conversely, areas with intermediate probability values exhibit more variability. Domain A shows generally less agreement between classification methods (salt-pepper noise effect), probably due to greater topographical heterogeneities (rapid changes) given by the mountain region. Additionally, the GAMB model generally demonstrates better overall consistency (highlighted also in Figures 2.9 and 2.10), which means generally better reliability in practical applications.

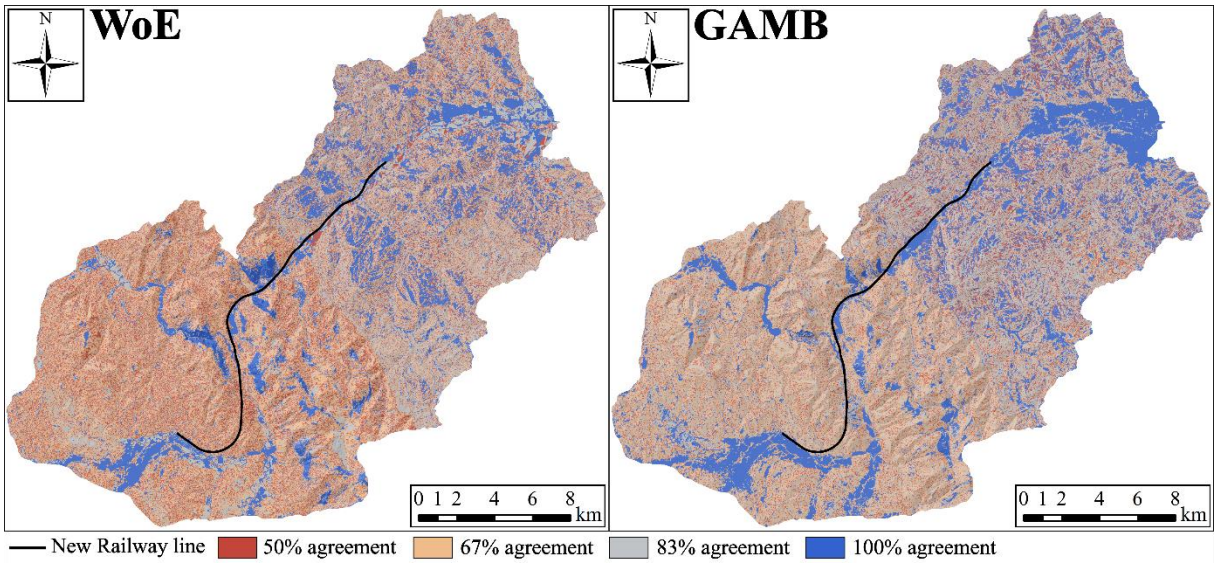


Figure 2.12. Agreement between reclassification methods for both methods.

2.6.3 Visualization along the railway

Figure 2.13 shows the values of landslide susceptibility along the railway line, with lines colored according to the type of landslide that predominantly affects the susceptibility levels. The background-colored fields indicate the four susceptibility classes within the two geological domains. As can be seen, there are notable differences between the outputs of WoE and GAMB, similar to those seen in Figure 2.7. The WoE model generally shows higher overall landslide susceptibility (P_{ovr}) in domain A, whereas the GAMB model indicates higher susceptibility in domain B. Between 4.5 and 6.5 km, the GAMB model indicates high susceptibility associated with Slide, whereas the WoE model classifies this section as low to moderate. At 11 to 11.5 km, the WoE model highlights a high susceptibility section related to Fall/Topple, which is not present in the GAMB model. Instead, the GAMB model emphasizes the section just after this, up to 12 km, with a focus on Slide. A significant critical peak section at 13 to 14 km in the WoE model is associated with FT landslides, but this is considerably reduced in the GAMB model. Moving to Domain B, the GAMB model presents slightly higher peaks in susceptibility. However, the main difference lies in the identification of the most influential landslide types: the GAMB model indicates Slow flow as the predominant type, while the WoE model identifies Complex landslides.

Figure 2.13 also directly compares the susceptibility profiles with the mapped landslide types along the proposed alignment (grey band at the top of the WoE and GAMB plots). From 0–4 km, WoE aligns better with mapped Slow Flow; from 4.5–6 km, GAMB captures the extended Slide that WoE underrates, and both miss a small Slow Flow near km 7. Around 13–14 km both flag a critical reach (WoE stronger, attributed to FT); at 14.4 km both recognise the mapped FT (GAMB High; WoE Moderate–High). At 14.5–15 km a mapped Complex landslide is rated

High by GAMB and lower by WoE; types are partly misclassified, but the section is consistently highlighted as hazardous.

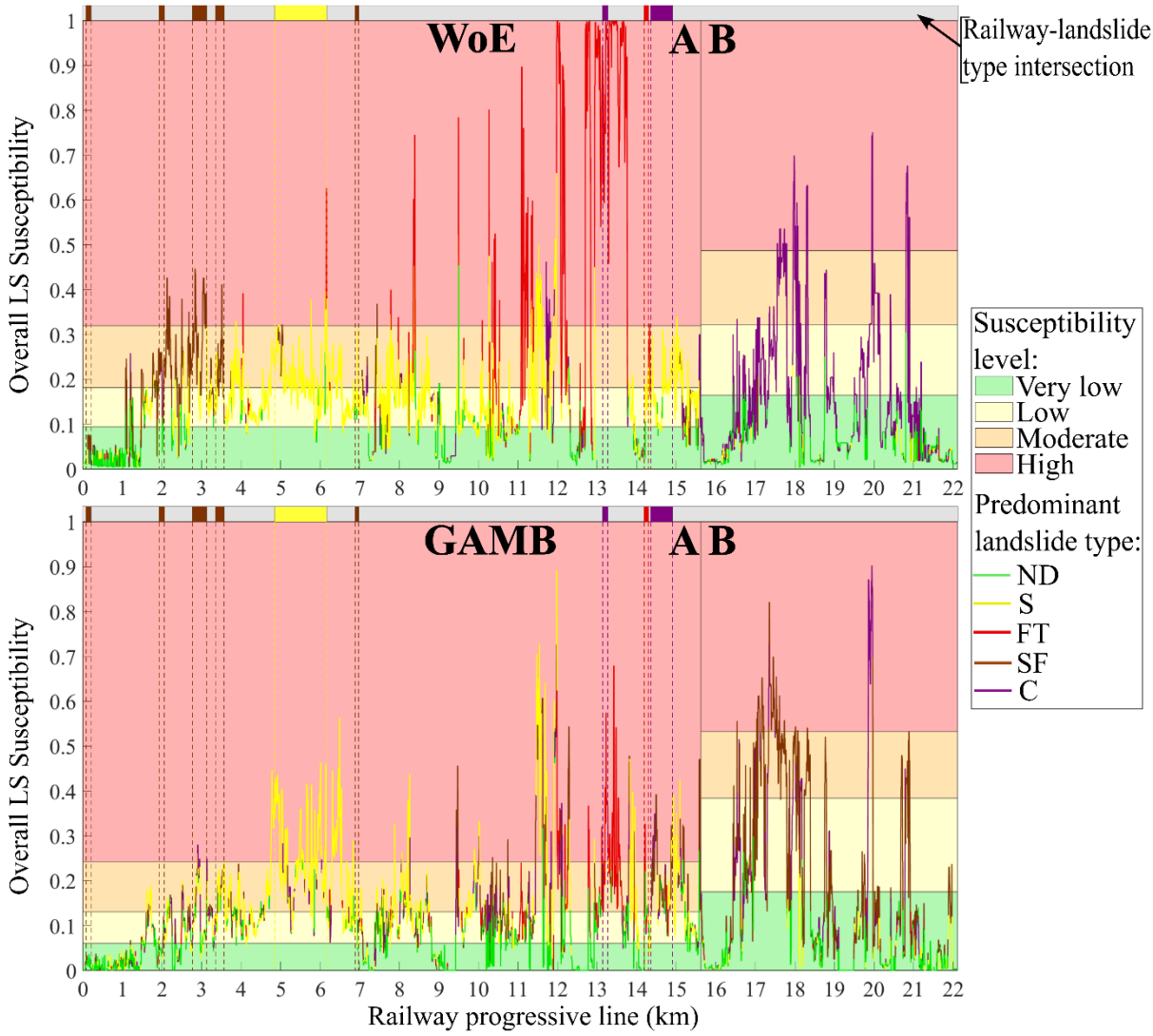


Figure 2.13. WoE and GAMB overall susceptibility along the railway. Colours on the lines denote the dominant landslide type; the grey bar at the top highlights railway segments intersecting mapped landslides, colour-coded by type.

To quantify how well the susceptibility profiles reproduce the observed distribution of landslides along the alignment, we computed a pixel-wise confusion matrix over the entire corridor. A pixel was counted as positive when its overall susceptibility class was high; the ground truth was the presence/absence of an intersection between mapped landslide polygons and the railway trace. Thus: TP = intersection and high susceptibility; TN = no intersection and susceptibility < high; FP = no intersection but high susceptibility; FN = intersection but susceptibility < high. The resulting overall accuracy, (TP+TN)/(TP+TN+FP+FN), is 0.797 for WoE and 0.837 for GAMB, values generally considered good for linear-infrastructure screening in heterogeneous terrain (Bordoni et al. 2018, Quinn et al. 2010). This evaluation is deliberately strict: if a mapped landslide intersects the line and the model does not label that segment as high, we count it as an FN. So, some mapped interactions are not placed in the top class. At the same time, the models flag other segments as high where no mapped landslide

intersects the line (FP). In a susceptibility framework, these FP segments are useful outputs as they point out to new stretches with high predisposition that our analysis brings to light. Therefore, even if not every mapped interaction is on a high susceptibility level, the profiles expand the list of priority reaches. In early design stages, when detailed geotechnical data are limited, these FP-prioritized segments provide a practical guide for field investigation, monitoring, and conservative design choices. However, it is impossible to determine the relative "accuracy" of a currently critical healthy slope (FP), as the susceptibility maps express the underlying predisposition to failure (and likely type), not event timing. As a result, comprehensive accuracy cannot be fully assessed at present: sections highlighted as critical today may host future landslides, while currently active zones may stabilise. In this sense, only continued observation over time will provide the most meaningful evaluation of both the susceptibility model and the critical sections it identifies.

Figure 2.14 illustrates the average landslide susceptibility within a 100-meter buffer alongside the railway. Here, the lines are color-coded based on the side of the buffer where the susceptibility is averaged: red for the left side and blue for the right side. In general, the railway seems to be more exposed on the right side in both models. In contrast, the left side features smaller or less critical portions. This information can be used to modify the original track in the preliminary phases of the project, avoiding the dangerous side. In our case, for example, both models exhibit a high degree of susceptibility on the left side, but not in the right, at 11.5-12 kms. Right after, at 12-12.2 km there is a peak (greater in WoE) of the left side and a smaller one of the right side. This switch from higher susceptibility on the left side to the right side

within 500 meters of line in both models (with different entities) is valuable information in preliminary tracing, perhaps suggesting a variation of the railway curvature.

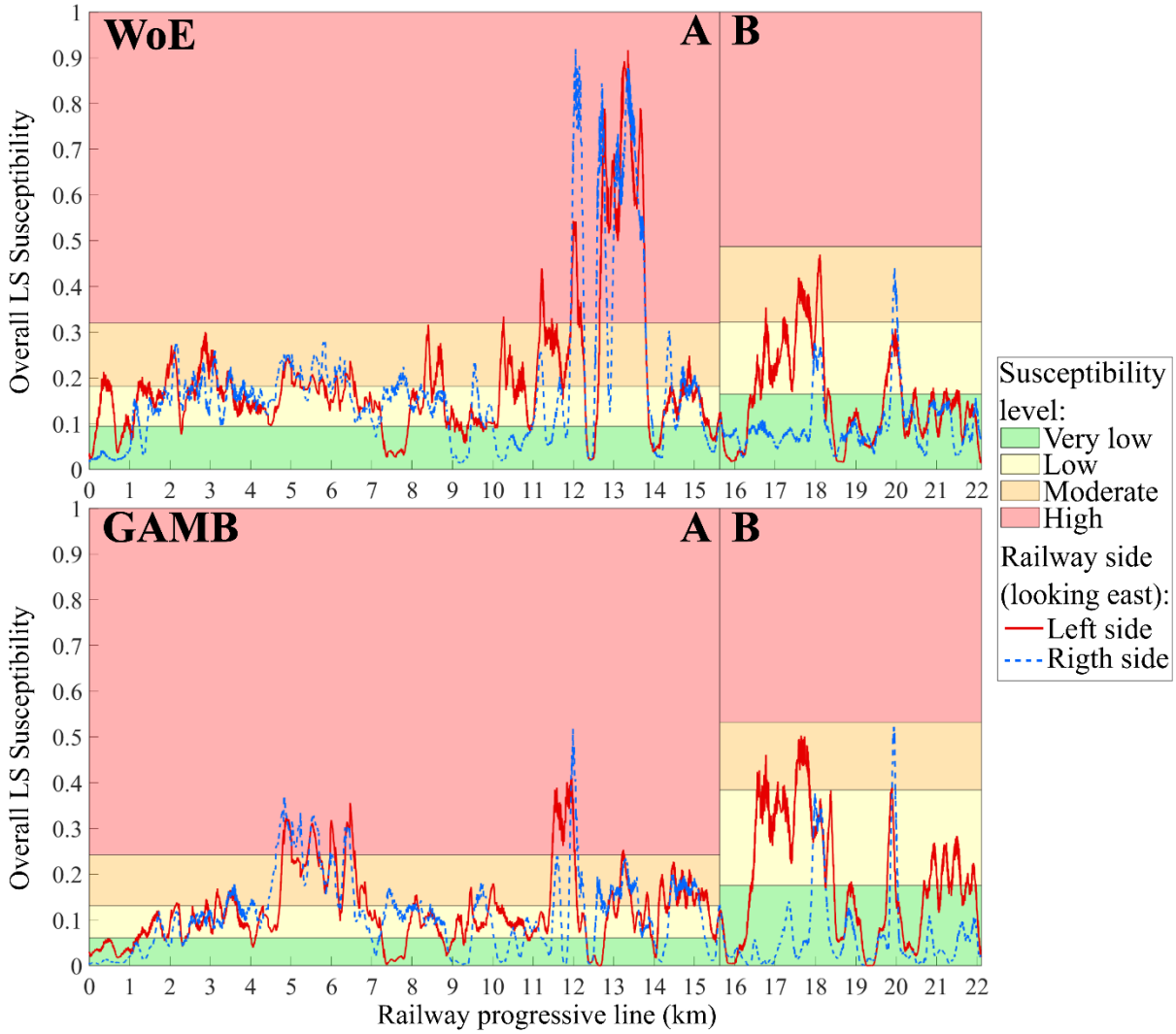


Figure 2.14. Averaged overall susceptibility along the two sides of the railway with a buffer of 100 m within an arc of ±60° from the perpendicular of the railway line.

2.7 Discussion

Differences between models

In our study area, the GAMB results are better from the metrics performance referring to the single landslide types. Moreover, GAMB shows better spatial consistency of the overall probability (P_{ovr}). This is visible in the histogram distribution (Figure 2.9), which turns into a better general agreement between reclassification method (Figure 2.12). This feature makes us aware that the GAMB better preserves the meaning of geomorphological homogeneity of pixels close to each other, resulting in a less "salt and pepper" effect, which is instead returned by the WoE, especially in the B domain.

On the other hand, the WoE appears better in detecting the most influencing landslide type per pixel regarding the Slow flow and Complex ones (see Table 2.5). GAMB drastically overestimates the Slow flow influence in domain B at the expense of the Complex type.

However, this tendency does not undermine the performance of the GAMB model for Slow Flow and Complex landslides. The sensitivity calculation, based on partial derivatives of the overall susceptibility with respect to each landslide type, reveals that Slow Flow and Complex are very close in influence for GAMB in Domain B. As a result, even minor input variations can shift the dominant type, indicating the model's difficulty in clearly separating them. This overlap likely reflects their shared geomorphological conditions, such as clay-rich lithologies and gentle slopes, and may also point to classification uncertainties in the original landslide inventory. Nonetheless, GAMB provides added interpretive value by highlighting areas where Slow Flow may be more likely than Complex, potentially refining the initial inventory classification.

Model	Metric	Not determinate	Slide	Fall/Topple	Slow Flow	Complex	Overall
WoE	**Sens/Truth (%)**	63.1	109.9	145.8	89.9	100.0	101.7
	Accuracy	0.92	0.78	0.96	0.85	0.72	0.62
GAMB	**Sens/Truth (%)**	132.8	97.4	109.5	295.2	41.4	135.3
	Accuracy	0.89	0.81	0.97	0.48	0.52	0.34

Table 2.5. Landslide type sensitivity attribution evaluation on actual landslide areas. Ratio sensitivity over ground truth in percentage (Sens/Truth), if under 100, landslide attribution by sensitivity is underestimating, if 100 perfect matches, if above it is overestimating. Accuracy represents the ratio between all correct classification (true positive and true negative) over the total.

The difference in behavior between the two can be explained by their Contrasts (S 2.6 – S 2.10) for WoE and Partial dependences (S 2.11 – S 2.15) for GAMB. The contrasts (CS) are the difference between W^+ and W^-, to indicate each category's contribution to instability. Partial dependence (PD) plots illustrate how the model's predictions change on average when a single predictor varies, while the effects of all other variables are averaged out (Friedman 2001). In practice, this is done by fixing one predictor (e.g. slope) to a chosen value for all pixels (e.g. slope = 10°), leaving the remaining predictors at their original values, and then averaging the predicted susceptibility across the dataset. The procedure is repeated for successive values of the predictor (e.g. 11°, 12°, …), thus constructing the plot. To account for uncertainty, we included 95% confidence intervals derived from 100 bootstrap resamples. This approach provides an intuitive way to interpret complex models by isolating the general effect of selected predictors (Nocentini et al. 2024, Schlögl et al. 2025), while adding statistical robustness through bootstrapped confidence intervals.

The main difference between WoE and GAMB is that the latter can examine the predictors as continuous values, apart from those categorized by nature (geology, land cover, and bedding plane–slope orientation geometrical relationship), which is visible in the CS and PD plots. Comparing each single predictor contribution between WoE and GAMB for each landslide type.

For Not determined landslides (WoE S 2.6 vs GAMB S 2.11), WoE shows strong categorical contrasts (geology/land cover) and a stepwise, generally decreasing slope effect (peak ~10°–12.5°); GAMB/PD moderates category extremes, reveals a nonlinear slope response with peaks at ~15° and ~40° (lower confidence), and captures a TWI peak–threshold (~10–15) and sharper curvature hotspots than CS smooths. For the Slide type (WoE S 2.7 vs GAMB S 2.12), WoE emphasises marls/limestones and mixed woods, with slope peaking at ~22°–28°; GAMB confirms the lithology but shifts land-cover influence toward urban green areas and moves the slope peak higher (~30°–40°), with PD showing a higher influence from flat (near-zero) planar curvature and from convex profile curvature (upward camber), plus a clearer TWI peak (~12–16). Both approaches converge in highlighting the greater predisposition of cataclinal slope settings (1, 2, 3), consistent with expected kinematics. With the Fall/Topple type (WoE S 2.8 vs GAMB S 2.13), WoE yields a near-monotonic slope increase (strong beyond ~35°) and favours meadows/sparse cover; GAMB/PD sharpens lithological control for compact limestones, introduces a slope threshold at ~40°–45°, highlights urban/excavated classes, confirms negative TWI with a drop at ~5–10, and stresses planar–convex curvature typical of cliffs. Both models also highlight the strong positive effect of anaclinal over-dip conditions (9), demonstrating that bedding planes can act as discontinuities from which vertical fractures propagate, thus favouring topples. For the Slow Flow type (S 2.9 vs S 2.14), both models agree on clay-rich/marly units; WoE favours arable land and gentle slopes (5°–15°) with concave curvature, while GAMB/PD broadens land-cover influence (including meadows), concentrates the slope response around ~15°, shows a clearer TWI band (~12–16), and shifts curvature signals toward planar (flat) and slightly convex profiles (with higher uncertainty). Lastly, for the Complex type (S 2.10 vs S 2.15), both models point to clay-rich/fine-sandstone units; WoE associates highs with arable/tall-tree areas and gentle slopes (5°–15°) and concave curvature, whereas GAMB/PD prefers meadows/mixed woods, adds a mid-slope peak (~25°–35°) before decline, shows TWI rising then stabilising (~16) without decreasing, supporting the interpretation that hydrological forcing contributes strongly to reactivation phases more than in other landslide types.

Overall, WoE and GAMB differ in how they convert predictors into susceptibility. WoE amplifies discrete class contrasts (geology, land cover) and produces stepwise trends for slope and TWI. GAMB yields smoother, nonlinear responses with clear thresholds (e.g., slope ~30°–40°, TWI stabilising >12–16) and sharper curvature effects around flat–convex forms. These differences matter most for Slow Flow and Complex types: WoE stresses clay-rich units, concavity, and gentle slopes, but underplays how moisture and topographic convergence combine, whereas GAMB raises susceptibility where slope–hydrology–curvature interact. GAMB captures these multi-predictor interactions that WoE cannot, explaining much of the divergence between outputs. However, this capacity comes with reduced transparency, since the combined (interaction) between nonlinear predictors effects are difficult to visualize directly in a 2D PD plot. Thus, discrepancies reflect both the models' formulations (interaction-sensitive GAMB vs contrast-based WoE) and the intrinsically multifactorial nature of landslides, which GAMB represents more fully. This complex behaviour appears in the type-specific

susceptibility maps (S 2.1- S 2.5) and is then integrated into the overall products in Figure 2.7 via the complementary probability approach.

However, Figure 2.15 illustrates the overall distribution of susceptibility classes along the proposed railway line. While the two models differ in performance and their ability to capture the dominant landslide types, their outcomes along the railway corridor show only minor differences, suggesting a limited impact on first decision-stage of the planning phase in a practical susceptibility assessments.

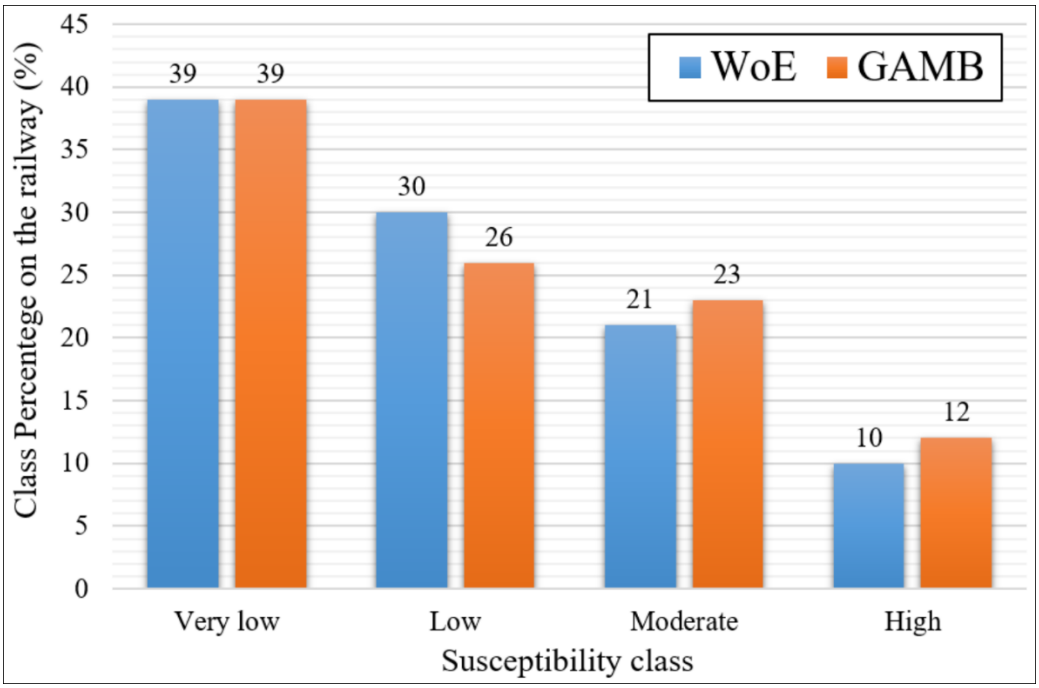


Figure 2.15. Percentages of the proposed railway exposed to each susceptibility classes.

In conclusion, we believe that the implementation of both and more models is recommended. In fact, some critical sectors that warrant particular attention might only be detected by one model, which may not be the one with the highest performance. Plots like those shown in Figures 2.13 and 2.14 are valuable for identifying such areas, regardless of overall model accuracy.

Long term susceptibility vs actual state of activity (EGMS)

The landslide susceptibility classes, ranging from very low to high, lack an intrinsic definition of hazard and represent an arbitrary stratification of a priori hazard levels. Consequently, the real-world significance of each susceptibility class in terms of actual risk or hazard remains speculative. This fundamental challenge of translating relative susceptibility classes into practical hazard implications has been largely overlooked in the literature. In particular, it is important to realize that a landslide susceptibility map, particularly when based on long-term landslide inventories, reflects the spatial propensity for landslide initiation, but does not capture the current state of activity of the slopes. Recently, Li et al. (2025) integrated this information into the analysis, but without comparing directly with the susceptibility results. To clarify this point, we compared the results of our susceptibility analysis with the actual ground velocity data provided by the European Ground Motion Service (EGMS; Costantini et al. 2021).

Specifically, we retrieved calibrated permanent scatterers (PSs) data (level 2B) covering our area of interest from 2015 to 2021, for both descending and ascending orbits, from the EGMS Copernicus website.

Our analysis then proceeded to explore two distinct types of relationships between PS velocity and susceptibility. Firstly, we sought to determine if a direct correlation exists between P_{ovr} and absolute PS velocities. This involved calculating the average absolute velocity of the PSs inside each landslide area together with the mean landslide susceptibility from the unique landslide-type maps. Subsequently, a scatter plot was generated, plotting mean velocities against mean susceptibility values, where each point represents an individual landslide area. The significance of the correlation was evaluated by fitting a second-order function and computing the corresponding R^2 value. Secondly, we examined the relationship between mean PS velocity and mean susceptibility of P_{ovr} within landslide-type areas, this time categorizing the data according to our four reclassified susceptibility classes. This analysis involved generating statistical data (summarized in a box plot) on the velocity within landslide areas for each class.

Given the nature of the PS data, which are able to read slowly to moderate ground deformations (less than 85.2 cm/yr from Crosetto et al 2021), we decided to restrict our approach only to landslide areas classified as Slow flow (SF-LS) and Complex (C-LS), to avoid misleading relationships. Finally, it is important to note that PS data tends to be concentrated in areas of human activity, introducing a bias towards regions where structures, often located in stable or deliberately stabilized zones, might not accurately reflect the expected velocity ranges in less interfered areas.

The comparison between EGMS velocity data (from 2015 to 2021) and landslide susceptibility aggregated per landslide area, is shown in Figure 2.16. As illustrated, no positive correlation emerges between displacement velocities and susceptibility values, highlighting a clear lack of relationship between the two. The R^2 values of the fitted second-order polynomial function are very close to zero, indicating no significant correlation.

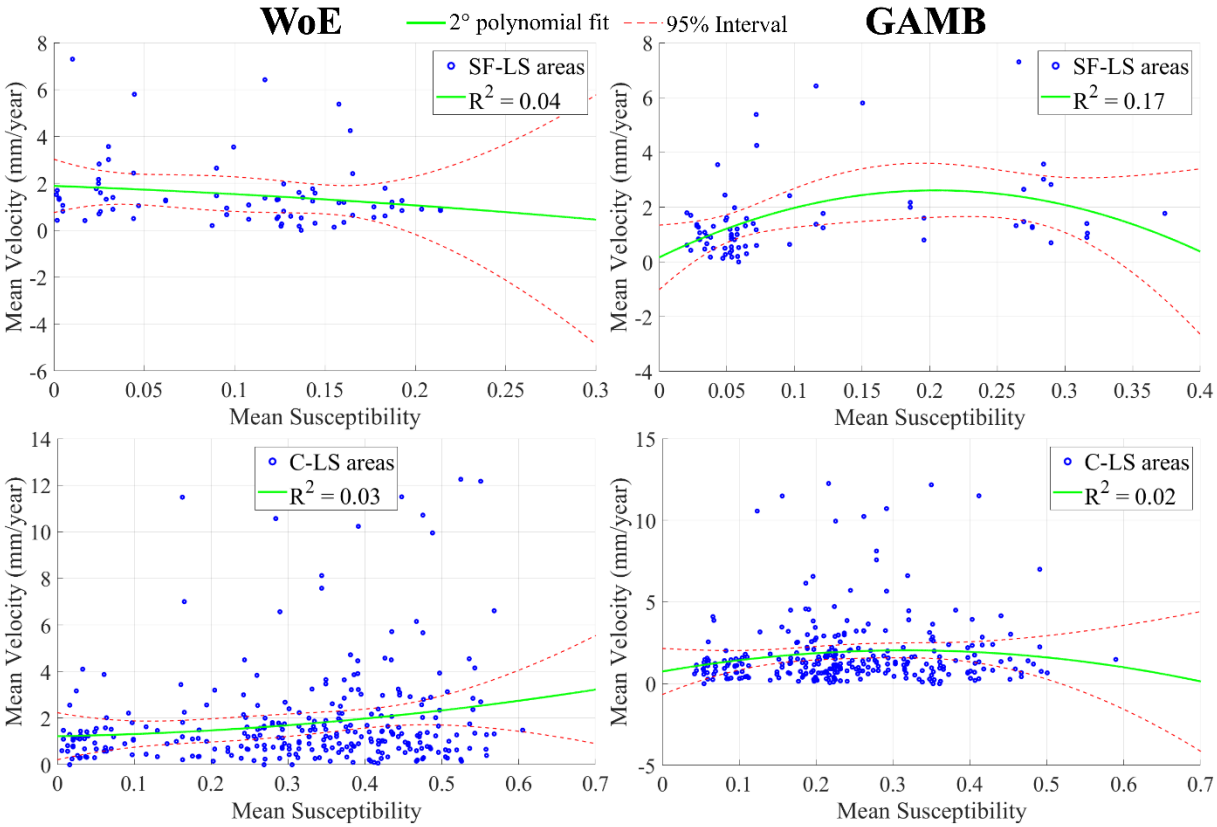


Figure 2.16. Correlations between mean velocity (mm/yr) and mean susceptibility in each slow flow (SF-LS, upper line) and complex (C-LS, bottom line) areas (blue dots), fitted by a second-order polynomial function (green line) with 95% confidence level interval (red dashed lines).

Furthermore, Figure 2.17 uses box plots to depict the relationship between EGMS velocities and susceptibility classes (see Par. 3.3). Also in this case, it is evident that there are no clear correlations between higher susceptibility classes and higher velocity values.

This comparison underscores the fundamental difference between long-term landslide susceptibility and actual ground displacement rates. Although this distinction is implicit in the analysis and well understood by specialists, it is often overlooked by stakeholders and therefore requires clear and explicit communication.

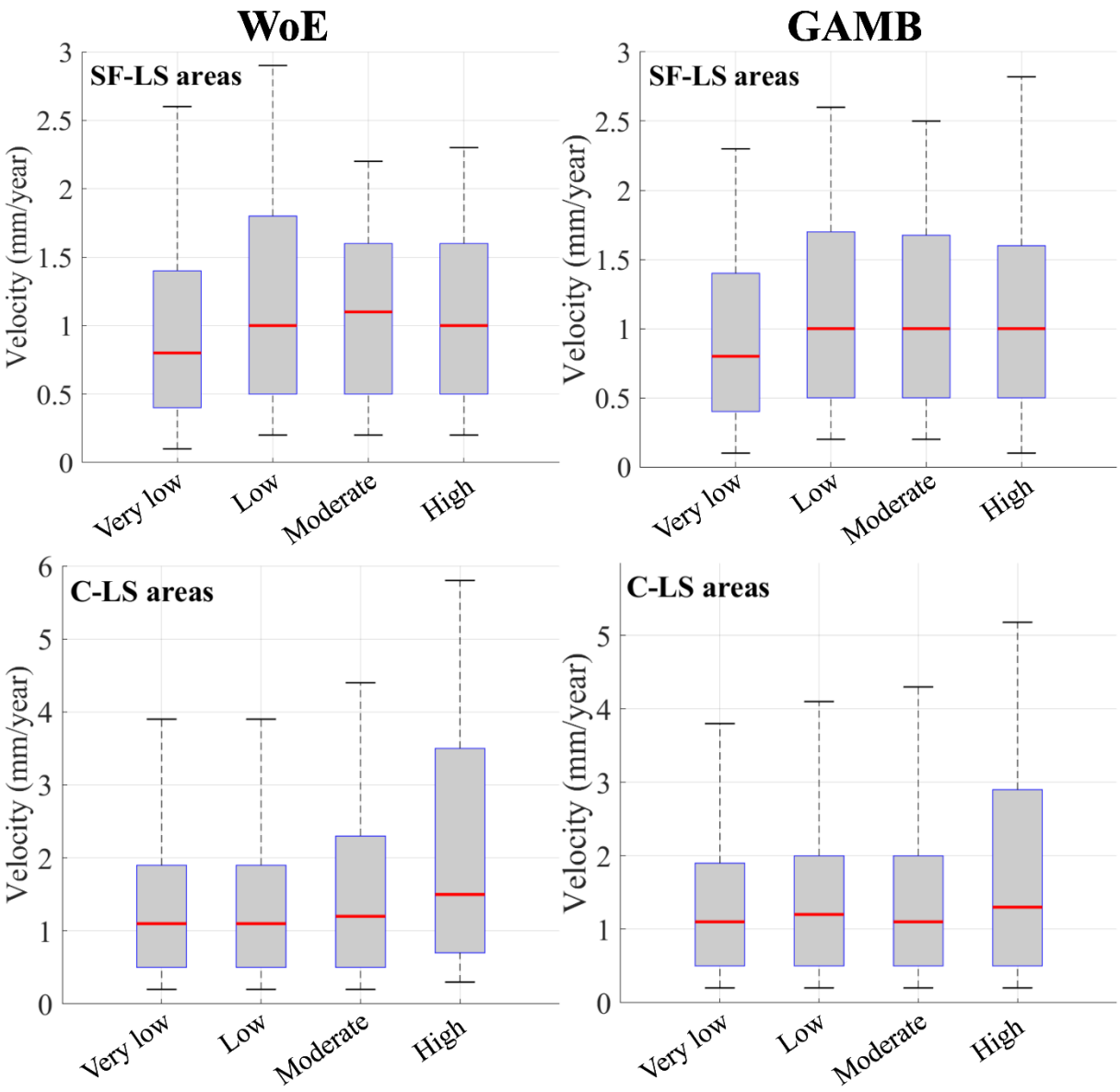


Figure 2.17. Box plot of the EGMS velocity values inside the slow flow (SF-LS) and complex (C-LS) areas grouped by susceptibility classes. Red-line is the median, box delineation at 25th and 75th percentile, lower and upper whisker at the 10th and 90th percentile.

In principle, any susceptibility model should be capable of identifying slopes where landslides are more likely to occur (Reichenbach et al., 2018). More specifically, these models detect areas that share similar characteristics, such as geology, geomorphology, and other relevant factors, with the slopes included in the training dataset. This approach is grounded in the principle of actualism, which holds that natural processes observed in the past will continue to operate in the future under comparable conditions (Aleotti & Chowdhury 1999). Consequently, the features of slopes that failed in the past serve as valuable indicators of potential future instability. Within this framework, the timescale of the analysis plays a critical role. When the training dataset reflects long-term landslide activity, as in our case, the resulting susceptibility map indicates the general predisposition of slopes to failure, rather than their current state of activity or immediate instability.

Limitations, improvements and practical implications

Our methodology tried to bring the landslide susceptibility/probability concepts in a real world applications scenario. One concern that is still tricky is the starting landslide inventory quality and completeness. A dataset that reflects the geomorphological landscape, encompassing a wide timescale in which landslides were formed, like ours, will produce results that are cautious but at the same time adaptable for different future scenarios. However, the main drawback is that these datasets can also be less detailed and sometimes subjective compared to event-based datasets. The interpretation of old geomorphological features, in particular, may lead to incorrectly mapped landslides, and so to less reliable susceptibility maps.

Regarding the delineation of susceptibility classes, we believe that the definition of relative thresholds per geological domain is necessary to accurately interpret landslide hazard and to effectively identify the most critical areas. In our case, for instance, applying uniform thresholds across the entire study area would cause domain B to be overrepresented in the higher susceptibility classes, thereby distorting the scale and masking critical areas in domain A. This approach also allows us to consider separately the different types of landslides, that should be analyzed separately given their differing processes, intensities, and velocities.

Among the methods tested for classifying susceptibility values, AUC optimization produced thresholds that were more balanced, neither excessively conservative nor overly permissive, making it potentially more appropriate for our context. Although this cannot be considered a universal solution, as it may not perform equally well in other regions, by integrating multiple methods into an ensemble approach, the resulting susceptibility classes benefit from a broader statistical basis. This reduces reliance on any single technique and improves robustness, making the final classification more reliable and less sensitive to the methodological choice.

Recommendations for future applications of our methodology include integrating additional models such as XGBoost, Support Vector Machines, and others. Similarly, expanding the ensemble of reclassification methods by incorporating approaches (e.g. Generalized Youden Index, Head/Tail breaks, or Geometrical Intervals) could further enhance robustness. As demonstrated in our study, different models and methods capture unique aspects of the territory, and their combined use can lead to more reliable and comprehensive susceptibility assessments.

2.8 Conclusion

Landslides pose a significant threat to infrastructure, with railways especially vulnerable due to their rigidity and linear extent. Integrating landslide susceptibility analysis during the feasibility phase of railway projects is therefore crucial for identifying high-risk areas and informing mitigation strategies. This study proposes a methodology tailored for railway planning, applied to a landslide-prone area in the Marche region (Italy), where a new railway line is under development.

The study area spans two distinct geological domains: Domain A (older marly and calcareous formations) and Domain B (younger clayey and arenaceous formations). We employed two susceptibility models (WoE and GAMB) using separate inventories for five landslide types. The resulting maps were merged into an overall susceptibility map via a complementary

probability approach, from which we also derived type-specific sensitivity. To enhance interpretability, we introduced an ensemble reclassification method combining six techniques.

The resulting maps of reclassified overall probability with an ensemble of methods and ranked sensitivity of each landslide type are reliable and easily consultable even by non-experts. The very low susceptibility class appropriately covers valley areas, while the high susceptibility class predominantly encompasses regions within Domain B correlated to Complex landslide type using WoE and to Complex and Slow Flow using GAMB; fewer areas within Domain A are highly susceptible and are reasonably associated with Fall/Topple and Slide types (in the carbonate ridge). The GAMB model consistently outperformed WoE in terms of prediction accuracy, showing higher AUC values (the best for Fall/Topple = 0.978 and the worst for Slide = 0.784 against the respective WoE values of Fall/Topple = 0.941 and Slide = 0.699) and smoother, more coherent susceptibility maps across all landslide types. The discrepancies between WoE and GAMB susceptibility results, particularly for Slow Flow and Complex landslides, reflect the ability of GAMB to capture nonlinear interactions among predictors.

At the same time, the WoE model achieved a more reliable landslide-type identification with the proposed sensitivity method, reaching an overall accuracy among different landslide types of 0.62 against 0.34 for GAMB. The difference is evident for Slow Flow and Complex in Domain B, although the two landslide types share common geomorphological features. In the reclassification step, Domain A shows generally less agreement between methods, highlighting more uncertainties in the area compared to Domain B. The GAMB model generally demonstrates better overall consistency compared to WoE, which indicates better reliability in practical applications. The proposed ensemble reclassification method naturally dampens extremes and avoids over-optimistic or over-pessimistic labels by assigning the model class across techniques. Among individual approaches, the AUC-optimisation reclassification consistently sits in the mid-range of class assignments, making it the most balanced option for our study area.

Final outputs of overall susceptibility plots along the railway route, supported by the distinct thresholds identified for each domain, highlight consistent yet location-dependent differences between WoE and GAMB. The landslide type-coloured profiles and the grey band of mapped intersections provide an effective, compact way to read model and inventory agreement along the line. The comparison with the actual railway–landslide intersections gives overall accuracy of 0.797 (WoE) and 0.837 (GAMB), suitable for corridor screening; and False Positive segments with relative expected landslide type are valuable as newly identified high-susceptibility stretches to prioritise. Moreover, comparative plots for both sides of the track (average susceptibility within 100 m buffers) indicate greater exposure on the right in both models and smaller/less critical portions on the left. Overall, the combined visualisations and metrics show that, while the models differ in emphasis and type attribution, together they exhibit similar percentages of susceptibility classes along the railway and thus provide complementary guidance for screening, investigation, and early-stage design decisions.

We also compared the results of our susceptibility analysis with the actual ground-velocity data provided by EGMS, focusing on Complex and Slow Flow landslides. The analysis concludes

that there is no correlation between susceptibility values and EGMS velocity data ($R^2 < 0.17$ for both Slow Flow and Complex), underlining the different meanings of long-term susceptibility and short-term activity of slopes. the resulting susceptibility map indicates the general predisposition of slopes to failure, rather than their current state of activity

In conclusion, this study demonstrates that, by accounting for geological differences and inventory limitations, susceptibility methods can provide a practical and scalable framework for improving railway design in landslide-prone regions.

2.9 References chapter 2

Aleotti, P., & Chowdhury, R. (1999). Landslide hazard assessment: Summary review and new perspectives. Bulletin of Engineering Geology and the Environment, 58, 21–44. https://doi.org/10.1007/s100640050066

Alvioli, M., Loche, M., Jacobs, L., Grohmann, C. H., Abraham, M. T., Gupta, K., Satyam, N., Scarigni, G., Bornaetxea, T., Rossi, M., Marchesini, I., Lombardo, L., Moreno, M., Steger, S., Camera, C. A. S., Bajni, G., Samodra, G., Wahyudi, E. E., Susyanto, N., Sinčić, M., Gazibara, S. B., Sirbu, F., Torzin, J., Schüßler, N., Mirus, B. B., Woodard, J. B., Aquilera, H., & Rivera-Rivera, J. S. (2024). A benchmark dataset and workflow for landslide susceptibility zonation. Earth-Science Reviews, 104927. https://doi.org/10.1016/j.earscirev.2024.104927

Ardizzone, F., Cardinali, M., Carrara, A., Guzzetti, F., & Reichenbach, P. (2002). Impact of mapping errors on the reliability of landslide hazard maps. Natural Hazards and Earth System Sciences, 2(1/2), 3–14. https://doi.org/10.5194/nhess-2-3-2002

Arslan, A., Tari, E., Ziatdinov, R., & Nabiyev, R. I. (2014). Transition curve modeling with kinematical properties: Research on log-aesthetic curves. Computer-Aided Design and Applications, 11(5), 509–517. https://doi.org/10.1080/16864360.2014.902680

Atkinson, P. M., & Massari, R. (1998). Generalised linear modelling of susceptibility to landsliding in the central Apennines, Italy. Computers & Geosciences, 24(4), 373–385. https://doi.org/10.1016/S0098-3004(97)00117-9

Ayalew, L., & Yamagishi, H. (2005). The application of GIS-based logistic regression for landslide susceptibility mapping in the Kakuda-Yahiko Mountains, Central Japan. Geomorphology, 65(1–2), 15–31. https://doi.org/10.1016/j.geomorph.2004.06.010

Ba, Q., Chen, Y., Deng, S., Yang, J., & Li, H. (2018). A comparison of slope units and grid cells as mapping units for landslide susceptibility assessment. Earth Science Informatics, 11, 373–388. https://doi.org/10.1007/s12145-017-0316-1

Bednarik, M., Magulová, B., Matys, M., & Marschalko, M. (2010). Landslide susceptibility assessment of the Kraľovany–Liptovský Mikuláš railway case study. Physics and Chemistry of the Earth, Parts A/B/C, 35(3–5), 162–171. https://doi.org/10.1016/j.pce.2009.12.001

Beven, K. J., & Kirkby, M. J. (1979). A physically based, variable contributing area model of basin hydrology. Hydrological Sciences Journal, 24(1), 43–69. https://doi.org/10.1080/02626667909491834

Bonham-Carter, G. (1994). Geographic information systems for geoscientists: Modelling with GIS (Vol. 13). Elsevier.

Bordoni, M., Persichillo, M. G., Meisina, C., Crema, S., Cavalli, M., Bartelletti, C., Galanti, Y., Barsanti, M., Giannecchini, R., & D'Amato Avanzi, G. (2018). Estimation of the susceptibility of a road network to shallow landslides with the integration of the sediment connectivity. Natural Hazards and Earth System Sciences, 18(6), 1735-1758. https://doi.org/10.5194/nhess-18-1735-2018, 2018

Bozzolan, E., Holcombe, E. A., Pianosi, F., Marchesini, I., Alvioli, M., & Wagener, T. (2023). A mechanistic approach to include climate change and unplanned urban sprawl in landslide susceptibility maps. Science of The Total Environment, 858, 159412. https://doi.org/10.1016/j.scitotenv.2022.159412

Brenning, A. (2005). Spatial prediction models for landslide hazards: Review, comparison and evaluation. Natural Hazards and Earth System Sciences, 5(6), 853–862. https://doi.org/10.5194/nhess-5-853-2005

Budimir, M. E. A., Atkinson, P. M., & Lewis, H. G. (2015). A systematic review of landslide probability mapping using logistic regression. Landslides, 12, 419-436. https://doi.org/10.1007/s10346-014-0550-5

Cantarino, I., Carrion, M. A., Goerlich, F., & Martinez Ibañez, V. (2019). A ROC analysis-based classification method for landslide susceptibility maps. Landslides, 16(2), 265–282. https://doi.org/10.1007/s10346-018-1071-6

Cardellini, S., Galdenzi, S., Guerrera, F., Pennacchioni, E., Piergiovanni, A., Pieruccini, P., Sandroni, P., Tosti, S., & Tramontana, M. (2009). Carta Geologica d'Italia alla scala 1: 50.000 Foglio 292 Jesi. https://www.isprambiente.gov.it/Media/carg/note_illustrative/292_Jesi.pdf

Carrara, A. (1983). Multivariate models for landslide hazard evaluation. Journal of the International Association for Mathematical Geology, 15(3), 403–426. https://doi.org/10.1007/BF01031290

Carrara, A., Crosta, G., & Frattini, P. (2008). Comparing models of debris-flow susceptibility in the alpine environment. Geomorphology, 94(3–4), 353–378. https://doi.org/10.1016/j.geomorph.2006.10.033

Catani, F., Lagomarsino, D., Segoni, S., & Tofani, V. (2013). Landslide susceptibility estimation by random forests technique: sensitivity and scaling issues. Natural Hazards and Earth System Sciences, 13(11), 2815–2831. https://doi.org/10.5194/nhess-13-2815-2013

Centamore, E., Chiocchini, M., Chiocchini, U., Dramis, F., Giardini, G., Jacobacci, A., Martelli, G., Micarelli, A., & Potetti, M. (1979). Note illustrative del Foglio 301 "Fabriano" alla scala 1:50.000. Serv. Geol. d'It.: pp. 51, Roma. https://www.isprambiente.gov.it/Media/carg/note_illustrative/301_Fabriano.pdf

Chen, W., Xie, X., Wang, J., Pradhan, B., Hong, H., Bui, D. T., Duan, Z., & Ma, J. (2017). A comparative study of logistic model tree, random forest, and classification and regression tree models for spatial prediction of landslide susceptibility. Catena, 151, 147–160. https://doi.org/10.1016/j.catena.2016.11.032

Chen, W., Peng, J., Hong, H., Shahabi, H., Pradhan, B., Liu, J., Zhu, A., Pei, X., & Duan, Z. (2018). Landslide susceptibility modelling using GIS-based machine learning techniques for Chongren County, Jiangxi Province, China. Science of the Total Environment, 626, 1121–1135. https://doi.org/10.1016/j.scitotenv.2018.01.160

Chester, M., & Horvath, A. (2010). Life-cycle assessment of high-speed rail: The case of California. Environmental Research Letters, 5(1), 014003. https://doi.org/10.1088/1748-9326/5/1/014003

Childs, C. (2004). Interpolating surfaces in ArcGIS spatial analyst. ArcUser, July–September, 32–35.

Chung, C. J. F., & Fabbri, A. G. (2003). Validation of spatial prediction models for landslide hazard mapping. Natural Hazards, 30(3), 451–472. https://doi.org/10.1023/B:NHAS.0000007172.62651.2b

Coltorti, M., Farabollini, P., Gentili, B., & Pambianchi, G. (1996). Geomorphological evidence for anti-Apennine faults in the Umbro-Marchean Apennines and in the peri-Adriatic basin, Italy. Geomorphology, 15(1), 33–45. https://doi.org/10.1016/0169-555X(95)00117-N

Costantini, M., Minati, F., Trillo, F., Ferretti, A., Novali, F., Passera, E., Dehls, J., Larsen, Y., Marinkovic, P., Eineder, M., Brcic, R., Siegmund, R., Kotzerke, P., Probeck, M., Kenyeres, A., Proietti, S., Solari, L., & Andersen, H. S. (2021, July). European ground motion service (EGMS). In 2021 IEEE international geoscience and remote sensing symposium IGARSS (pp. 3293-3296). IEEE. 10.1109/IGARSS47720.2021.9553562

Crosetto, M., Solari, L., Balasis-Levinsen, J., Bateson, L., Casagli, N., Frei, M., Oyen, A., Moldestad, D. A., & Mróz, M. (2021). Deformation monitoring at European scale: The Copernicus ground motion service. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XLIII-B3-2021, 141–146. https://doi.org/10.5194/isprs-archives-XLIII-B3-2021-141-2021

Cruden, D. M. (1996). Landslide types and processes. Transportation Research Board Special Report, 247, 36–75.

Dahal, A., Tanyas, H., van Westen, C., van der Meijde, M., Mai, P. M., Huser, R., & Lombardo, L. (2024). Space–time landslide hazard modeling via Ensemble Neural Networks. Natural Hazards and Earth System Sciences, 24(3), 823–845. https://doi.org/10.5194/nhess-24-823-2024

De Kemp, E. A. (1998). Three-dimensional projection of curvilinear geological features through direction cosine interpolation of structural field observations. Computers & Geosciences, 24(3), 269–284. https://doi.org/10.1016/S0098-3004(97)00066-6

Dou, H. Q., Huang, S. Y., Jian, W. B., & Wang, H. (2023). Landslide susceptibility mapping of mountain roads based on machine learning combined model. Journal of Mountain Science, 20(5), 1232-1248. https://doi.org/10.1007/s11629-022-7657-2

Fang, Z., Wang, Y., van Westen, C., & Lombardo, L. (2024). Space–time landslide susceptibility modeling based on data-driven methods. Mathematical Geosciences, 56(6), 1335–1354. https://doi.org/10.1007/s11004-023-10105-6

Fell, R., Corominas, J., Bonnard, C., Cascini, L., Leroi, E., & Savage, W. Z. (2008). Guidelines for landslide susceptibility, hazard and risk zoning for land use planning. Engineering Geology, 102(3–4), 85–98. https://doi.org/10.1016/j.enggeo.2008.03.022

Frattini, P., Crosta, G., & Carrara, A. (2010). Techniques for evaluating the performance of landslide susceptibility models. Engineering Geology, 111(1–4), 62–72. https://doi.org/10.1016/j.enggeo.2009.12.004

Friedman, J. H. (2001). Greedy function approximation: a gradient boosting machine. Annals of statistics, 1189-1232. https://doi.org/10.1214/aos/1013203451

Froude, M. J., & Petley, D. N. (2018). Global fatal landslide occurrence from 2004 to 2016. Natural Hazards and Earth System Sciences, 18(8), 2161–2181. https://doi.org/10.5194/nhess-18-2161-2018

Goetz, J. N., Guthrie, R. H., & Brenning, A. (2011). Integrating physical and empirical landslide susceptibility models using generalized additive models. Geomorphology, 129(3–4), 376–386. https://doi.org/10.1016/j.geomorph.2011.03.001

Grelle, G., Revellino, P., Donnarumma, A., & Guadagno, F. M. (2011). Bedding control on landslides: A methodological approach for computer-aided mapping analysis. Natural Hazards and Earth System Sciences, 11(5), 1395–1409. https://doi.org/10.5194/nhess-11-1395-2011

Guinau, M., Vilajosana, I., & Vilaplana, J. M. (2007). GIS-based debris flow source and runout susceptibility assessment from DEM data – A case study in NW Nicaragua. Natural Hazards and Earth System Sciences, 7(6), 703–716. https://doi.org/10.5194/nhess-7-703-2007

Guo, Z., Tian, B., Zhu, Y., He, J., & Zhang, T. (2024). How do the landslide and non-landslide sampling strategies impact landslide susceptibility assessment?—A catchment-scale case study from China. Journal of Rock Mechanics and Geotechnical Engineering, 16(3), 877-894. https://doi.org/10.1016/j.jrmge.2023.07.026

Guo, Z., Zeng, T., Zhang, Y., Yu, W., Wang, L., Guo, Z., & Glade, T. (2025). A novel hybrid model integrating high resolution remote sensing and stacking ensemble techniques for landslide susceptibility mapping: Application to event-based landslide inventory. Geomorphology, 109886. https://doi.org/10.1016/j.geomorph.2025.109886

Guzzetti, F., Cardinali, M., & Reichenbach, P. (1996). The influence of structural setting and lithology on landslide type and pattern. Environmental & Engineering Geoscience, 2(4), 531–555. https://doi.org/10.2113/gseegeosci.2.4.531

Guzzetti, F., Carrara, A., Cardinali, M., & Reichenbach, P. (1999). Landslide hazard evaluation: A review of current techniques and their application in a multi-scale study, Central Italy. Geomorphology, 31(1–4), 181–216. https://doi.org/10.1016/S0169-555X(99)00078-1

Guzzetti, F., Galli, M., Reichenbach, P., Ardizzone, F., & Cardinali, M. (2006). Landslide hazard assessment in the Collazzone area, Umbria, Central Italy. Natural Hazards and Earth System Sciences, 6(1), 115–131. https://doi.org/10.5194/nhess-6-115-2006

Hanley, J. A., & McNeil, B. J. (1982). The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology, 143(1), 29-36. https://doi.org/10.1148/radiology.143.1.7063747

Hastie, T., & Tibshirani, R. (1990). Exploring the nature of covariate effects in the proportional hazards model. Biometrics, 46(4), 1005–1016. https://doi.org/10.2307/2532444

Jacobs, L., Kervyn, M., Reichenbach, P., Rossi, M., Marchesini, I., Alvioli, M., & Dewitte, O. (2020). Regional susceptibility assessments with heterogeneous landslide information: Slope unit- vs. pixel-based approach. Geomorphology, 356, 107084. https://doi.org/10.1016/j.geomorph.2020.107084

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). An introduction to statistical learning (Vol. 112, p. 18). New York: Springer. https://doi.org/10.1007/978-1-4614-7138-7

Kavzoglu, T., Sahin, E. K., & Colkesen, I. (2014). Landslide susceptibility mapping using GIS-based multi-criteria decision analysis, support vector machines, and logistic regression. Landslides, 11, 425–439. https://doi.org/10.1007/s10346-013-0391-7

Li, J., Zhou, Z., & Ma, W. (2025). Assessment of landslide susceptibility along the Lanzhou-xinjiang high-speed railway: A case study of Menyuan-Shandanmachang. Transportation Geotechnics, 50, 101473. https://doi.org/10.1016/j.trgeo.2024.101473

Lima, P., Steger, S., Glade, T., & Mergili, M. (2023). Conventional data-driven landslide susceptibility models may only tell us half of the story: potential underestimation of landslide

impact areas depending on the modeling design. Geomorphology, 430(108638), 10-1016. https://doi.org/10.1016/j.geomorph.2023.108638

Liu, K., Wang, M., Cao, Y., Zhu, W., & Yang, G. (2018). Susceptibility of existing and planned Chinese railway systems subjected to rainfall-induced multi-hazards. Transportation Research Part A: Policy and Practice, 117, 214–226. https://doi.org/10.1016/j.tra.2018.08.030

Loche, M., Alvioli, M., Marchesini, I., Bakka, H., & Lombardo, L. (2022). Landslide susceptibility maps of Italy: Lesson learnt from dealing with multiple landslide types and the uneven spatial distribution of the national inventory. Earth-Science Reviews, 232, 104125. https://doi.org/10.1016/j.earscirev.2022.104125

Lombardo, L., Cama, M., Conoscenti, C., Märker, M., & Rotigliano, E. (2015). Binary logistic regression versus stochastic gradient boosted decision trees in assessing landslide susceptibility for multiple-occurring landslide events: Application to the 2009 storm event in Messina (Sicily, Southern Italy). Natural Hazards, 79, 1621–1648. https://doi.org/10.1007/s11069-015-1915-8

Lombardo, L., & Mai, P. M. (2018). Presenting logistic regression-based landslide susceptibility results. Engineering Geology, 244, 14–24. https://doi.org/10.1016/j.enggeo.2018.07.019

Lombardo, L., Opitz, T., Ardizzone, F., Guzzetti, F., & Huser, R. (2020). Space-time landslide predictive modeling. Earth-Science Reviews, 209, 103318. https://doi.org/10.1016/j.earscirev.2020.103318

Lou, Y., Caruana, R., Gehrke, J., & Hooker, G. (2013). Accurate intelligible models with pairwise interactions. In Proceedings of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (pp. 623–631). https://doi.org/10.1145/2487575.2487579

Marchesini, I., Althuwaynee, O., Santangelo, M., Alvioli, M., Cardinali, M., Mergili, M., Reichenbach, P., Perucci, S., Balducci, V., Agostino, I., Esposito, R., & Rossi, M. (2024). National-scale assessment of railways exposure to rapid flow-like landslides. Engineering Geology, 332, 107474. https://doi.org/10.1016/j.enggeo.2023.107474

Martinello, C., Cappadonia, C., Conoscenti, C., Agnesi, V., & Rotigliano, E. (2021). Optimal slope units partitioning in landslide susceptibility mapping. Journal of Maps, 17(3), 152–162. https://doi.org/10.1080/17445647.2021.1916790

Martinello, C., Cappadonia, C., Conoscenti, C., & Rotigliano, E. (2022). Landform classification: A high-performing mapping unit partitioning tool for landslide susceptibility assessment—A test in the Imera River basin (Northern Sicily, Italy). Landslides. https://doi.org/10.1007/s10346-022-01870-6

Meentemeyer, R. K., & Moody, A. (2000). Automated mapping of conformity between topographic and geological surfaces. Computers & Geosciences, 26(7), 815–829. https://doi.org/10.1016/S0098-3004(99)00126-2

Merghadi, A., Yunus, A. P., Dou, J., Whiteley, J., ThaiPham, B., Bui, D. T., Avtar, R., & Abderrahmane, B. (2020). Machine learning methods for landslide susceptibility studies: A comparative overview of algorithm performance. Earth-Science Reviews, 207, 103225. https://doi.org/10.1016/j.earscirev.2020.103225

Meyer, N. K., Schwanghart, W., Korup, O., Romstad, B., & Etzelmüller, B. (2014). Estimating the topographic predictability of debris flows. Geomorphology, 207, 114–125. https://doi.org/10.1016/j.geomorph.2013.10.030

Moreno, M., Lombardo, L., Crespi, A., Zellner, P. J., Mair, V., Pittore, M., van Westen, C. J., & Steger, S. (2024). Space-time data-driven modeling of precipitation-induced shallow landslides in South Tyrol, Italy. Science of the Total Environment, 912, 169166. https://doi.org/10.1016/j.scitotenv.2023.169166

Nocentini, N., Rosi, A., Piciullo, L., Liu, Z., Segoni, S., & Fanti, R. (2024). Regional-scale spatiotemporal landslide probability assessment through machine learning and potential applications for operational warning systems: a case study in Kvam (Norway). Landslides, 21(10), 2369-2387. https://doi.org/10.1007/s10346-024-02287-9

Ori, G. G., Serafini, G., Visentin, C., Ricci Lucchi, F., Casnedi, R., Colalongo, M. L., & Mosna, S. (1991). The Pliocene-Pleistocene Adriatic Foredeep (Marche and Abruzzo, Italy): an integrated approach to surface and subsurface geology.

Pasang, S., & Kubíček, P. (2020). Landslide susceptibility mapping using statistical methods along the Asian Highway, Bhutan. Geosciences, 10(11), 430. https://doi.org/10.3390/geosciences10110430

Peng, L., Niu, R., Huang, B., Wu, X., Zhao, Y., & Ye, R. (2014). Landslide susceptibility mapping based on rough set theory and support vector machines: A case of the Three Gorges area, China. Geomorphology, 204, 287–301. https://doi.org/10.1016/j.geomorph.2013.08.013

Petschko, H., Brenning, A., Bell, R., Goetz, J., & Glade, T. (2014). Assessing the quality of landslide susceptibility maps – Case study Lower Austria. Natural Hazards and Earth System Sciences, 14(1), 95–118. https://doi.org/10.5194/nhess-14-95-2014

Pourghasemi, H. R., & Rossi, M. (2017). Landslide susceptibility modeling in a landslide-prone area in Mazandaran Province, north of Iran: A comparison between GLM, GAM, MARS, and M-AHP methods. Theoretical and Applied Climatology, 130(1), 609–633. https://doi.org/10.1007/s00704-016-1919-5

Quinn, P. E., Hutchinson, D. J., Diederichs, M. S., & Rowe, R. K. (2010). Regional-scale landslide susceptibility mapping using the weights of evidence method: An example applied to linear infrastructure. Canadian Geotechnical Journal, 47(8), 905–927. https://doi.org/10.1139/T10-019

Rényi, A. (2007). Probability theory. Courier Corporation. https://doi.org/10.1007/978-94-010-1795-4

Reichenbach, P., Rossi, M., Malamud, B. D., Mihir, M., & Guzzetti, F. (2018). A review of statistically-based landslide susceptibility models. Earth-Science Reviews, 180, 60–91. https://doi.org/10.1016/j.earscirev.2018.03.001

Ricci-Lucchi, F. (1975). Depositional cycles in two turbidite formations of northern Apennines. Journal of Sedimentary Research, 45(1), 3–43. https://doi.org/10.1306/212F6CB7-2B24-11D7-8648000102C1865D

Rosi, A., Frodella, W., Nocentini, N., Caleca, F., Havenith, H. B., Strom, A., Saidov, M., Bimurzaev, G. A., & Tofani, V. (2023). Comprehensive landslide susceptibility map of Central Asia. Natural Hazards and Earth System Sciences, 23(6), 2229–2250. https://doi.org/10.5194/nhess-23-2229-2023

Rossi, M., Guzzetti, F., Reichenbach, P., Mondini, A. C., & Peruccacci, S. (2010). Optimal landslide susceptibility zonation based on multiple forecasts. Geomorphology, 114(3), 129–142. https://doi.org/10.1016/j.geomorph.2009.06.020

Santangelo, M., Marchesini, I., Cardinali, M., Fiorucci, F., Rossi, M., Bucci, F., & Guzzetti, F. (2015). A method for the assessment of the influence of bedding on landslide abundance and types. Landslides, 12, 295–309. https://doi.org/10.1007/s10346-014-0475-1

Schafer, D. H., & Barkan, C. P. L. (2008). A prediction model for broken rails and an analysis of their economic impact. In Proceedings of the American Railway Engineering and Maintenance-of-Way Association (AREMA) Annual Conference.

Schlögl, M., & Matulla, C. (2018). Potential future exposure of European land transport infrastructure to rainfall-induced landslides throughout the 21st century. Natural Hazards and Earth System Sciences, 18(4), 1121–1132. https://doi.org/10.5194/nhess-18-1121-2018

Schlögl, M., Spiekermann, R., & Steger, S. (2025). Towards a holistic assessment of landslide susceptibility models: insights from the Central Eastern Alps. Environmental Earth Sciences, 84(4), 113. https://doi.org/10.1007/s12665-024-12041-y

Sciarra, M., Coco, L., & Urbano, T. (2017). Assessment and validation of GIS-based landslide susceptibility maps: A case study from Feltrino stream basin (Central Italy). Bulletin of Engineering Geology and the Environment, 76(2), 437–456. https://doi.org/10.1007/s10064-016-0954-7

Simoni, S., Zanotti, F., Bertoldi, G., & Rigon, R. (2008). Modelling the probability of occurrence of shallow landslides and channelized debris flows using GEOtop-FS. Hydrological Processes, 22(4), 532–545. https://doi.org/10.1002/hyp.6886

Sinčić, M., Bernat Gazibara, S., Krkač, M., & Mihalić Arbanas, S. (2022). Landslide susceptibility assessment of the City of Karlovac using the bivariate statistical analysis. Rudarsko-geološko-naftni zbornik, 37(2), 149–170. https://doi.org/10.17794/rgn.2022.2.13

Steger, S., Brenning, A., Bell, R., Petschko, H., & Glade, T. (2016). Exploring discrepancies between quantitative validation results and the geomorphic plausibility of statistical landslide susceptibility maps. Geomorphology, 262, 8–23. https://doi.org/10.1016/j.geomorph.2016.03.005

Steger, S., Moreno, M., Crespi, A., Gariano, S. L., Brunetti, M. T., Melillo, M., Peracucci, S., Marra, F., de Vugt, L., Zieher, T., Rutzinger, M., Mair, V., & Pittore, M. (2024). Adopting the margin of stability for space–time landslide prediction–A data-driven approach for generating spatial dynamic thresholds. Geoscience Frontiers, 15(5), 101822. https://doi.org/10.1016/j.gsf.2023.101822

Sujatha, E. R. (2021). An integrated landslide susceptibility model to assess landslides along linear infrastructure for environmental management. Environmental Earth Sciences, 80(12), 447. https://doi.org/10.1007/s12665-021-09747-8

Sur, U., Singh, P., & Meena, S. R. (2020). Landslide susceptibility assessment in a lesser Himalayan road corridor (India) applying fuzzy AHP technique and earth-observation data. Geomatics, Natural Hazards and Risk, 11(1), 2176–2209. https://doi.org/10.1080/19475705.2020.1836038

Taalab, K., Cheng, T., & Zhang, Y. (2018). Mapping landslide susceptibility and types using Random Forest. Big Earth Data, 2(2), 159–178. https://doi.org/10.1080/20964471.2018.1472392

Tarquini, S., Isola, I., Favalli, M., Mazzarini, F., Bisson, M., Pareschi, M. T., & Boschi, E. (2007). TINITALY/01: A new Triangular Irregular Network of Italy. Annals of Geophysics, 50(3), 407–425. https://doi.org/10.4401/ag-4424

Tarquini, S., Vinci, S., Favalli, M., Doumaz, F., Fornaciai, A., & Nannipieri, L. (2012). Release of a 10-m-resolution DEM for the Italian territory: Comparison with global-coverage DEMs and anaglyph-mode exploration via the web. Computers & geosciences, 38(1), 168-170. https://doi.org/10.1016/j.cageo.2011.04.018

Titti, G., van Westen, C., Borgatti, L., Pasuto, A., & Lombardo, L. (2021). When enough is really enough? On the minimum number of landslides to build reliable susceptibility models. Geosciences, 11(11), 469. https://doi.org/10.3390/geosciences11110469

Titti, G., Sarretta, A., Lombardo, L., Crema, S., Pasuto, A., & Borgatti, L. (2022). Mapping susceptibility with open-source tools: A new plugin for QGIS. Frontiers in Earth Science, 10, 842425. https://doi.org/10.3389/feart.2022.842425

Trigila, A., Iadanza, C., & Spizzichino, D. (2010). Quality assessment of the Italian Landslide Inventory using GIS processing. Landslides, 7(4), 455–470. https://doi.org/10.1007/s10346-010-0213-0

Trigila, A., Iadanza, C., Esposito, C., & Scarascia-Mugnozza, G. (2015). Comparison of Logistic Regression and Random Forests techniques for shallow landslide susceptibility assessment in Giampilieri (NE Sicily, Italy). Geomorphology, 249, 119–136. https://doi.org/10.1016/j.geomorph.2015.06.001

Van Rijsbergen, C. J. (1979). Information retrieval: theory and practice. In Proceedings of the Joint IBM/University of Newcastle upon Tyne Seminar on Data Base Systems (Vol. 79, pp. 1–14).

Van Westen, C. J., Rengers, N., & Soeters, R. (2003). Use of geomorphological information in indirect landslide susceptibility assessment. Natural Hazards, 30(3), 399–419. https://doi.org/10.1023/B:NHAZ.0000007097.42735.9e

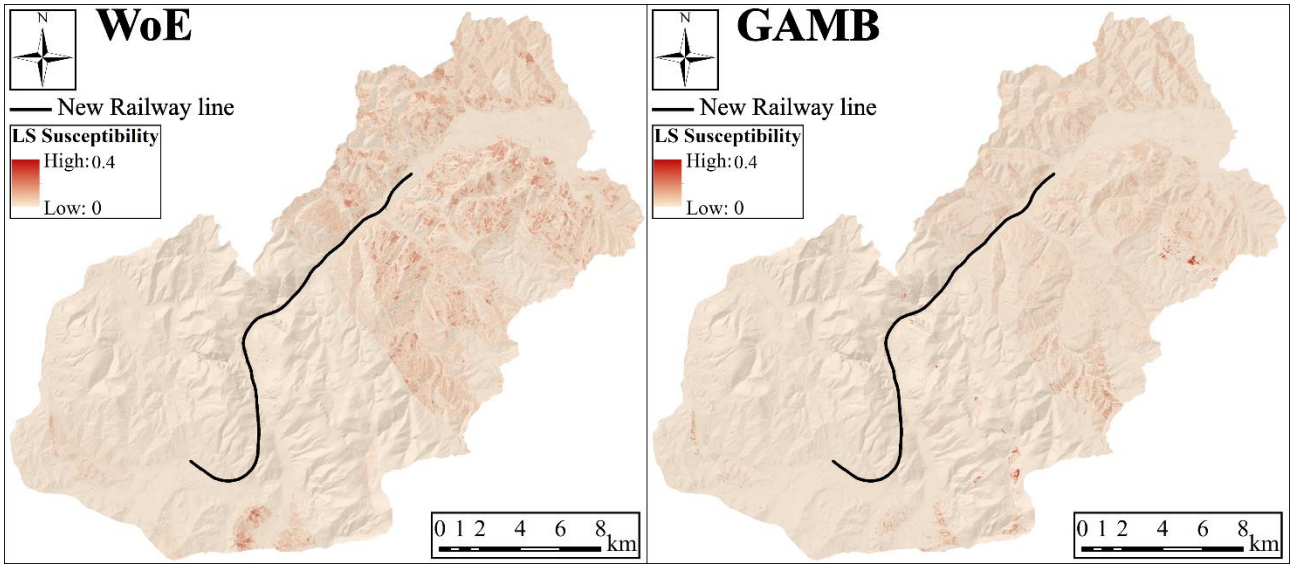
Yildiz, A., Zhao, H., & Kowalski, J. (2023). Computationally-feasible uncertainty quantification in model-based landslide risk assessment. Frontiers in Earth Science, 10, 1032438. https://doi.org/10.3389/feart.2022.1032438

Yu, L., Wang, Y., & Pradhan, B. (2024). Enhancing landslide susceptibility mapping incorporating landslide typology via stacking ensemble machine learning in Three Gorges Reservoir, China. Geoscience Frontiers, 15(4), 101802. https://doi.org/10.1016/j.gsf.2024.101802

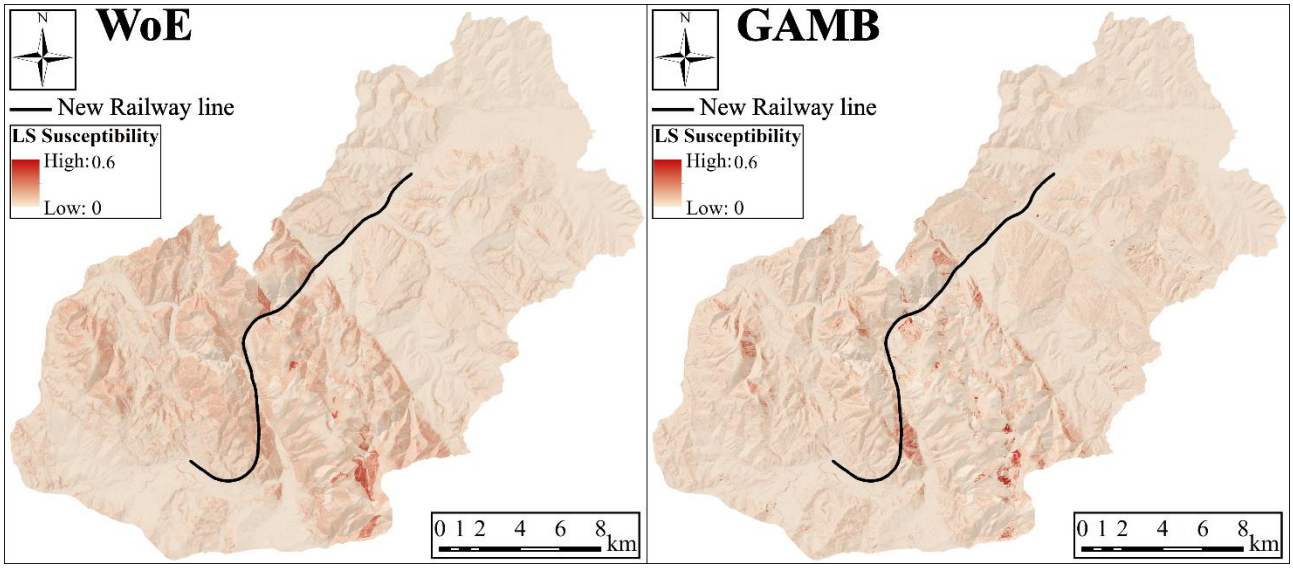
Yu, X., Wang, Y., Niu, R., & Hu, Y. (2016). A combination of geographically weighted regression, particle swarm optimization, and support vector machine for landslide susceptibility mapping: A case study at Wanzhou in the Three Gorges Area, China. International Journal of Environmental Research and Public Health, 13(5), 487. https://doi.org/10.3390/ijerph13050487

Zêzere, J. L. (2002). Landslide susceptibility assessment considering landslide typology: A case study in the area north of Lisbon (Portugal). Natural Hazards and Earth System Sciences, 2(1/2), 73–82. https://doi.org/10.5194/nhess-2-73-2002

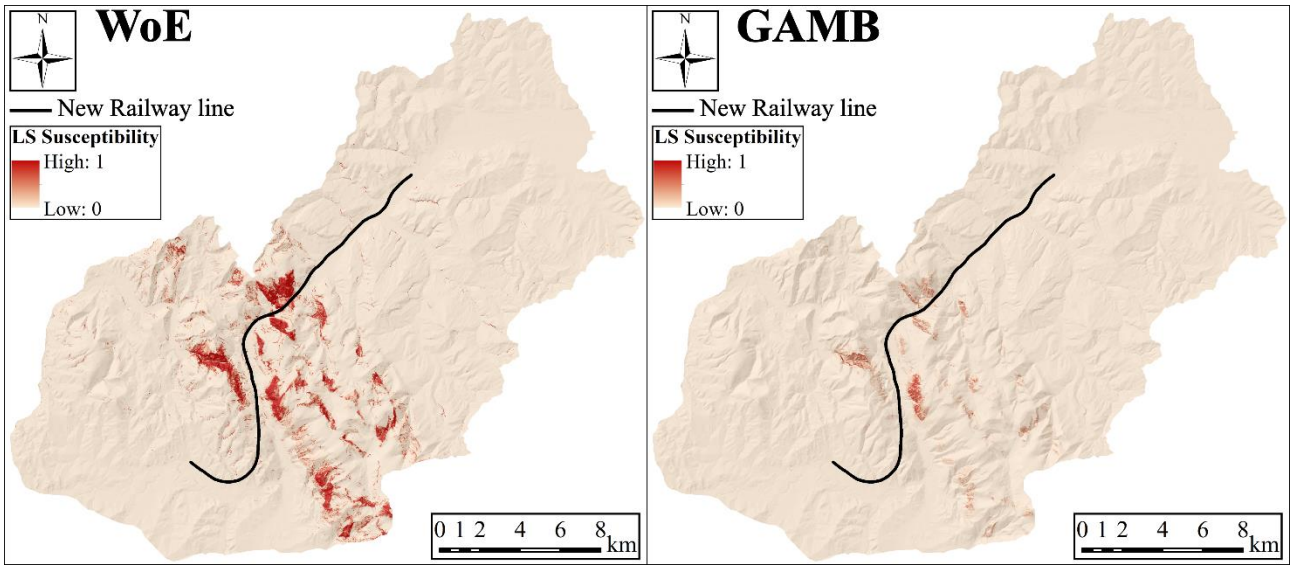
2.10 Supplementary material chapter 2



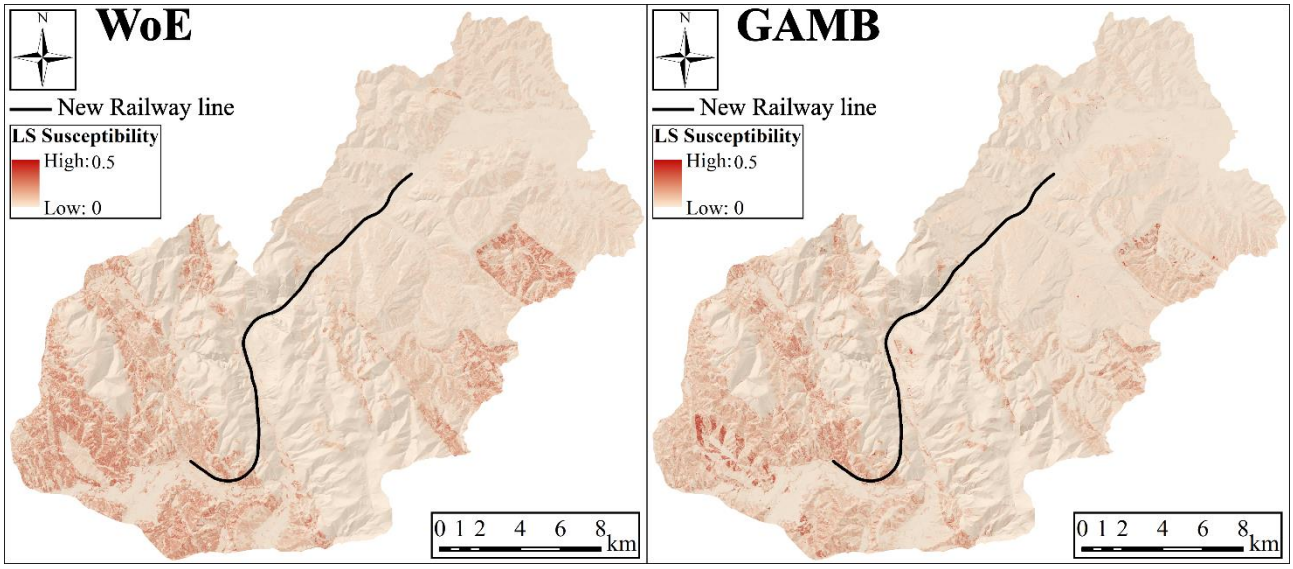
S 2.1. Not determinate (ND) landslide type susceptibility with both models (WoE and GAMB). Colors stretched from 0 to 0.4.



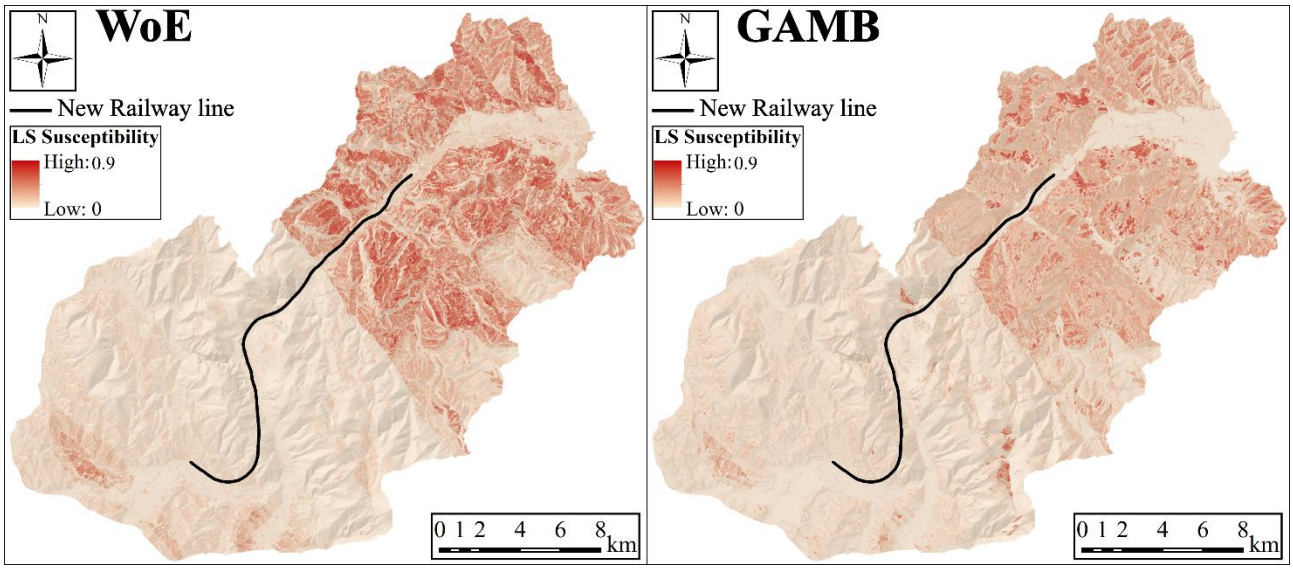
S 2.2. Slide (S) landslide type susceptibility with both models (WoE and GAMB). Colors stretched from 0 to 0.6.



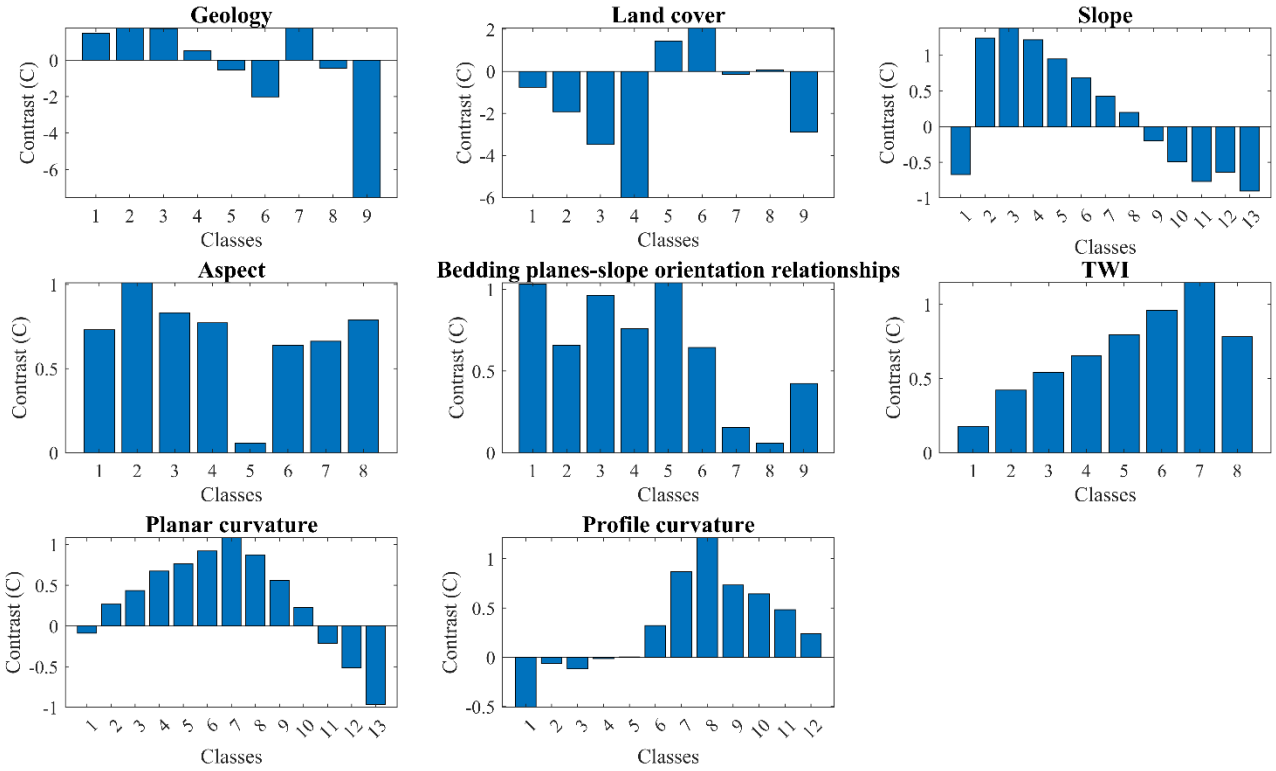
S 2.3. Fall/Topple (FT) landslide type susceptibility with both models (WoE and GAMB). Colors stretched from 0 to 1.



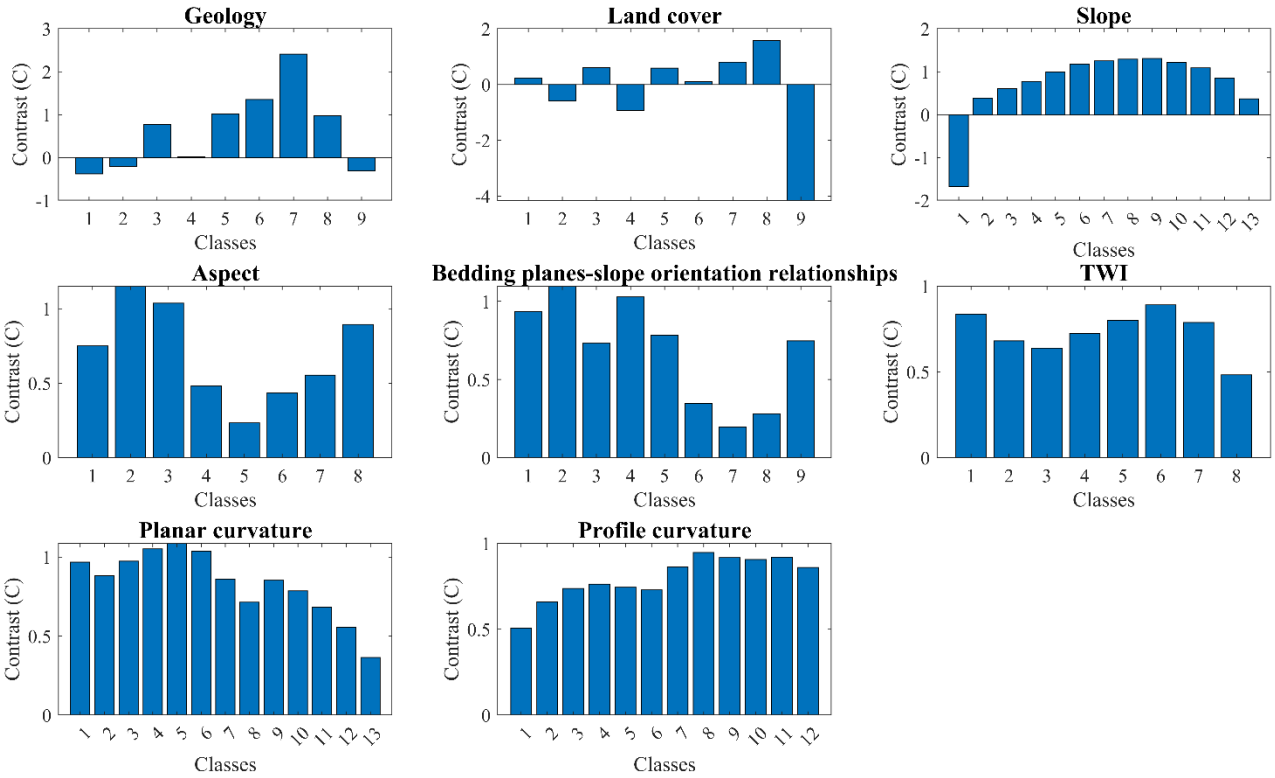
S 2.4. Slow flow (SF) landslide type susceptibility with both models (WoE and GAMB). Colors stretched from 0 to 0.5.



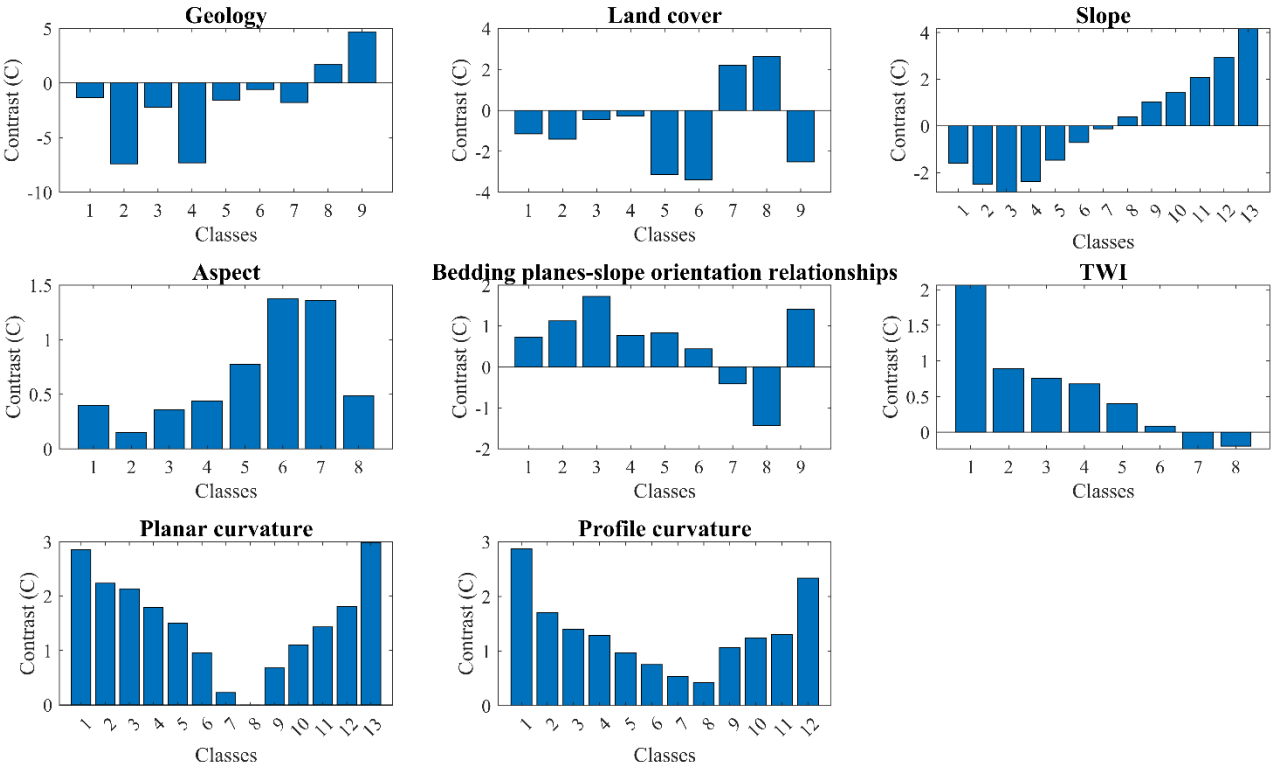
S 2.5. Complex (C) landslide type susceptibility with both models (WoE and GAMB). Colors stretched from 0 to 0.9.



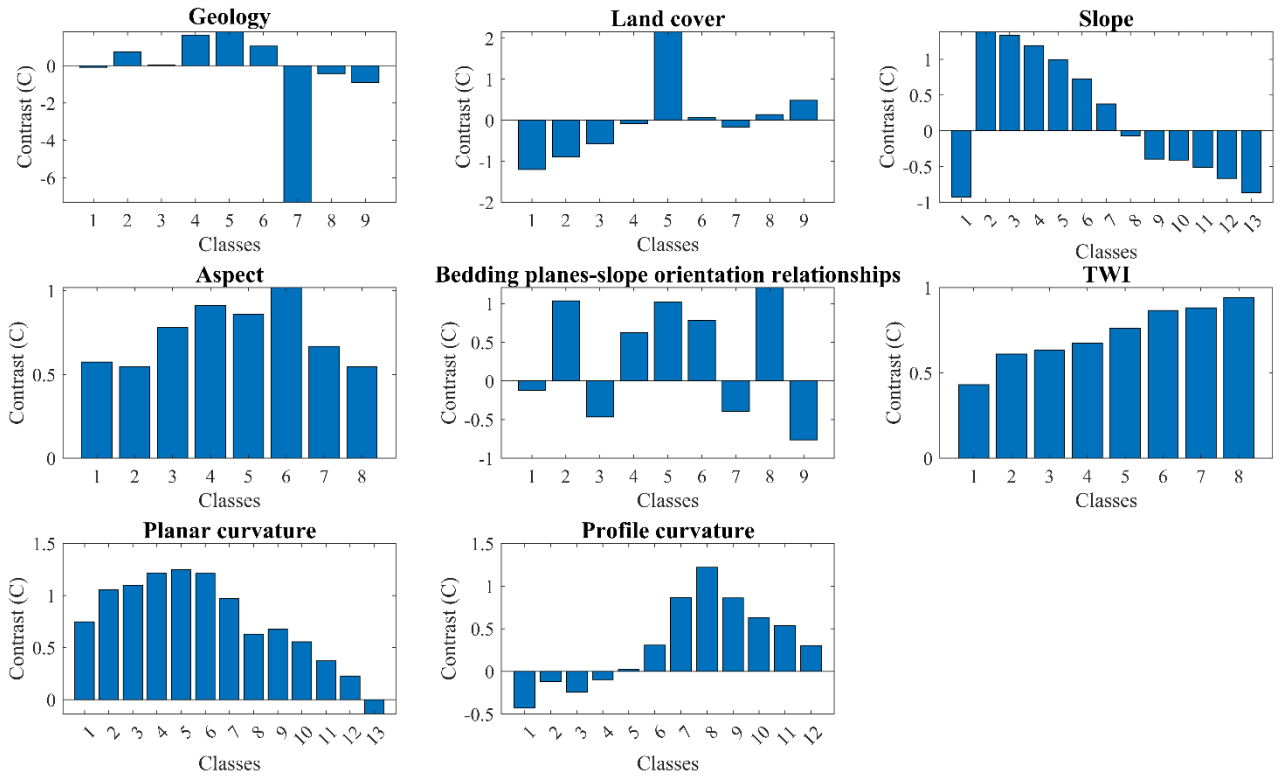
S 2.6. WoE Contrast (CS) of Not Determinate (ND) Landslide



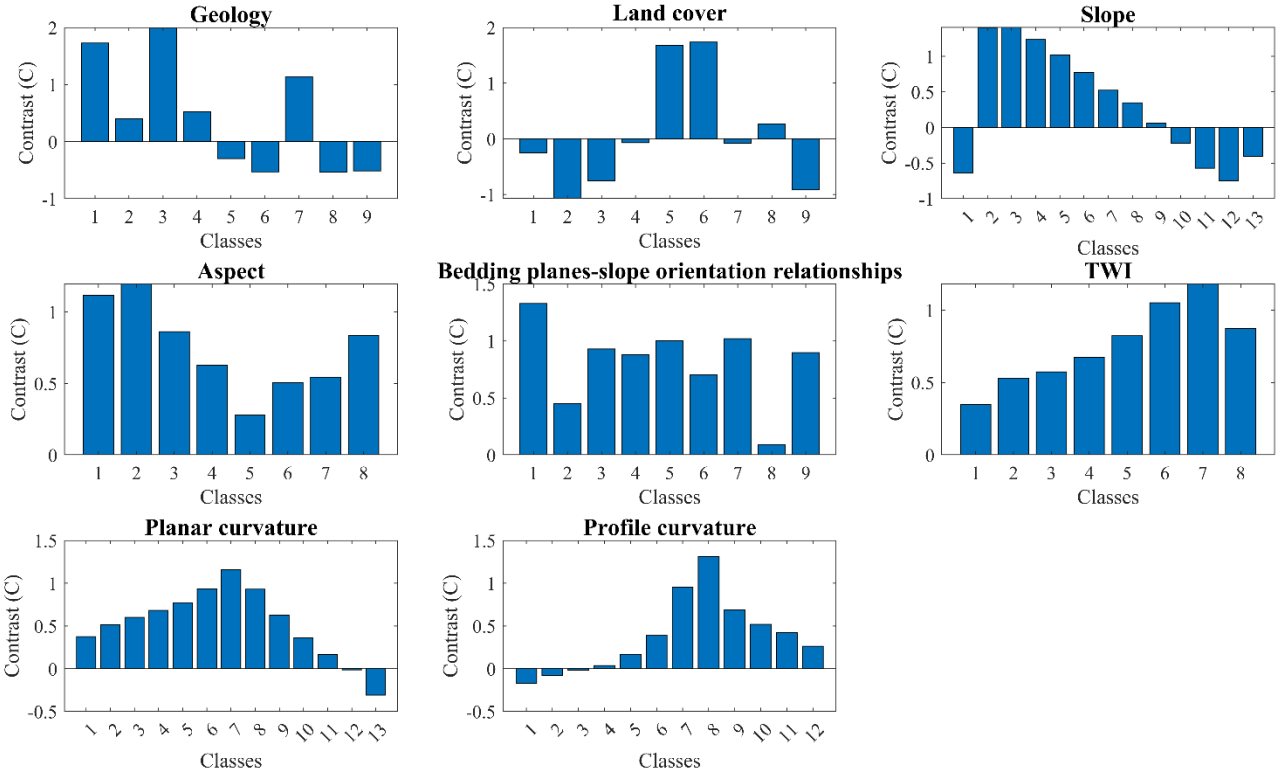
S 2.7. WoE Contrast (CS) of Slide (S) Landslide



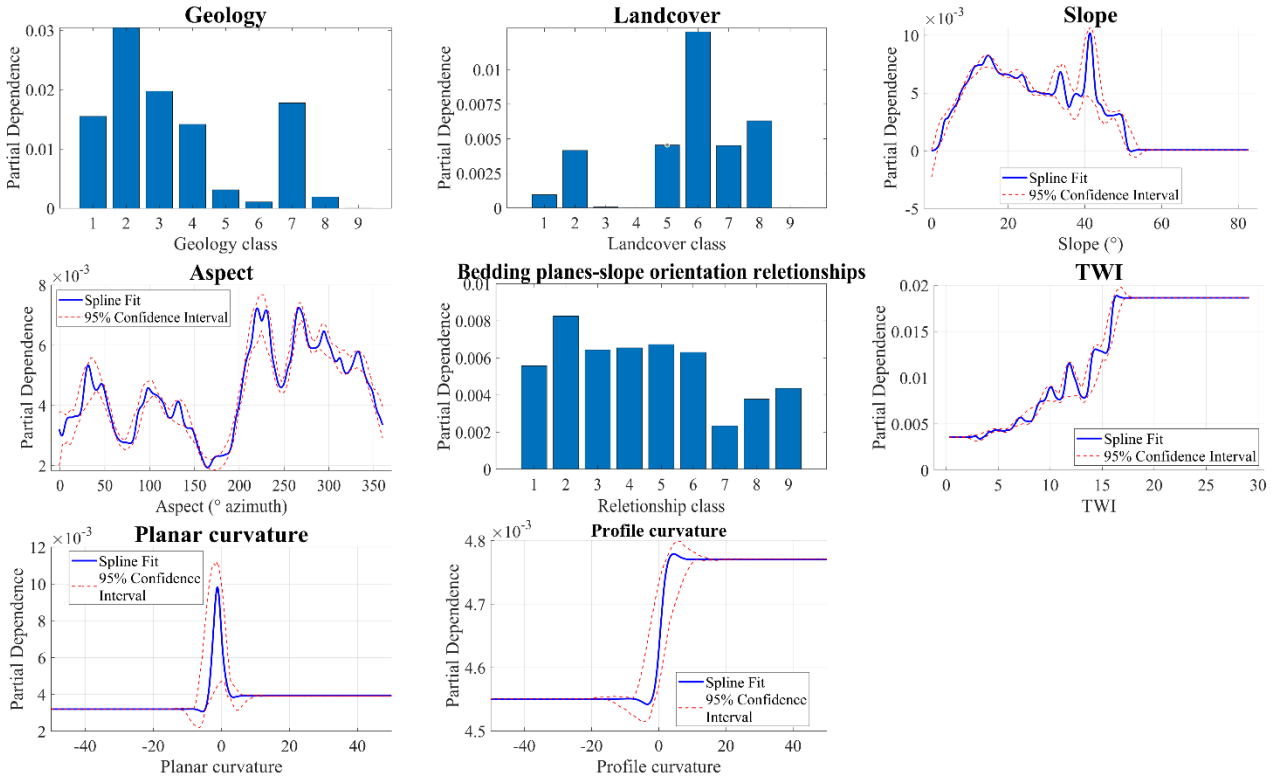
S 2.8. WoE Contrast (CS) of Fall/Topple (FT) Landslide



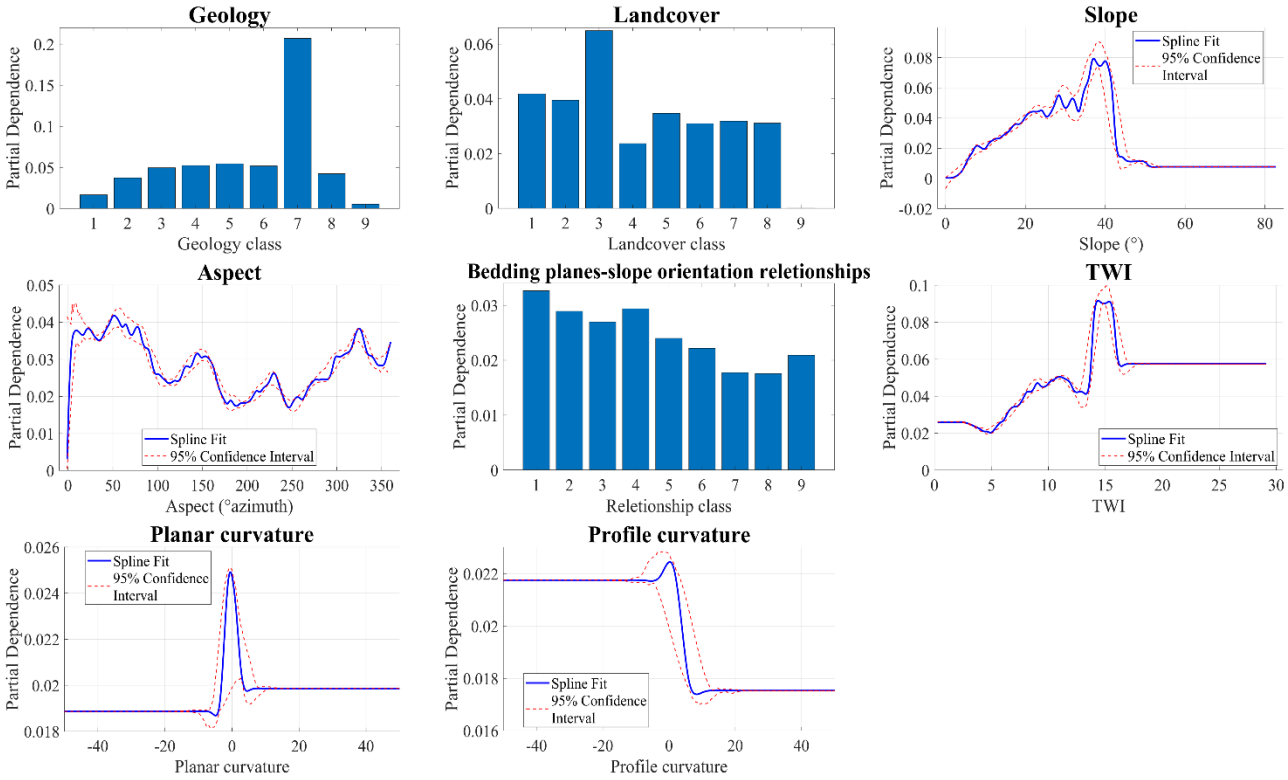
S 2.9. WoE Contrast (CS) of Slow Flow (SF) Landslide



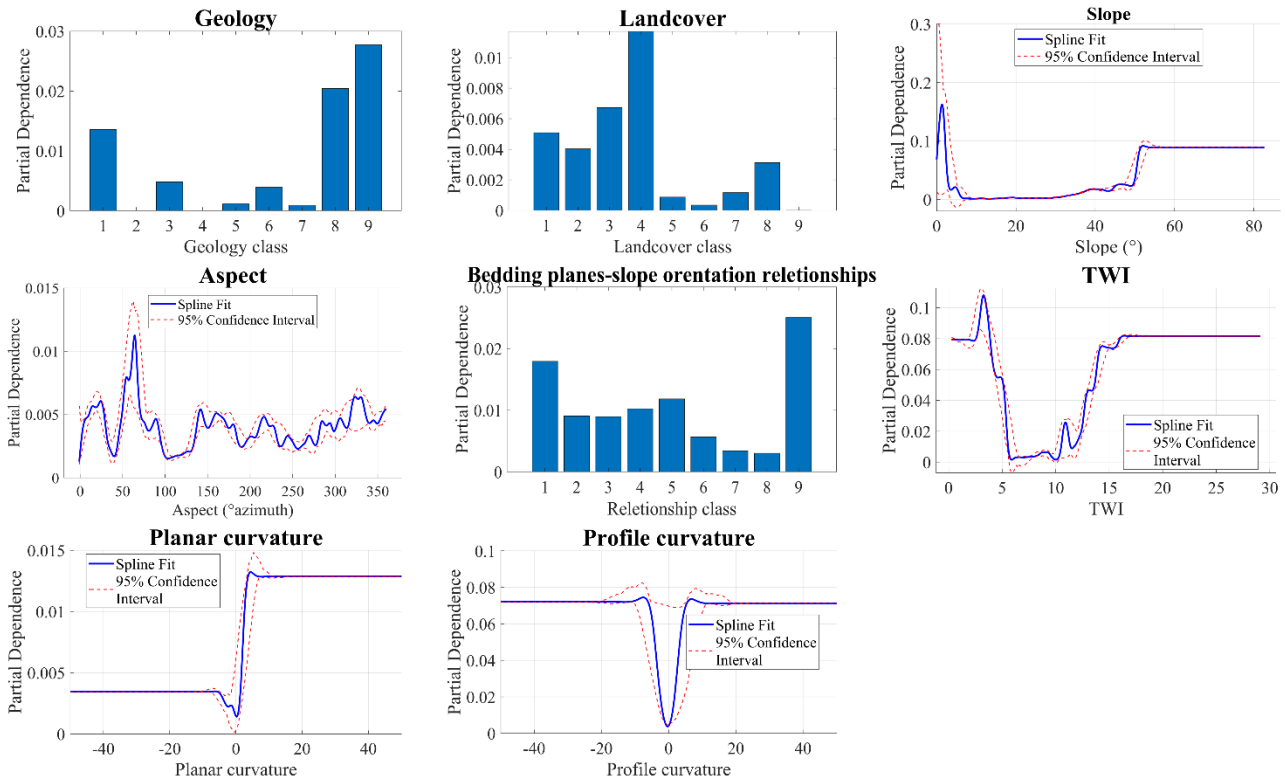
S 2.10. WoE Contrast (CS) of Complex (C) Landslide



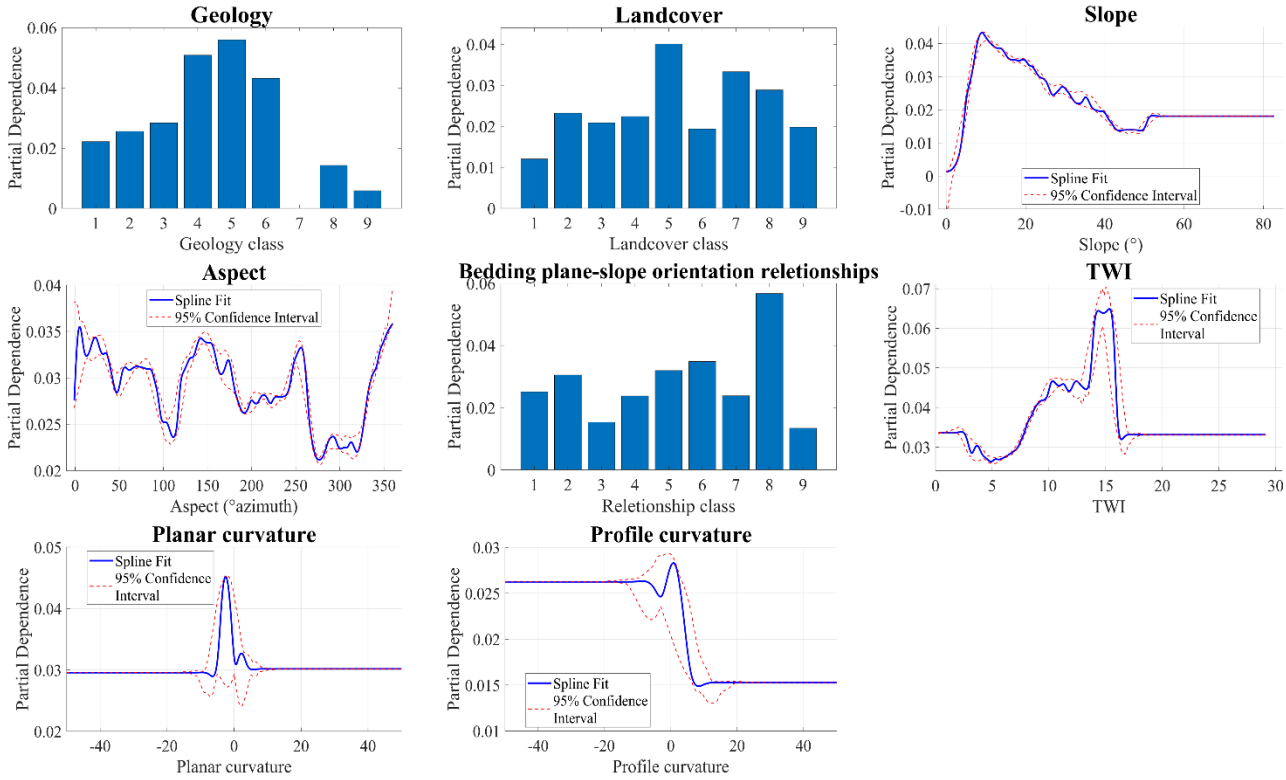
S 2.11. GAMB partial dependences (PD) of Not Determinate (ND) Landslide



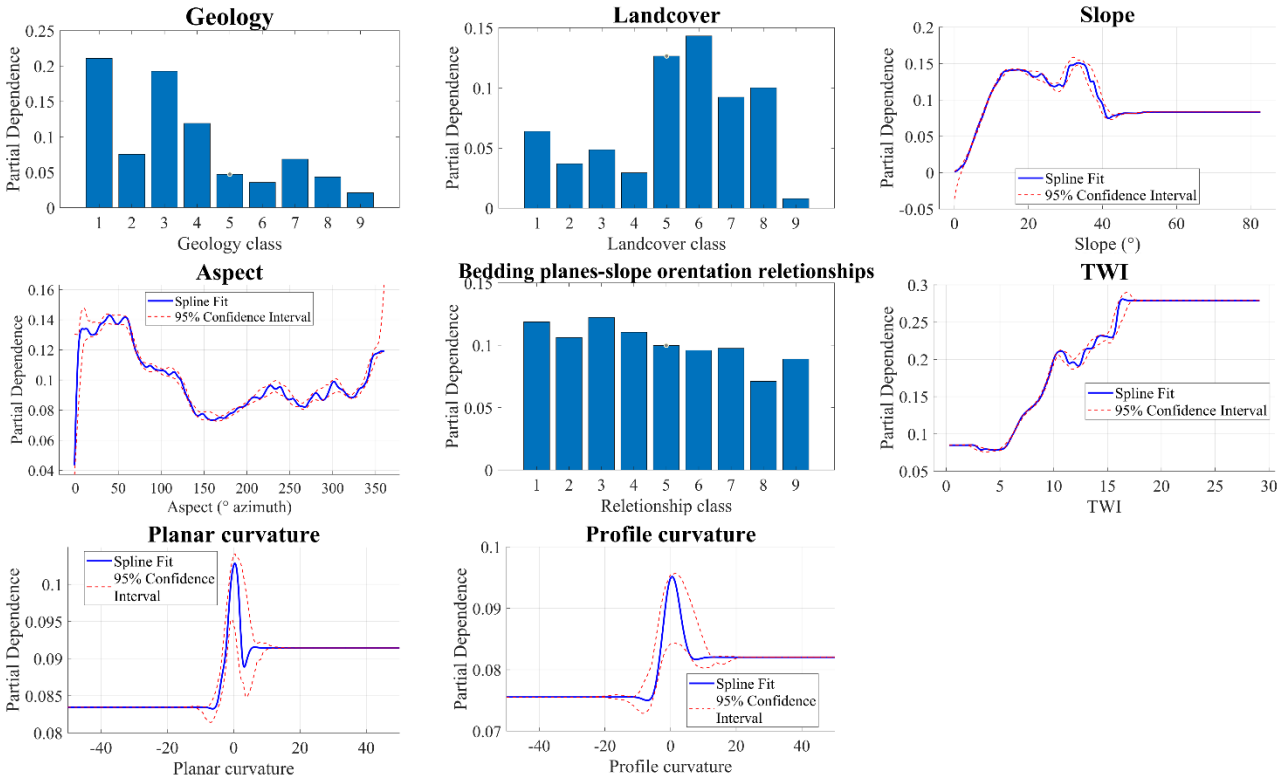
S 2.12. GAMB partial dependences (PD) of Slide (S) Landslide



S 2.13. GAMB partial dependences (PD) of Fall/Topple (FT) Landslide



S 2.14. GAMB partial dependences (PD) of Slow Flow (SF) Landslide



S 2.15. GAMB partial dependences (PD) of Complex (C) Landslide

3. Chapter 3: RER2023: the landslide inventory dataset of the May 2023 Emilia-Romagna event

3.1 Preface

This Chapter 3 explores a significant event in hydrogeological risk management: the comprehensive mapping and analysis of the May 2023 landslide event in the Emilia-Romagna region. Following the extreme rainfall events, this work became necessary to identify and assess the landslides triggered by these intense events. The project was conducted in collaboration with the Emilia-Romagna regional authorities, the University of Modena and Reggio Emilia, and the Po River District Authority.

The dataset presented here, compiled through a coordinated effort involving manual mapping of high-resolution aerial imagery and advanced geospatial techniques, represents a key tool for both immediate recovery operations and long-term hazard assessment. The resulting inventory, comprising over 80,000 landslides classified in 6 types, provides an unprecedented level of detail and serves as a crucial resource for refining future landslide susceptibility models in the region.This work is part of a larger, multi-institutional collaboration; my contributions focused on the direct manual mapping of landslides, active discussion of the landslide classification methodology, and the integration of rapid mapping approaches with manual mapping techniques.

The results of this work are directly linked to landslide susceptibility analyses. As introduced in the previous chapters, a reliable landslide inventory is essential as the starting point for statistical and machine learning models. Therefore, the creation of a detailed inventory is strongly recommended to support susceptibility and probability analyses aimed at preventing future catastrophic landslide events, as further applied and discussed in Chapter 4.

The chapter is based on a paper currently under review in the journal Earth System Science Data (ESSD):

Berti M.[1]*, Pizziolo M.[2], Scaroni M.[2], Generali M.[2], Critelli V.[3], Mulas M.[3], Tondo M.[3], Lelli F.[3], Fabbiani C.[3], Ronchetti F.[3], Ciccarese G.[1], Dal Seno N.[1], Ioriatti E.[1], Rani R.[1], Zuccarini A.[1], Simonelli T.[4], Corsini A.[2], 2025. RER2023: the landslide inventory dataset of the May 2023 Emilia-Romagna event. https://doi.org/10.5194/essd-17-1055-2025

[1] Department of Biological, Geological, and Environmental Sciences (BiGea), University of Bologna, Via Zamboni 67, 40126, Bologna, Italy

[2] Regione Emilia-Romagna, Area Geologia, Suoli e Sismica, Bologna, Italy.

[3] Department of Chemical and Geological Sciences, University of Modena and Reggio-Emilia, Modena, Italy.

[4] Autorità di Bacino Distrettuale del Fiume Po, Parma, Italy.

3.2 Abstract

Landslide inventories play a vital role in assessing susceptibility, hazards, and risks and are essential for developing resilience strategies in mountainous areas. This importance is amplified in the context of climate change as existing inventories might not adequately reflect changing stability conditions. In May 2023, the Emilia-Romagna region of Italy was hit by two major rainfall events, leading to widespread flooding and the triggering of thousands of landslides. Predominantly, these were shallow debris slides and debris flows, occurring on slopes previously deemed to be stable based on historical data, with no prior landslides recorded. Our team supported the Civil Protection Agency through field surveys and mapping efforts to pinpoint and record these landslides, prioritizing areas critical to immediate public safety and focusing on thorough mapping for future recovery planning. The outcome is a detailed map of all landslides induced by these events, manually identified using high-resolution aerial photography (0.2 m pixel resolution; four bands – RGB and near-infrared (NIR)) and categorized with the help of a 3D viewer. This comprehensive landslide inventory, comprising 80,997 landslide polygons, has been made openly accessible to the scientific community (https://doi.org/10.5281/zenodo.13742643, Pizziolo et al., 2024).

3.3 Introduction

Landslide inventories are crucial for susceptibility, hazard, and risk assessments and management (Soeters and Van Westen, 1996; Fell et al., 2008; Galli et al., 2008; Corominas et al., 2014). In Europe, landslide inventories are compiled on a national to a regional basis (Van Den Eeckhaut and Hervás, 2012) and can be supported by advanced landslide recognition and monitoring techniques (Guzzetti et al., 2012; Jaboyedoff et al., 2012; Amatya et al., 2021; Catani, 2021; Bhuyan et al., 2023). Landslide inventories should be as complete and spatially accurate as possible, and, also, they should consistently distinguish and classify different landslide types. These factors are important for improving frequency–area analyses (Malamud et al., 2004) and for obtaining reliable statistically based landslide susceptibility maps thanks to complete input data (Steger et al., 2017; Gaidzik and Ramírez-Herrera, 2021) and disjunct analysis of landslide types (Zêzere, 2002).

Generally, inventories of large-scale landslides are quite complete due to the landslides' geomorphic features that remain evident long after their occurrence and the fact that they can have slow movements detectable by remote sensing (Bertolini et al., 2017; Rosi et al., 2018; Luetzenburg et al., 2022; Ardizzone et al., 2023). On the contrary, regional or national inventories might not include a complete record of past shallow rainfall-induced landslides unless they were mapped soon after occurrence, i.e., before becoming hardly recognizable due to vegetation growth, rill erosion, or land cultivation (Guzzetti et al., 2004; Crozier, 2005; Cardinali et al., 2006; Zieher et al., 2016; Hao et al., 2020, Santangelo et al., 2023; Pittau et al., 2024). Therefore, it is crucial that existing national and regional landslide inventories be updated following each significant rainfall event to gather data that not only prove to be essential but also enhance the systematic application of landslide susceptibility maps in land use planning (Fell et al., 2008). This is a challenging and important task as the incidence of

shallow rainfall-induced landslides is likely to increase in Europe due to climate change (Gariano and Guzzetti, 2016; Handwerger et al., 2022; Aufli˘c et al., 2023).

In the Emilia-Romagna region (northern Italy), land use planning and land use restrictions are based on an inventory map of landslides at 1:10,000 scale (Bertolini et al., 2005) and on a catalogue of thousands of records referring to the activation or reactivation of landslides in the past (Piacentini et al., 2018). These documents are updated after every occurrence or reactivation of large-scale landslides and after multiple occurrences of shallow rainfall-induced landslides. In recent years, updates have been necessary to include hundreds of debris flows triggered during the rainstorm events that hit Parma Province in September 2014 (Corsini et al., 2017) and Piacenza Province in October 2015 (Scorpio et al., 2018; Ciccarese et al., 2020). In May 2023, the entire southern sector of Emilia-Romagna (from the provinces of Rimini to Reggio Emilia) has been hit by two consecutive exceptional rainfall events that triggered thousands of first-failure landslides. Ferrario and Livio (2024) provided an initial screening of these landslides through visual inspection of Planet satellite images (3 m resolution) and limited field surveys, while Notti et al. (2024) utilized an unsupervised identification method on Sentinel-2 images (10 m resolution). Both inventories were quickly compiled soon after the emergency, utilizing satellite imagery with resolutions that are relatively low considering the small size of the landslides from May 2023. Indeed, many of these landslides were so small that they were challenging to detect even at a 3 m resolution. Additionally, landslides were not classified or categorized by type of movement or material. Consequently, these two datasets were not intended to achieve – nor did they achieve – the level of completeness, consistency, and accuracy required for updating the official landslide inventory.

In this paper, we present the landslide inventory dataset of the May 2023 Emilia-Romagna events which has been designated as the reference map by the Emilia-Romagna region and the Po River Authority for the "Special Plan for interventions against situations of hydrogeological instability" (Piano Speciale Dissesto Idrogeologico – Eventi di Maggio 2023, https://www.adbpo.it/piano-speciale-dissesto-idrogeologico/, last access: 10 March 2025) (approved in a preliminary version in April 2024) to aid the Commission for Reconstruction in implementing the recovery phase. This spatial dataset is based on expert-based identification, mapping, and classification of landslides in high-resolution aerial images taken shortly after the second event (0.2 m resolution, RGB, and near-infrared). Particular attention has been given to the consistency of landslide type classification, which required the development and application of an algorithm for data harmonization across areas surveyed by different operators.

3.4 The May 2023 Emilia-Romagna event

The Emilia-Romagna region is located in northern Italy and stretches from the Apennine Mountains to the Po River Valley and eastward to the Adriatic Sea (Figure 3.1a). It is one of Italy's most economically prosperous regions, with a strong industrial base in automotive, machinery, food processing, and ceramics. The Po River Valley plays a vital role in agriculture, while the eastern coastline is a hub for both domestic and international tourism. By contrast, the Apennine Mountains present a more subdued economic landscape. The economic activity and population in these mountains have declined since the 1960s, and now there is a focus on

agrotourism, ecotourism, and niche markets. Regional initiatives are underway to foster economic growth in these mountainous areas.

In May 2023, the Emilia-Romagna region was struck by two exceptional rainfall events. The first, from 1 to 3 May, delivered approximately 200 mm of rain over a span of 48 h. Only 2 weeks later, on 16–17 May, a second event matched this intensity, with rainfall totals reaching 200–250 mm within another 48 h window. The recurrence interval for a single 2 d event was estimated to exceed 100–300 years, but the combined effect of these two closely timed events far surpassed 500 years (Brath et al., 2023). Both events impacted roughly the same area in the eastern part of the region (Figure 3.1).

These rainfall events led to extensive flooding across the Po Plain and triggered thousands of landslides in the Apennines. The total damages have been estimated to surpass EUR 9 billion, affecting roads, railways, buildings, and cultural heritage sites, along with the destruction of bridges, power facilities, and communication lines. Additionally, agricultural fields, farming operations, and cultivated slopes saw significant disruption over an area of about 1000 km². A total of 15 people lost their lives due to the flooding, and 2 lost their lives due to landslides.

The Emilia-Romagna region and the Italian government promptly responded to the event, mobilizing all necessary resources. The primary focus was on the Po Plain area, which is densely populated and houses the majority of industrial and agricultural activities. Consequently, the severe issues caused by landslides in the mountainous regions were initially overlooked. Over time, the significance of these issues became apparent, but even a year after the disaster, the situation remained critical. The impact of landslides in the Apennines has been especially severe due to the local economy's vulnerability, the extensive damages to infrastructure, and the significant land loss, all of which have slowed and complicated the recovery process.

We assisted local and national agencies and working groups in addressing the problems caused by landslides. The initial 2 weeks following the event were primarily focused on field surveys and rapid assessment of the most critical situations that demanded immediate actions to ensure public safety. Subsequently, our efforts shifted towards landslide mapping. In a first stage, it was crucial to identify the roads and buildings affected by landslides to coordinate emergency interventions and to perform an initial damage assessment. Afterward, we completed the landslide inventory to develop a comprehensive map detailing all landslides triggered by the event across the area. This map has been officially designated as the landslide map for the May 2023 event by the Po River Authority and the Emilia-Romagna region, and it is currently being used by the Commission for Reconstruction in implementing the recovery phase.

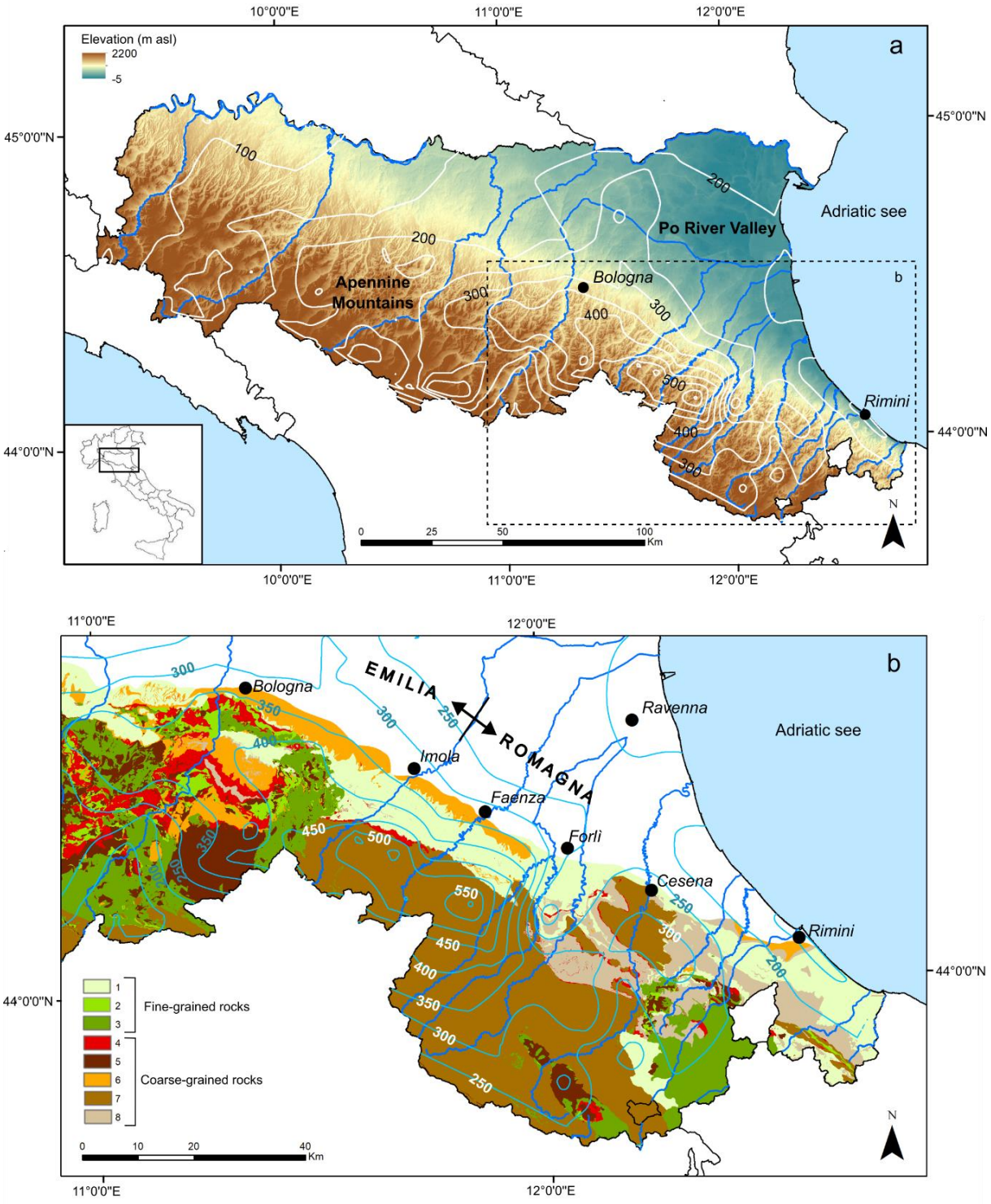


Figure 3.1. ***a****) Overview map of the Emilia-Romagna region (Italy), illustrating elevation and cumulative rainfall isohyets from May 1-17, 2023.* ***b****) Detailed view of the area most impacted by the event, featuring the geological units referenced in Table 3.1.*

3.5 Methods

This section describes the methodology used to develop the landslide inventory for the May 2023 event. It covers the classification of lithological units, the identification and mapping of

landslides, their classification, and the quality control and data harmonization procedures implemented.

3.5.1 Litho-technical units classification

In the study area, bedrock geology significantly influences the morphology of the slopes; the mechanical properties of the weathered soil layer; the vegetation cover; and, consequently, the proneness to slope instability. As a matter of fact, these factors played a crucial role in the behavior of the slopes during the May 2023 event, controlling the type and density of landslides. Consequently, bedrock geology is an essential base layer of our landslide inventory.

The geological map of the Emilia-Romagna region, created by the regional Geological Survey at a 1:10,000 scale (AGSS-RER, 1986), includes more than 600 geological formations. These formations are distinguished by unique features that signify variations in depositional environments, composition, or geological age. The variety and detail of these formations illustrate the region's complex geological history and the precision employed in the map's creation.

For our inventory, we categorized all geological formations into eight distinct units, as depicted in Figure 3.1b and detailed in Table 3.1. These units were delineated by merging lithological characteristics with their respective structural domains, recognizing that the same rock type can display varying structural and mechanical properties depending on its location within the orogenic sequence. For example, flysch rocks within the Ligurian domain (unit 5) are generally more fractured, less resistant, and prone to deep-seated landslides compared to those in the Tuscan–Umbrian domain (unit 7) due to the extensive tectonic stress they endured in the accretionary wedge.

The eight units identified were further divided into two broad categories: fine-grained rock masses (units 1 to 3) and coarse-grained rocks (units 4 to 8). This categorization aids in the preliminary differentiation of the types of weathered-soil covers these rocks produce, which experienced widespread landslides during May 2023. Coarse-grained rocks typically produce granular soils composed of sand, gravel, and cobbles, with smaller amounts of silt and clay, aligning with the "debris" category in the Cruden and Varnes (1996) classification. In contrast, fine-grained rocks lead to the formation of fine soils predominantly made up of silt and clay, fitting the "earth" classification. These two categories, debris and earth, are utilized to classify landslides that occurred on soil-covered slopes.

Unit ID	Lithology	Domain	Structural position	Geological Age
1	Clays, silty clays, and marly clays	Padano-Adriatic	Outer Foredeep	Pliocene to Pleistocene
2	Marls and marly clays	Epiligurian	Wedge-top basins	Oligocene to Miocene
3	Clay shales, clay breccias, tectonized clays, olistostromes	Ligurian	Accretionary wedge	Cretaceous to Eocene

4	Massive rocks: basalts, serpentines, limestones, arenites	Ligurian, Epiligurian	Accretionary wedge Wedge-top basins	Cretaceous to Miocene
5	Flysch rocks made of rhythmic alternations of sandstones, limestones, pelites, and shales	Ligurian, Epiligurian	Accretionary wedge Wedge-top basins	Cretaceous to Eocene
6	Weakly cemented sandstones and conglomerates	Padano-Adriatic	Outer Foredeep	Pliocene to Pleistocene
7	Flysch rocks made of rhythmic alternations of sandstones and pelites	Tuscan-Umbrian	Inner Foredeep	Miocene
8	Weakly cemented sandstones with interbedded pelitic layers	Padano-Adriatic	Outer Foredeep	Pliocene to Pleistocene

Table 3.1. Classification of the geological formations in the Emilia-Romagna region into eight units, based on their lithological composition and geological structural domains. Units 1 to 3 consist mainly of fine-grained rocks, while units 4 to 8 are primarily composed of coarse-grained rocks.

3.5.2 Landslide identification and mapping

Landslide identification and mapping were conducted by means of photo-interpretation of high-resolution aerial images. These images were captured using a Leica DMC III sensor aboard a Cessna 402C aircraft, flying at approximately 4700m.a.s.l. The images, taken shortly after the second rainfall on 23 May 2023, have a 0.2m resolution and include four bands: RGB and near-infrared.

The mapping process was organized as follows. The total area was segmented based on the administrative boundaries of the municipalities. These sections were then distributed among three institutions: the University of Bologna, the University of Modena and Reggio Emilia, and the Geological Survey of the Emilia-Romagna Region. Each institution assigned 4 mappers, with a total of 12 individuals being involved in the effort. Landslide detection was conducted in GIS environment by comparing pre- and post-event images with an on-screen zoom of approximately 1 V 1000 (Figure 3.2a and b). Once a landslide was spotted, further inspection was conducted using near-infrared (NIR) and normalized difference vegetation index (NDVI) images (Figure 3.2c and d) and supplemented by a 3D viewer featuring high-resolution images overlaid on a 30m DEM (Copernicus GLO-30, ESA, 2024; Figure 3.2e). Viewing the slope from different angles enhanced the delineation of the affected area and the interpretation of the type of movement. Following this analysis, each landslide was classified into the specified classes described in the next section. The digital mapping of the landslide polygon was then executed at scales ranging from 1 V 800 to 1 V 200, depending on the landslide's size, ensuring precise tracing of the affected perimeter. In this final stage, the regional topographic map at 1 V 10000 scale was utilized to further verify alignment with the existing topography (Figure 3.2f). While the delineation of the landslides was carried out at a large scale for precise mapping

of the boundaries, the final inventory is designed to be appropriate for consultation at a scale of 1 V 2000.

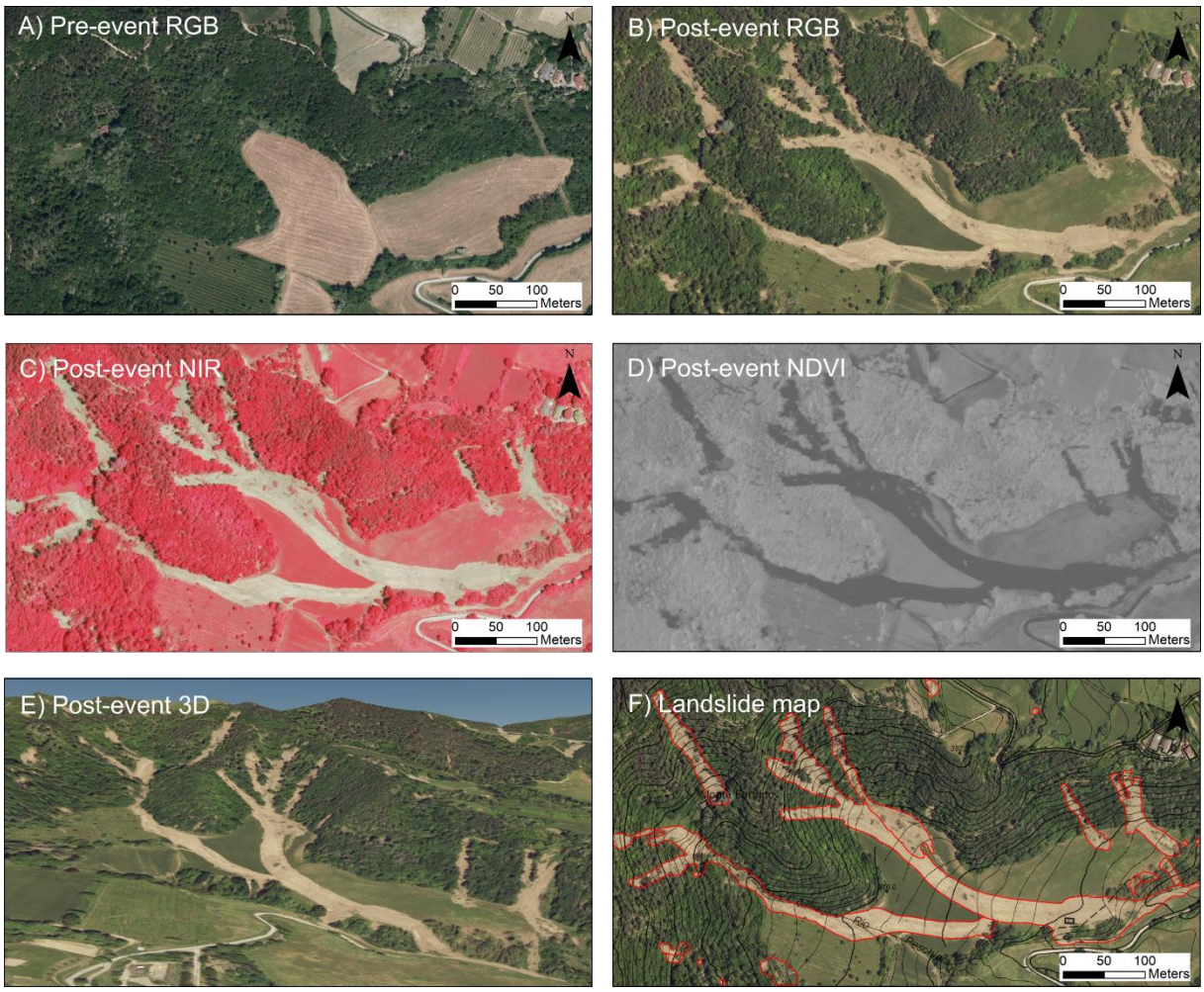


Figure 3.2. Example of manually identifying and mapping landslides. (a) Pre-event image (AGEA aerial photos, April–July 2020, 0.2m resolution), (b) post-event image (23 May 2023, 0.2m resolution), (c) near-infrared (NIR) images derived from post-event images, (d) normalized difference vegetation index (NDVI) images derived from post-event images (dark colors signify absence of vegetation), (e) 3D visualization of post-event images, (f) landslide map manually created using (a) to (e) and aligned with the pre-event topographic map at a 1 V 10000 scale.

Overall, the process of identifying landslides was fairly objective. Each landslide cleared vegetation, uncovered distinct patches of bare soil or bedrock, and led to deposits of loose material. The post-event images, taken only 10 d after the second rainfall, clearly displayed these geomorphological markers, eliminating any ambiguity in recognizing the landslides from the event. Additionally, most of the landslides were new occurrences, not present in the pre-event images; even in case of re-activations, it was straightforward to identify the newly affected areas. This gives us confidence that our dataset exclusively contains landslides from the May 2023 event. It is important to emphasize the distinct nature of this dataset. The existing landslide map of the Emilia-Romagna region (https://geoportale.regione-emilia-romagna.it/, last access: 3 March 2025) encompasses all landslides identified in the region via photo

interpretation, historical data, and field surveys. This map provides a comprehensive overview of landslide activity across a broad geological time frame, accounting for various climatic, seismic, and morphological conditions. In contrast, our dataset specifically captures the landslide response to a single critical meteorological event. Additionally, the criteria used to distinguish different types of landslides are specifically designed for this scenario, as explained in the subsequent section (see Sect. 3.3).

While identifying the landslides was straightforward, defining their exact boundaries was more subjective. Many landslides became fluidized upon failure, with the distal debris spreading among trees without removing vegetation, thus complicating the mapping of the deposit (Figure 3.3a). Moreover, several slopes experienced complete removal of soil cover by adjoining shallow failures, blurring the distinction between individual slides (Figure 3.3b). In these cases, we chose to interpret the landslide boundaries rather than just tracing the visible debris edges. For fluidized slides, the polygons were adjusted by connecting visible debris patches to include areas obscured by vegetation. For coalescent slides, we attempted to map each individual slide by identifying distinctive arcuate shapes along the detachment scarps that signified separate failures. Although this approach introduced some subjectivity into the manual mapping process, it was necessary to create a dataset suitable for analyzing the morphometric features of the landslides.

The manual mapping process was demanding and laborintensive. Initially, we focused on mapping landslides in the areas most severely impacted by the event, particularly around roads and urban centers. This priority was set to align with the Civil Protection Agency's needs to identify damages during the emergency response. This initial phase of mapping, which produced several "damage maps" for the affected municipalities, spanned the first 2 months following the disaster. Subsequently, the landslide inventory was expanded to cover the entire area over the following months. The complete process took approximately 6 months to finish.

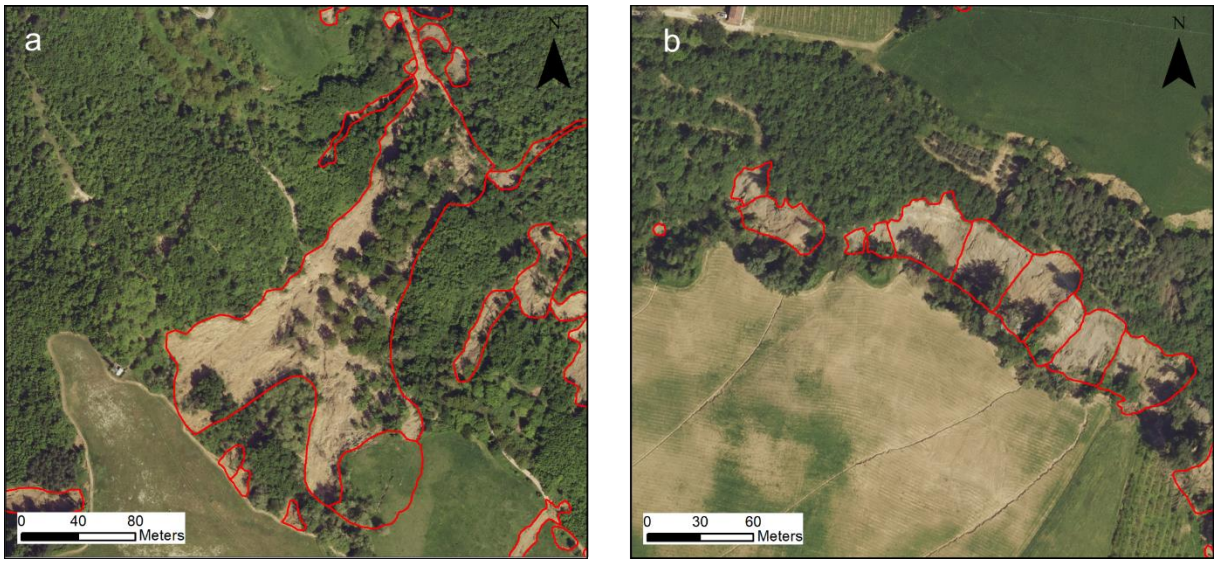


Figure 3.3. Examples of uncertain landslide boundary delineations. (a) Debris slides interspersed among trees without clearing the vegetation, (b) multiple coalescent debris slides segmented into individual slides.

3.5.3 Landslides classification

Right from the start of our work, the classification of landslides triggered by the event was identified as a crucial task. We utilized the well-known Cruden and Varnes (1996) classification system, which categorizes landslides based on two primary criteria: the type of movement and the type of material. Most of the landslides from the May 2023 event were categorized either as debris slides, where mixed granular material moves along a plane of weakness as a relatively coherent mass, or as debris flows, where the material moves in a fluid-like manner over greater distances. However, we encountered challenges in further distinguishing between various types of debris slides and debris flows, a differentiation not addressed by the standard classification.

Figures 3.4 and 3.5 illustrate the problem. According to the Cruden and Varnes (1996) classification, all the landslides depicted in Figure 3.4 can be classified as debris flows. Yet, clear differences are apparent between the upper (a1–3) and the lower three (b1–3). The latter are typical debris flows that start on a steep slope and stop as the slope decreases; the former, while starting similarly on steep slopes, demonstrate extensive propagation and complete fluidization of the deposit as they travel much further. None of these cases involve a well-defined channel; thus, the classification proposed by Hungr et al. (2014) that distinguished channelized-debris avalanches flows from unchannelized-debris avalanches does not apply here. A similar challenge presents itself with debris slides (Figure 3.5). Current classifications fail to distinguish between slides of different degrees of mobility, a distinction that is clearly visible in the field. In fact, some slides exhibited high mobility, completely clearing the vegetation (upper pictures, a1–3), while others show low mobility, indicated by minimal vegetation damage (lower pictures, b1–3). Understanding the conditions that lead to these diverse behaviors is crucial for hazard assessment and necessitates differentiating these phenomena.

An additional classification challenge involves rock-block slides. These landslides impacted the homoclinal slopes of the Marnoso-Arenacea Formation (unit 7 in Fig. 3.1b) and manifested as massive, translational rock-slab slides along bedding planes. While classifying these landslides poses no issues, it was necessary to distinguish between rock slides based on their degree of evolution. Some experienced movements ranging from several meters to tens of meters, signaling paroxysmal failures (Fig. 3.6a1–a3), whereas others shifted merely a few centimeters, indicative of incipient, undeveloped failures (Fig. 3.6b1–b3). The latter represent highly dangerous zones prone to potential collapse and thus required special attention.

These classification challenges were extensively discussed by our team. We ultimately decided to adhere to the Cruden and Varnes (1996) classification system to define the primary types of landslides. These include debris slides (DSs), debris flows (DFs), and rock-block slides (RSs) in the coarse-grained units (Fig. 3.1b), along with earth slides (ESs) and earth flows (EFs) in the fine-grained units. Then, we introduced the informal subclasses of high-mobility debris slides (DS1), low-mobility debris slides (DS2), long-runout debris flows (DF1), limited-runout debris flows (DF2), fully developed rock slides (RS1), and incipient rock slides (RS2) to capture the varied behaviors observed in the field. Subclasses were not assigned to earth slides (ESs) and earth flows (EFs) because landslides in areas with fine-grained soils were significantly less

frequent and had milder impacts. This variation in response is likely tied to the distinct hydrological behaviors of fine-grained soils within the study area. Previous studies, which include statistical analyses of critical rainfall (Rossi et al., 2012; Berti et al., 2012) and field monitoring of unstable slopes (Berti et al., 2010; Berti and Simoni, 2012), indicate that these clay-rich soils are more prone to failure during extended periods of rainfall rather than during brief, intense downpours that generally cause surface runoff and flooding.

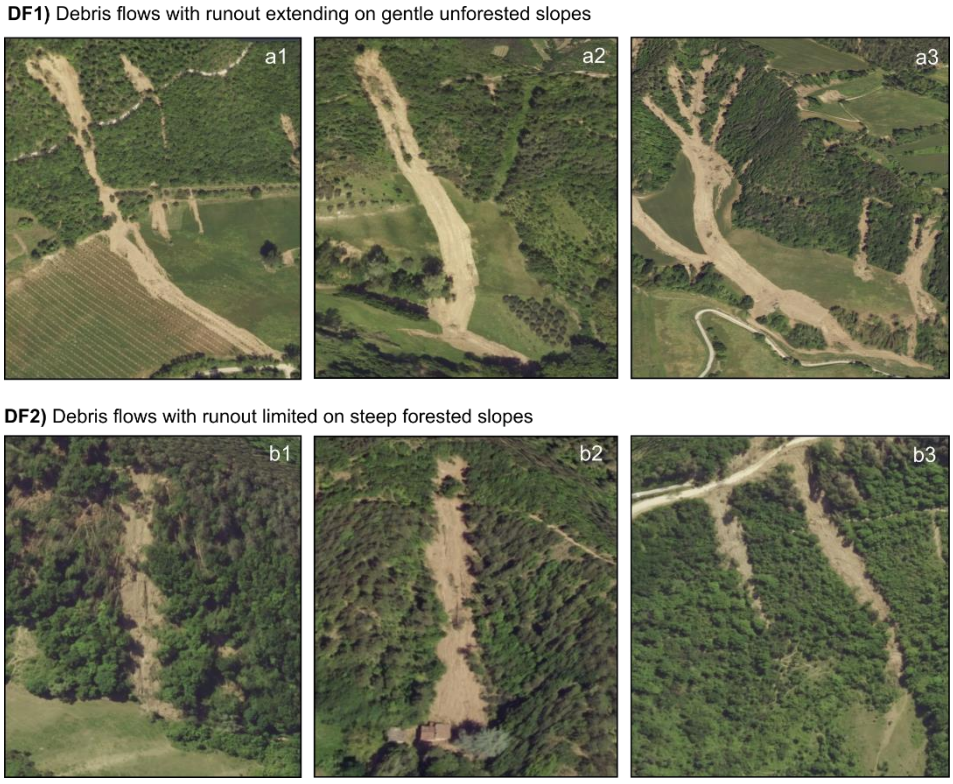


Figure 3.4. Representative images of the two distinct types of debris flows caused by the May 2023 event.

Figure 3.5. Representative images of the two distinct types of debris slides caused by the May 2023 event.

Figure 3.6. Representative images of the two distinct types of rock-block slides caused by the May 2023 event.

3.5.4 Quality control

To ensure the consistency of the results, an experienced geologist conducted a comprehensive review of the entire area after the completion of the manual mapping. This critical review focused on several key aspects to verify that all mappers adopted the same standards of detail and accuracy that were established at the start of the work. Guided by a 1×1 km grid, the reviewer assessed the following: (i) any missed landslides, (ii) the precision in outlining landslide boundaries, (iii) the consistency in interpreting vegetated areas, (iv) the adherence to the established classification criteria, and (v) the accurate segmentation of individual slide events.

The findings from this review were summarized in a report sent to the 12 mappers. The report ranked the need for revisions in each municipality from “small” to “high” and included a detailed explanation of the necessary adjustments, along with screenshots highlighting the errors detected. Each mapper was then tasked with revising their section of the manual inventory based on this feedback. This review and revision phase lasted approximately 2 months.

Following these adjustments, the landslide map was significantly improved. Although some variations persisted in the resolution of digitization and in interpreting boundaries obscured by vegetation, the primary discrepancies were effectively addressed and resolved. The only remaining issue was the variation in landslide classification among mappers, which was evident when comparing the inventory maps of different municipalities. To address this, we developed the automatic procedure detailed in the next section.

3.5.5 Data harmonization

As mentioned earlier, the criteria for landslide classification were extensively discussed among us. We collectively examined and analyzed approximately 50 complex cases of landslides with uncertain classifications, particularly those that straddle the categories of slides and flows and incipient rock-block slides. This was done to synchronize the understanding among mappers and to promote a consistent approach to classifying these phenomena. These preparatory efforts resulted in a substantial homogeneity in the classification of rock-block slide phenomena. However, notable variations were still evident in the final map between different types of debris slides and flows, which were influenced by the subjectivity of the mappers. Some determined the classification based on the landslide’s shape, while others determined the classification on the presence of flow-like features in the deposition area, and yet others made this determination based on the texture of the debris within the landslide polygon. On the other hand, such a distinction is inherently subjective due to the gradual transition between slides and flows, making it challenging to establish a clear-cut boundary or to define strict classification rules. The same issue arose with classifying landslides based on their degree of mobility. While it was evident that many landslides exhibited complete fluidization and high mobility, significant discrepancies persisted in how the mappers categorized these events.

To address this issue, we implemented the automated procedure depicted in the flowchart of Figure 3.7. This procedure utilizes standardized criteria to ensure uniform classification of

landslides throughout the area and to correct inevitable errors in such a large dataset. The automated procedure was applied to all landslides except for rock-block slides, which have distinctive features that all mappers clearly and consistently recognized. Four key steps were identified to achieve a consistent classification of material type, movement type, degree of mobilization of debris slides, and degree of fluidization of debris flows.

Material type (debris or earth)

The initial step was verifying the classification of material types. As previously mentioned, landslides on soil-covered slopes were divided into debris and earth categories according to the Cruden and Varnes (1996) classification. In our study area, this classification is clearly defined by the underlying bedrock geology; debris is derived from coarse-grained rock units, and earth is derived from fine-grained units (Fig. 3.1b).

All mappers employed this classification system, referencing the geological map of the Emilia-Romagna region at a 1:10,000 scale, which provided an objective and standardized framework for classifying material types. However, manual mapping led to inconsistencies due to human error and subjective judgments, particularly when categorizing landslides spanning multiple material types. To address these issues, we overlaid landslide polygons on the lithological map (Fig. 3.1b) and classified each landslide as either debris or earth based on its polygon's centroid location. This classification was achieved through a simple spatial join between the landslide data and the lithological map within the GIS environment.

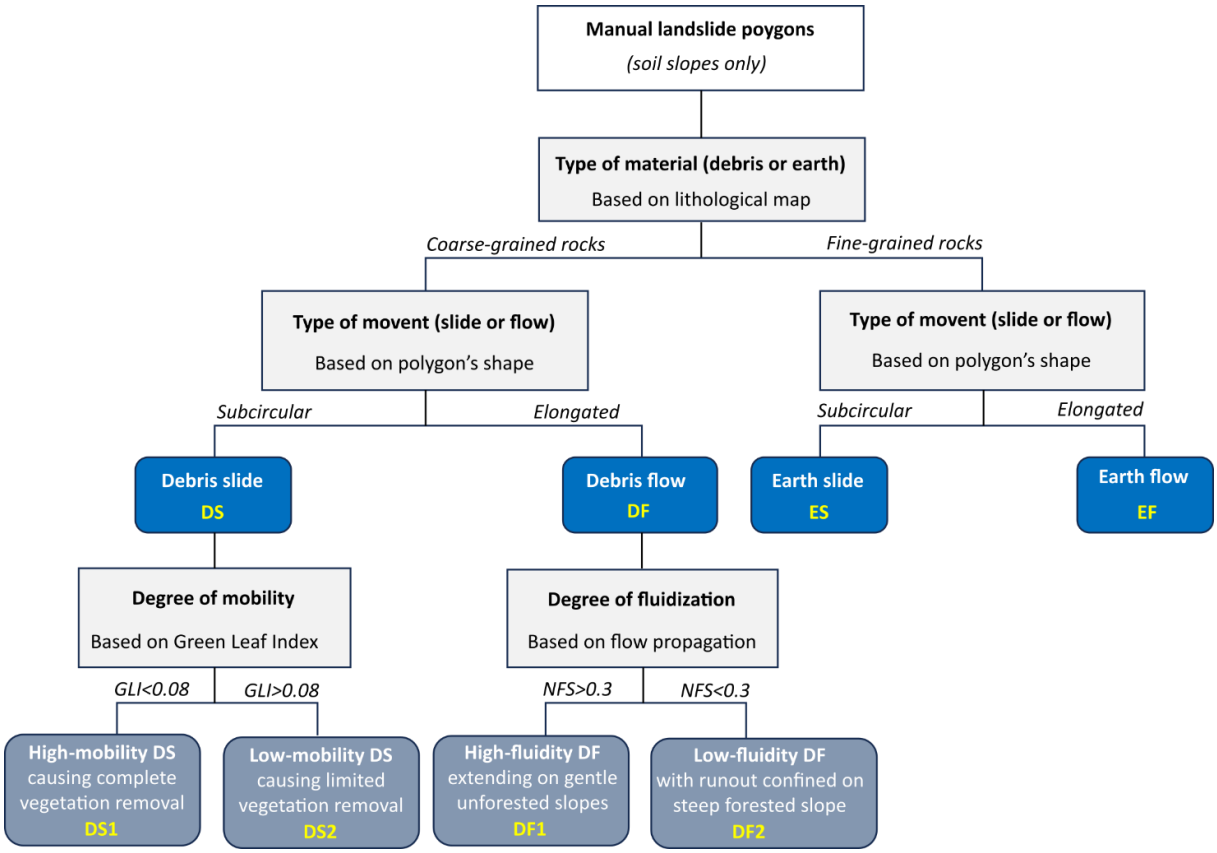


Figure 3.7. Flowchart depicting the process used to ensure data quality and standardize the classification of manually mapped landslides.

Type of movement (slide or flow)

To standardize the distinction between slides and flows, we employed a standard convolutional neural network (CNN) specifically designed to recognize the distinct shapes of slides and flows. The CNN was trained with data from the municipality of Casola Valsenio. This area was chosen due to its highly accurate manual mapping and because it served as the initial training ground for the mappers, where classification challenges were collaboratively discussed.

The CNN features an input layer, two convolutional layers (each with batch normalization and ReLU activation), and subsequent maximum-pooling layers to reduce image dimensions and enhance feature extraction. The input layers process 300×300 pixel black-and-white images of individual landslides, with landslide areas marked in white. These features are then categorized into slide or flow through a fully connected layer, followed by a softmax layer that determines the class probability. To enhance the model's ability to generalize, we implemented various data augmentation techniques, including random horizontal reflections, rotations ranging from −90° to 90°, scaling from 80% to 120% of the original size, and translations up to 10 pixels.

The network was trained using a randomly selected half of the 4156 debris slides and 1115 debris flows identified in Casola Valsenio, while the other half was utilized to fine-tune the network's hyperparameters and to test and assess the model's performance. These evaluations showed that the CNN effectively replicates expert classifications of slides and flows. Utilizing the Adam optimizer with an initial learning rate of 0.001 over 100 epochs, the CNN reaches an F1 score of 0.80 on the testing dataset, indicating robust accuracy in terms of both precision and recall. Figure 3.8 displays the confusion matrix obtained for the testing dataset, alongside a selection of landslide images that were correctly and incorrectly classified by the neural network. Of course, as is clearly evident when looking at the false-positive and false-negative cases, the CNN cannot overcome the inherent ambiguity in classifying landslides that fall between slides and flows, particularly those that are only partially fluidized and whose polygon shapes are neither distinctly sub-circular nor clearly elongated. However, by implementing the network across all the polygons, we ensure that the classification criteria agreed upon in Casola Valsenio are consistently applied throughout the entire area.

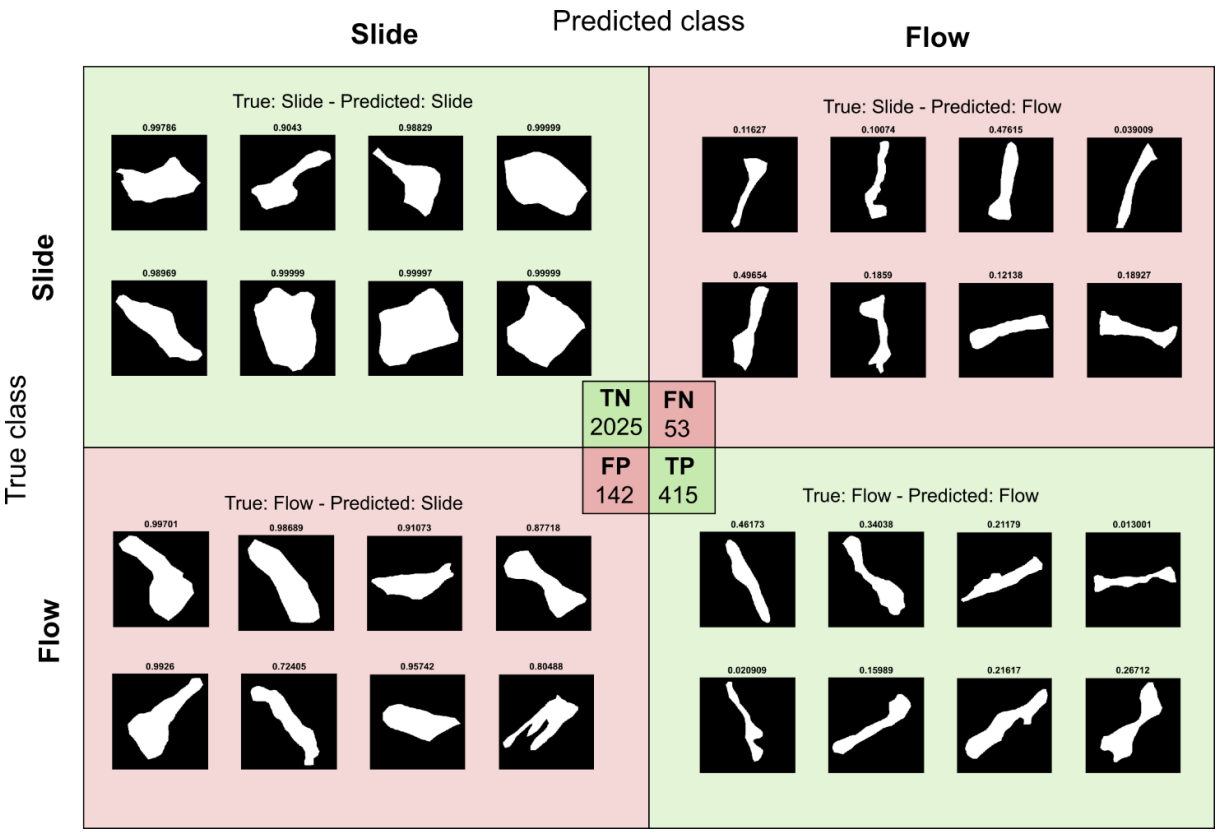


Figure 3.8. Outcomes from the Convolutional Neural Network model applied to differentiate slides from flows using the shape of the polygons. The figure displays the confusion matrix for the testing dataset, which includes 50% of the landslides manually mapped in the Casola Valsenio municipality: TN=True Negative; FN=False Negative; FP=False Positive; TP=True Positive. The small polygons in each category represent example landslides that are correctly (TP, TN) or incorrectly (FP, FN) classified.

Degree of mobility of debris slides

Debris slides were classified by mappers into two categories based on their apparent mobility (Figure 3.5). The class DS1, indicating high-mobility slides, was assigned to slides that showed extensive internal disruption and complete removal of vegetation. The class DS2, denoting low-mobility slides, was assigned to slides with minimal internal deformation and little impact on vegetation. While mappers collectively agreed on these criteria for classifying debris slides, discrepancies arose due to variations in personal judgment. To standardize this assessment, we evaluated the mobility of debris slides by examining the remaining vegetation cover after movement. The green-leaf index (GLI) was used to quantify the amount of green vegetation within a landslide polygon:

$$\mathrm{GLI} = \frac{(2 \cdot \mathrm{Green} - \mathrm{Red} - \mathrm{Blue})}{(2 \cdot \mathrm{Green} + \mathrm{Red} + \mathrm{Blue})} \qquad [4.1]$$

Here, green, red, and blue denote the reflectance values from the respective color bands. The GLI ranges between −1 and 1, with higher values indicating a denser presence of green leaves.

To identify a suitable GLI threshold for distinguishing the two classes, we analyzed the frequency distribution of GLI values for DS1 and DS2 in the Casola Valsenio dataset. As shown in Figure 3.9, a distinct separation is observed in the higher categories: 99% of the 4144 high-mobility debris slides (DS1) have GLI values under 0.08, indicating that they are primarily bare soils with minimal or no vegetation. In contrast, 34% of the 125 low-mobility slides (DS2) exceed this threshold, suggesting the presence of vegetation. Such a threshold therefore allows for an effective classification of DS1, but it risks misclassification of DS2.

The challenge in distinguishing the two classes stems from the inherent subjectivity involved, especially when vegetation is only partially removed. Like the difficulty in differentiating between slides and flows, no automated method can fully address this issue. However, in our case, the occurrence of DS2 is significantly less frequent than that of DS1. Consequently, we have chosen to set a GLI threshold of 0.08, acknowledging that this may lead to some misclassification errors with DS2. This approach classifies nearly intact vegetation slides as low mobility (DS2) and those with partial vegetation as high mobility (DS1). The resulting F1 score is 0.86, and this threshold has proven to be stable whether computed on a randomly selected subsample or a specific segment of the Casola Valsenio area.

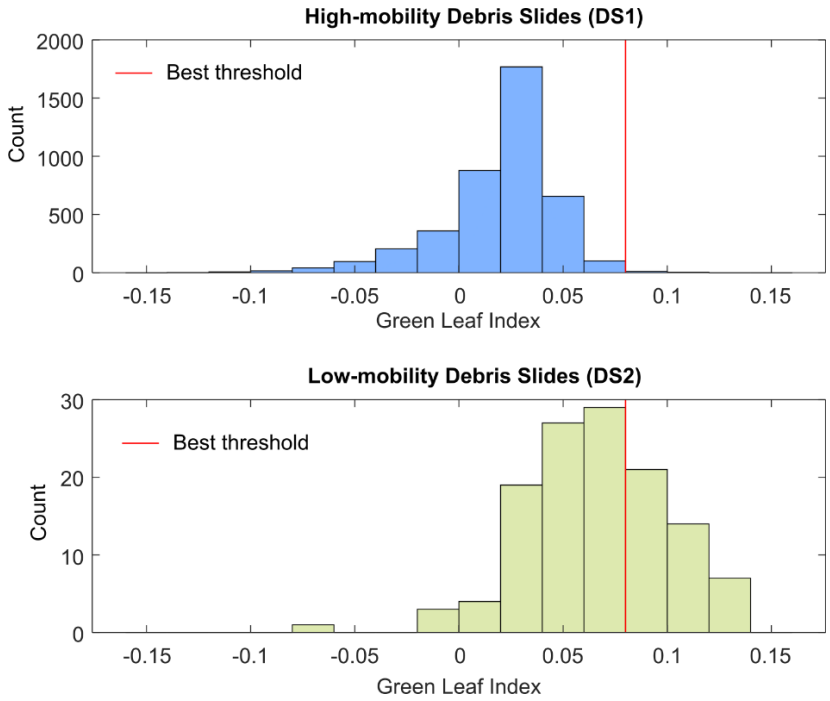


Figure 3.9. Comparison of the frequency distributions of the Green Leaf Index for high-mobility debris slides (upper) and low-mobility debris slides (lower) manually mapped in Casola Valsenio municipality. The red line marks the optimal threshold distinguishing the two landslide types.

Degree of fluidization of debris flows

Debris flows were divided into two distinct classes to highlight differences in fluidization and runout (Figure 3.4). The DF1 category was used for long-runout debris flows, marked by fluidized deposits spreading over relatively flat terrain. Conversely, DF2 was used for debris flows with more limited fluidization, typically confined to steep, forested slopes. Mappers used

these criteria but also looked at factors such as the size of the debris flow, the presence of a channel, or the location of the initiation area. As a result, the manual classification of DF1 and DF2 was notably inconsistent.

The classification problems are evident when attempting to define an automatic standardization procedure. Both DF1 and DF2 exhibit elongated shapes and the absence of vegetation within the landslide areas. Consequently, previous methods that rely on polygon shape or vegetation cover are not applicable. One potential approach could involve using the mean slope of the landslide area, which is generally lower for DF1. However, this metric could introduce bias into the dataset, particularly when comparing the morphological characteristics of the different landslides.

After experimenting with various factors and machine learning techniques, we decided on a simple, reproducible method. Using the Casola Valsenio dataset again, we determined that a reliable indicator of debris mobility is the percentage of the landslide area that extends over non-forested slopes (NFSs). NFSs encompass all slopes lacking forest cover, which, in most cases, are shrub and/or grassy areas, areas with sparse or no vegetation, and agricultural lands. Mappers typically classified debris flows that overrun these areas as DF1. NFS is simply given by

$$\mathrm{NFS} = \frac{\mathrm{A_{NF}}}{\mathrm{A_{Tot}}} \cdot 100 \qquad [4.2]$$

Here, $\mathrm{A_{NF}}$ is the landslide area on non-forested slopes, and $\mathrm{A_{Tot}}$ is the total area of the landslide. $\mathrm{A_{NF}}$ was detected by overlapping the landslide polygon with the soil use coverage SU2014 provided by the Emilia-Romagna region. This coverage was derived from aerial images captured between May and September 2014 using four bands at a 0.5 m resolution and classified according to the Corine Land Cover directive. All the slopes not categorized as 311 (broad-leaved forest), 312 (coniferous forest), or 313 (mixed forest) were identified as non-forested.

In Casola Valsenio, a total of 1053 debris flows were documented. Among these, 471 were notably fluid and mobile (DF1), whereas the remaining 582 exhibited less mobility (DF2). The non-forested-slope (NFS) values varied distinctly between the two classes, with DF1 generally displaying higher NFS values (Figure 3.10). An NFS threshold of 0.3 has proven to be effective in distinguishing between DF1 and DF2: 83% of the DF1 category exceeds this threshold, whereas 82% of the DF2 category falls below it. The corresponding F1 score is 0.82, reflecting a high degree of accuracy.

The harmonization procedure described above resulted in significant modifications to the initial manual classifications. Approximately 50% of the debris slides with limited mobility (DS2) were reclassified as debris slides with high mobility (DS1) due to either heavy or partial clearing of vegetation cover by the movement. About 25% of debris flows (DF1 and DF2) were reclassified as debris slides (DS1 or DS2) due to the limited elongation of the deposit, and about 60% of earth flows (EFs) were reclassified as earth slides (ESs) for the same reason. It is important to stress that the harmonization process should not be viewed as an automatic

classification but rather as an effort to apply consistent classification criteria across the entire area.

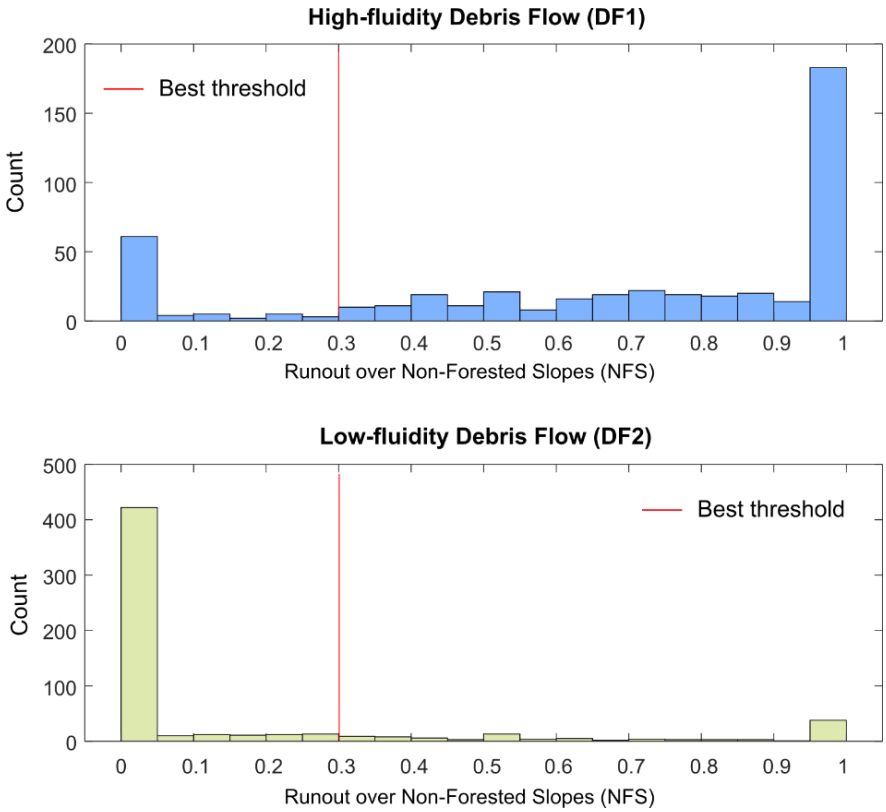


Figure 3.10. Comparison of debris flows with high fluidity (upper) and low fluidity (lower) in the Casola Valsenio municipality, analyzed through the ratio of runout over non-forested slopes (NFS). The red line indicates the optimal threshold for distinguishing between the two types of landslides.

3.6 The landslide inventory dataset

The landslide inventory for the 2023 Emilia-Romagna event includes 80,997 landslide polygons, each categorized according to the classification described in Sect. 3.2. The inventory encompasses landslides triggered by the combined rainfall events of 1–3 and 15–16 May 2023, without distinguishing between the two events. Differentiation between the events is feasible only in specific small areas where high-resolution images were available after the first rainfall; however, this distinction is not included in the current dataset.

Figure 3.11 provides an overview of the inventory, showing landslide points (Fig. 3.11a) and a kernel density map (Fig. 3.11b). Notably there is a strong, though not perfect, correlation between cumulative rainfall and landslide density. In the eastern part of the region, known as Romagna, the 300 mm rainfall isohyet roughly outlines the area where landslide density exceeds 40 landslides per square kilometre. In contrast, the western part of the region, known as Emilia, has a landslide density below 40 landslides per square kilometre despite receiving the same amount of rainfall. This difference can be attributed to the distinct geological settings of the two areas. As shown in Fig. 3.1b, the Romagna region is primarily characterized by a Miocene flysch (Marnoso-Arenacea Formation, unit 7), which results in steep slopes and

coarse-grained weathered soil. Meanwhile, the Emilia region has a more complex geological setting, including extensive areas of fine-grained rocks that responded less intensely to these rainfall events.

In the Romagna region, landslide density reached an impressive level of over 250 landslides per square kilometre. The zone most heavily affected, with more than 40 landslides per square kilometre, stretches across roughly 800 km² and covers the outer sector of the Marnoso-Arenacea Formation (red area in Figure 3.12). About 64% of the landslides occurred within this zone. The landslide index, which is the ratio of landslide area to total area, reaches 20%–25% in this area. These figures are particularly significant considering the fact that they represent the percentage of the area destabilized during a single episode.

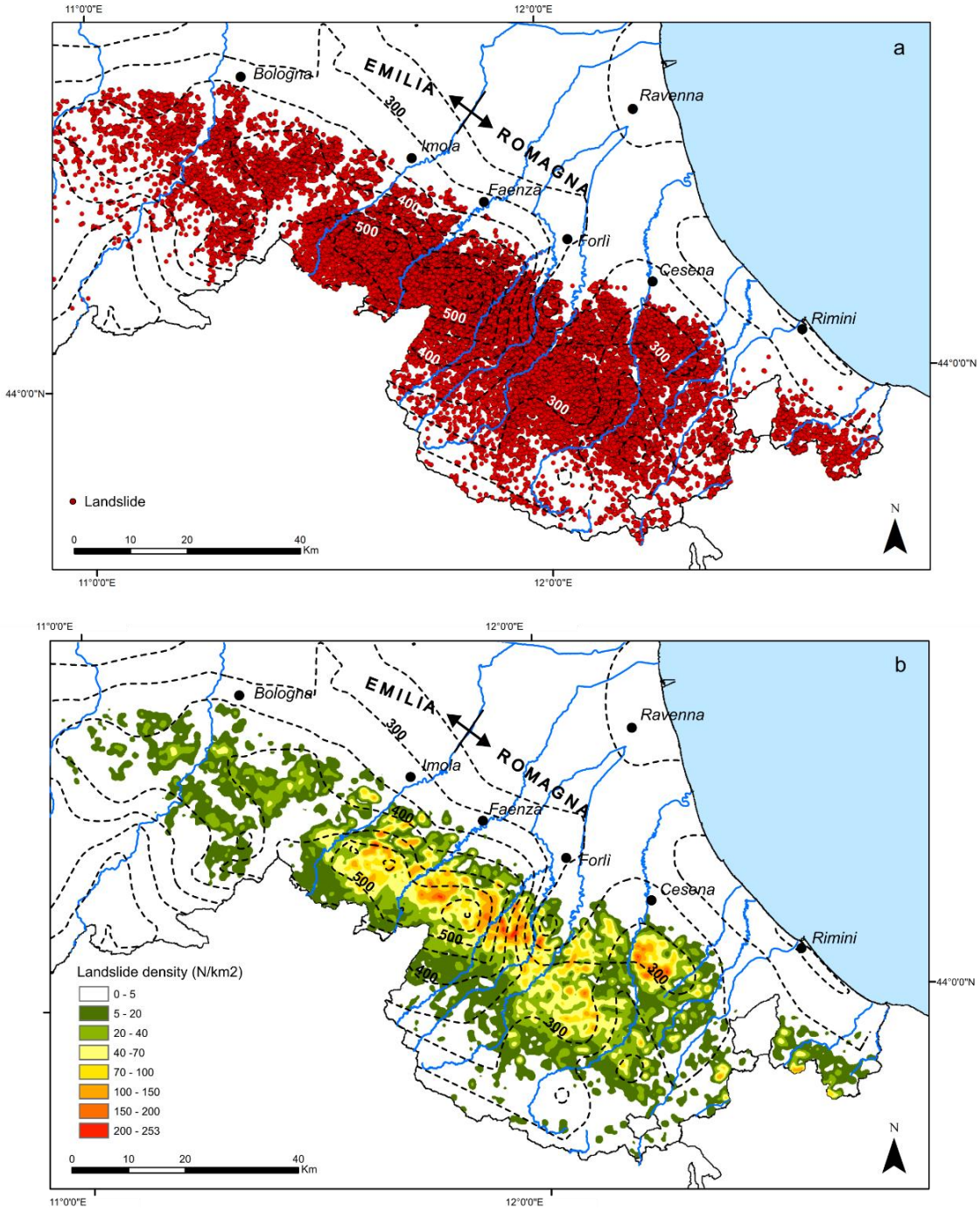


Figure 3.11. ***a****) Map showing the distribution of the 80997 landslides triggered by the May 2023 event in the Emilia-Romagna region, manually mapped and represented as individual points.* ***b****) Density map calculated as the number of landslides per 1 km² cell. The dashed lines in both images (a and b) represent the cumulative rainfall isohyete of the combined events.*

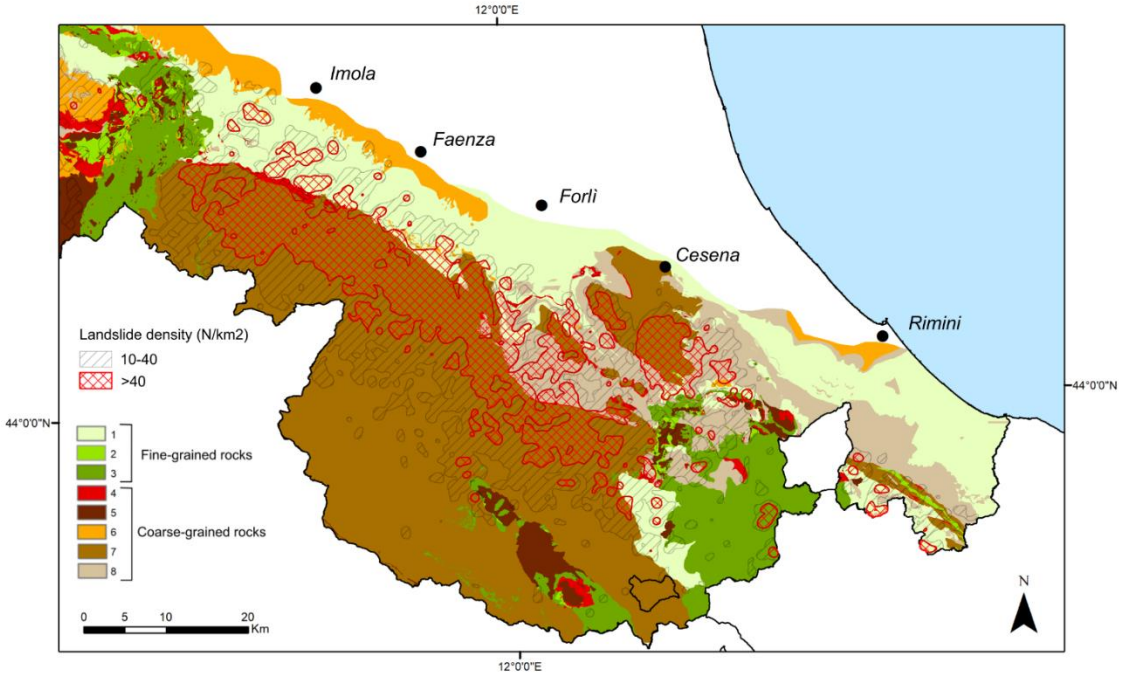


Figure 3.12. Detail of the Romagna area showing two classes of landslide density (between 10 and 40 landslide/km2 and more than 40 landslide/km^2) with the eight lithological units in background.

A deeper examination of the harmonized inventory underscores the occurrence and main features of various landslide types, providing insights into their spatial distribution and contributing factors. Figure 3.13 and Figure 3.14 display several statistical details about the count, dimensions, and slope angles of these landslides. The results from these diagrams are discussed below, enhanced with additional observations from manual mapping and field surveys.

Debris slides (DSs) represent 66% of all landslides by number and 49% by area, marking them as the most common type triggered by the event (Fig. 3.13). These landslides were generally small to very small (area of less than 1,000 m², Fig. 3.14) and typically occurred on steep slopes with inclines exceeding 25–30°. In the region, many of these slopes are covered with forests as they are unsuitable for farming; hence, while root reinforcement and rainwater interception by the tree canopies exist, they were insufficient to prevent these failures. Approximately 94% of the slides were fast-moving and became liquefied after traveling a short distance (DS1). A minor fraction (6%) moved as a coherent mass, showing considerably less internal disruption (DS2). Slides with high mobility caused extensive damage to roads, buildings, and infrastructure and transported large amounts of debris and wood into rivers. Conversely, low-mobility slides predominantly occurred on milder slopes and near roadways. These slides might indicate early-stage slides that had not fully developed, secondary failures behind landslide headscarps, or slides involving rotational movements.

Debris flows (DFs) are the second most frequent type of landslide, constituting 15% of the total count and 29% by area (Fig. 3.13). DFs were consistently initiated by debris slides on slopes that are generally steeper than 25–30°. Currently, it remains difficult to ascertain why some

debris slides transformed into debris flows while others did not. However, it is evident that the predominant failure mechanism during the May 2023 event was shallow sliding of the weathered soil cover. Together, debris flows and slides represent 81% of the landslides catalogued in the inventory. After the initial failure, debris flows generally traveled without following predefined channels and cleared the vegetation, forming straight, elongated rectangular shapes. About 50% of these flows had relatively limited runouts halting along steep slopes (DF2), while the rest displayed significantly higher mobility, spreading extensively across gentle slopes due to complete fluidization of the material (DF1). A notable feature of the debris flows triggered during the event, especially DF1, was their relatively low destructive power. In many cases, these flows approached buildings and roads without causing substantial damage and spread over grassy fields without harming the vegetation. The limited damage caused by these flows can be linked to their composition, primarily liquefied sand and silt without large cobbles or boulders. This composition enabled them to flow downslope as a dense slurry without a destructive bouldery front. The typical dimensions of debris flows range from very small to medium (Fig. 3.14).

Rock-block slides (RSs) constitute less than 2% by number and 5% by area of all landslides (Fig. 3.13), but they left the most profound impression on the public and media. These landslides occurred on homoclinal slopes within the Marnoso-Arenacea Formation (lithological unit 7) and developed as planar slides along bedding planes that aligned with the slope. The thickness of the displaced rock mass varied from about 2 to over 30 m, and several slides extended over areas larger than 10 ha. Compared to debris flows and debris slides, rock-block slides affected more gentle slopes, typically less than 15°, and were bigger in size (class medium, Fig. 3.14). Their large volume and high velocity and the fact that they occurred on sloping lands that were heavily urbanized and farmed made these landslides a major concern during the event. All rock slides initially traveled as coherent rock blocks, moving translationally for several to tens of metres. However, some slides disintegrated during their motion, transforming into rapid flows of debris and fragmented rock. This disintegration typically occurred when the displaced blocks tumbled down an existing scarp or struck a lateral slope, causing the material to break apart. These fragmented rock-block slides were highly mobile and covered long distances.

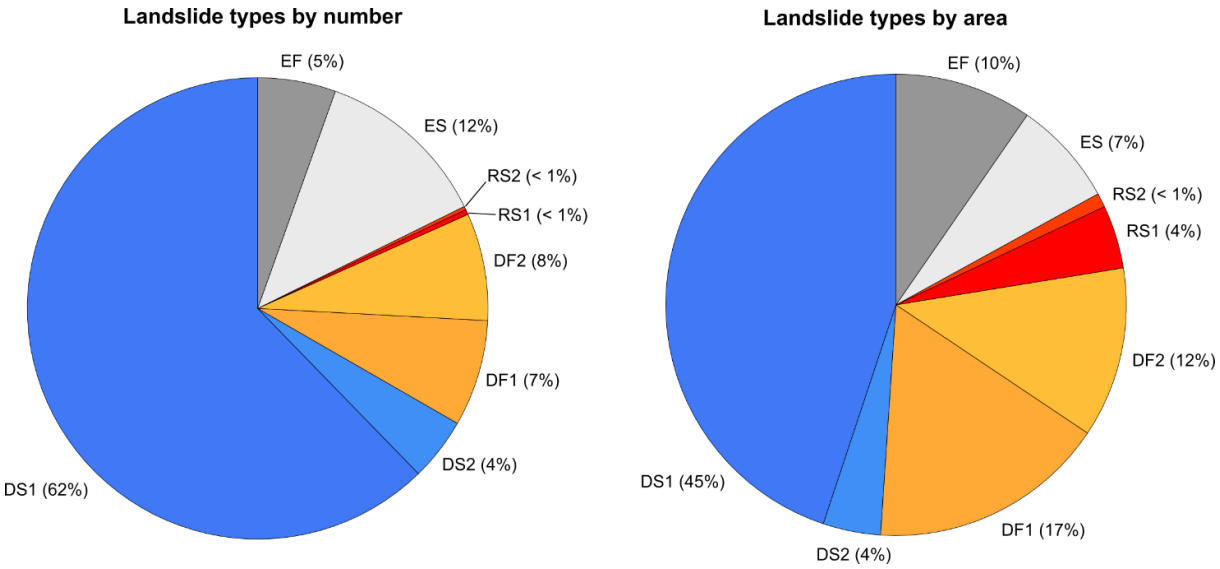


Figure 3.13. Pie charts showing the percentage of landslide types by number (left) and by area (right).

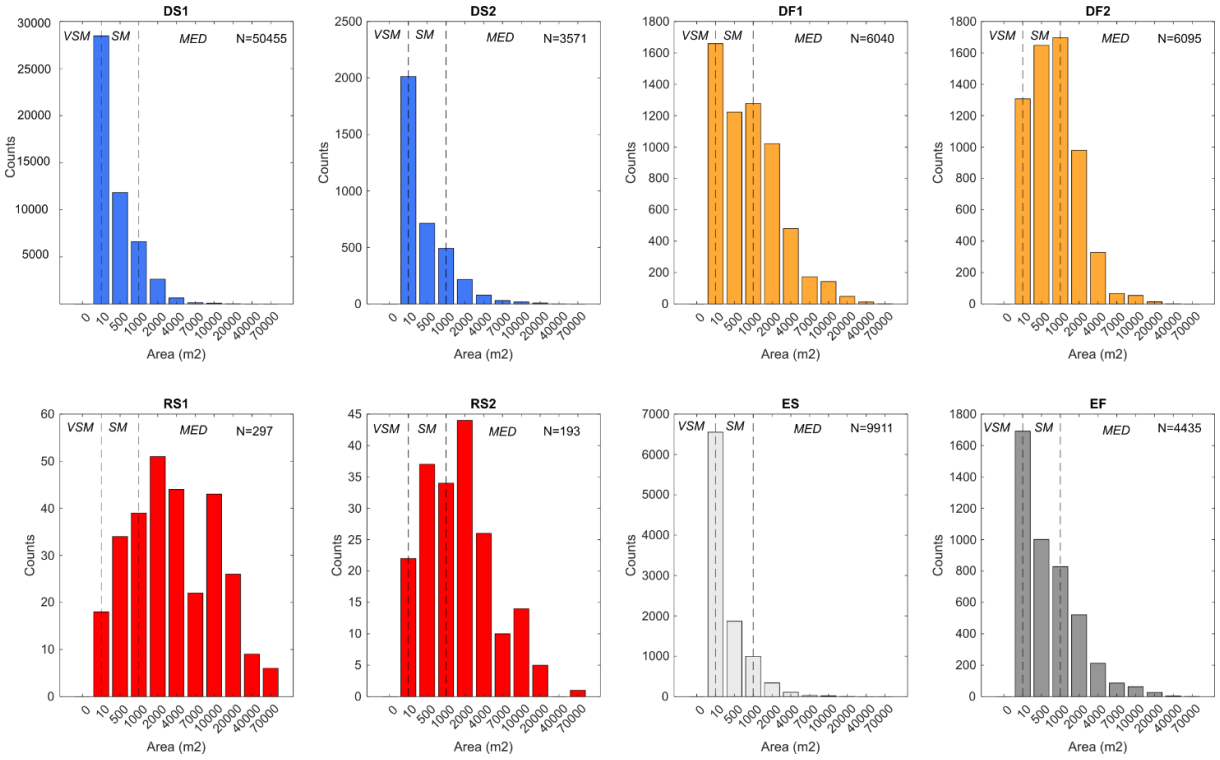


Figure 3.14. Frequency histograms depicting size of various landslide types. The classes labeled at the top of each chart (VSM=very small; SM=small; MED=medium) correspond to the size classification proposed by McColl and Cook (2024).

Earth slides and earth flows accounted for 12% and 5% of the total number of landslides and 7% and 10% of the total area affected, respectively (Fig. 3.13). These landslides predominantly occurred within fine-grained units, specifically the Pliocene clays (unit 1) and Cretaceous clay shales (unit 3), as illustrated in Fig. 3.1b. These regions generally experienced fewer landslides, with less severe impacts compared to areas dominated by coarse-grained rocks in the southern parts. The distinct patterns of landslides in areas with coarse- and fine-grained lithological units

are clearly illustrated in the sample maps of Figure 3.15. While landslides are commonly found on steep slopes in both cases, the coarse-grained units also show that even gentle slopes are impacted by extensive long-runout debris flows (DF1 in Fig. 3.15a) and rock-block slides (RS1). Conversely, in the fine-grained units (Fig. 3.15b), earth flows and earth slides are primarily concentrated in the badlands areas, with gentle slopes remaining largely unaffected. Moreover, in Emilia-Romagna, it is typical for earth slides and especially for earth flows to occur repeatedly at the same locations as reactivations of dormant landslides. This recurring pattern was evident during the May 2023 event, with most landslides appearing as reactivations of previously known landslides, which were already familiar to the local communities. In contrast, nearly all the landslides in the coarse-grained units – including debris slides, debris flows, and rock-block slides – represented first-time failures and occurred unexpectedly on slopes previously free of documented landslides.

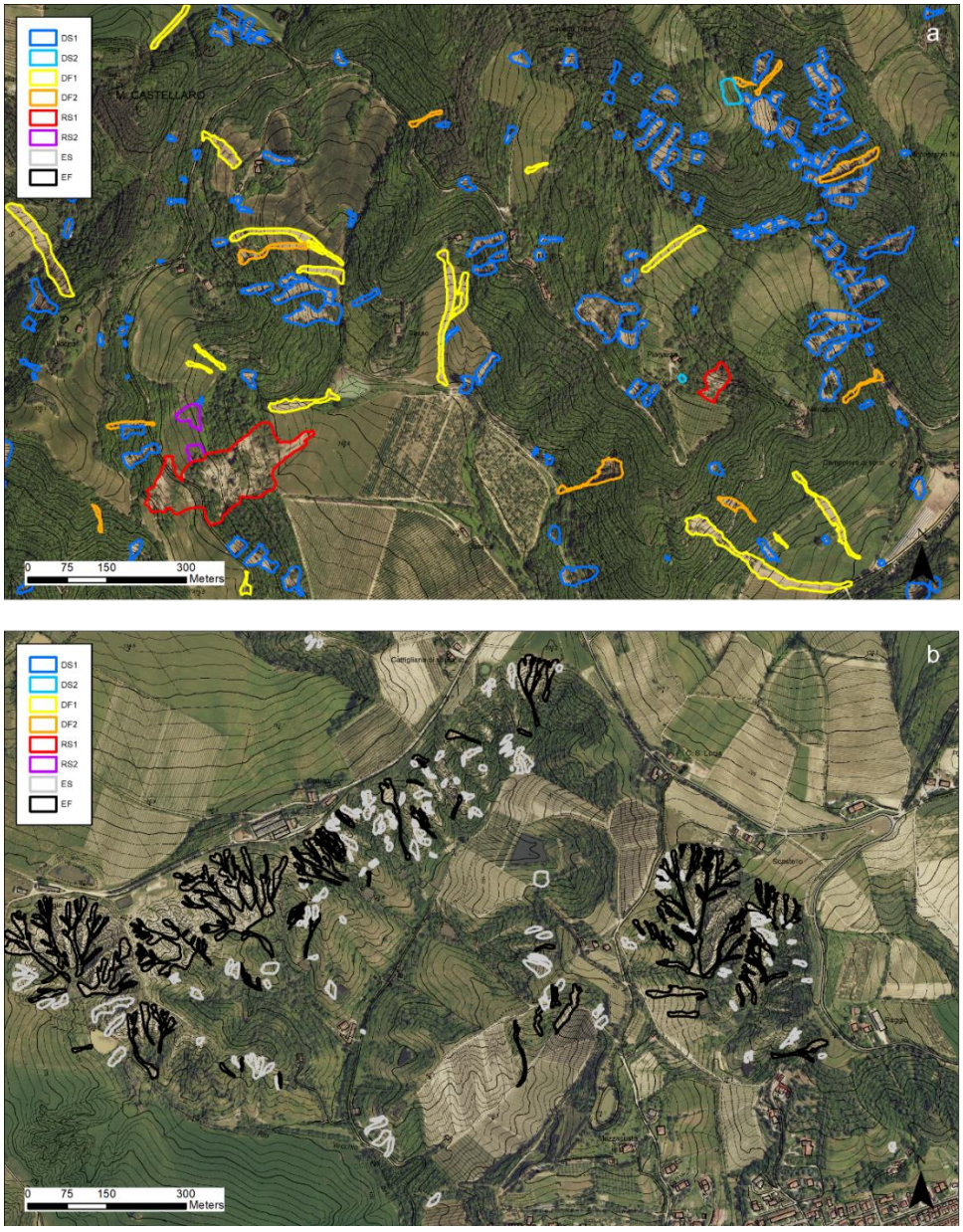


*Figure 3.15. Example images from the landslide inventory showing two representative areas in coarse-grained units (**a**) and in fine-grained units (**b**).*

3.7 Limitations and future updates

The landslide inventory was carried out with great care under the constraints imposed by the emergency situation and ongoing recovery efforts. Utilizing high-resolution 0.2 m imagery and an automated harmonization process facilitated a detailed and consistent record across the region. However, as is typical with any expert-driven inventory, errors and inconsistencies are unavoidable. Missed landslides, or false negatives, are likely to occur in shadowed areas like river gorges or steep slopes, as well as in forested areas where landslides have occurred without clearing the vegetation. Additionally, landslides with minimal ground displacement, although clearly visible on-site, may not be discernible in aerial images and thus may remain undetected. False positives – areas mistakenly identified as landslides – are also possible but are expected to be fewer. These may include anthropogenic debris accumulations, excavation activities, or plowed fields that alter the soil surface in a manner similar to landslides or landslides that happened after the pre-event images from the period of April–July 2020 but before the May 2023 event. Nevertheless, we estimate that the combined total of missed or incorrectly identified landslides might constitute less than 1% of the total inventory.

The primary limitation of the inventory likely lies in the accuracy of the landslide boundaries. Not all mappers across the area had sufficient time to delineate the landslide polygons with high-resolution detail, resulting in some boundaries appearing jagged and imprecise upon closer inspection. The data quality procedures brought this issue to light, but redrawing all the rough-edged polygons would be excessively time-consuming. Given that the locations of these landslides are accurate, we chose to publish and make available the inventory in its current state and defer any refinements to future updated versions that, besides refining polygons, will also incorporate changes recommended by local authorities, which are currently being considered. In April 2024, the Emilia-Romagna region shared the landslide inventory with all municipalities affected by the event, requesting feedback on any overlooked landslides. This information is now being gathered, and a first update is scheduled for completion by the end of 2025. Initial feedback primarily concerns small landslides that caused damage to private or public properties but were not detected in aerial photographs due to minimal displacement. These new landslides will be included in version 2 of the inventory, which will be available in the same Zenodo repository.

3.8 Data availability

The landslide inventory is freely accessible in the Zenodo repository (https://doi.org/10.5281/zenodo.13742643, Pizziolo et al., 2024). The dataset is available as an ESRI (Environmental Systems Research Institute) shapefile and is compatible with GIS software. The shapefile encompasses several attributes: polygon ID (IDC), landslide type as manually classified by the operator (ClassMan), geological unit of the polygon's centroid (Lito), green-leaf index (GLI), percentage of deposit over non-forested slopes (NFS), and landslide type after applying the harmonization algorithm (ClassNew).

3.9 Concluding remarks

The dataset of the May 2023 Emilia-Romagna event encompasses more than 80,000 rainfall-induced landslides (mostly first-failure) distributed over an area of more than 6,000 km², with density reaching as high as 200 landslides per square kilometre. Despite some inherent limitations and potential areas for improvement in the dataset, we believe that our landslide inventory offers significant value to the scientific community and to the involved institutions for several reasons.

Firstly, it documents the response of a large area to an exceptional meteorological event, likely linked to ongoing climate change. This can support the scientific community in proving that multiple occurrences of rainfall-related landslides are likely to become more frequent in the coming years and can make decision-makers more aware of the fact that even slopes that have been unaffected by landslides in the past cannot be considered to be free of risk for the future.

Secondly, the Emilia-Romagna region's relatively straightforward geological framework makes it ideal for conducting geospatial analyses of landslide susceptibility and for proving that these analyses can be adopted to support land use planning in addition to landslide inventories. Actually, the Emilia-Romagna region's geoportal provides free access to an extensive range of spatial data, including DEMs, lithology, land use, and rainfall data, all of which can be integrated into our landslide map to test both traditional and machine-learning-based predictive tools.

Thirdly, the predominance of shallow planar failures in this event provides an excellent case for testing physically based slope stability models and for highlighting the relevance of such types of landslides in the study area so as to promote a much more careful evaluation of the possible impact of such phenomena on existing infrastructure networks and for designing new assets.

In conclusion, we warmly invite interested colleagues to contact us with any questions and/or specific needs or to initiate a collaborative research effort that could transform a tragic event into an opportunity to enhance our understanding of landslide risk assessment.

3.10 References chapter 3

AGSS-RER (Area Geologia, Suoli e Sismica, Regione Emilia-Romagna). (1986). Carta Geologica dell'Appennino emiliano-romagnolo, scala 1:10.000. Regione Emilia-Romagna. Retrieved March 3, 2025, from https://geoportale.regione.emilia-romagna.it/

Amatya, P., Kirschbaum, D., Stanley, T., & Tanyas, H. (2021). Landslide mapping using object-based image analysis and open source tools. Engineering Geology, 282, 106000. https://doi.org/10.1016/j.enggeo.2021.106000

Ardizzone, F., Bucci, F., Cardinali, M., Fiorucci, F., Pisano, L., Santangelo, M., & Zumpano, V. (2023). Geomorphological landslide inventory map of the Daunia Apennines, southern Italy. Earth System Science Data, 15, 753–767. https://doi.org/10.5194/essd-15-753-2023

Aufli˘c, M. J., Bezak, N., Šegina, E., Frantar, P., Gariano, S. L., Medved, A., & Peternel, T. (2023). Climate change increases the number of landslides at the juncture of the Alpine, Pannonian and Mediterranean regions. Scientific Reports, 13, 23085. https://doi.org/10.1038/s41598-023-50314-x

Berti, M., & Simoni, A. (2010). Field evidence of pore pressure diffusion in clayey soils prone to landsliding. Journal of Geophysical Research, 115, F03031. https://doi.org/10.1029/2009JF001463

Berti, M., & Simoni, A. (2012). Observation and analysis of near-surface pore pressure measurements in clay shale slopes. Hydrological Processes, 26, 2187–2205. https://doi.org/10.1002/hyp.9316

Berti, M., Martina, M. L. V., Franceschini, S., Pignone, A., Simoni, A., & Pizziolo, M. (2012). Probabilistic rainfall thresholds for landslide occurrence using a Bayesian approach. Journal of Geophysical Research, 117, F04006. https://doi.org/10.1029/2012JF002367

Bertolini, G., Guida, M., & Pizziolo, M. (2005). Landslides in Emilia-Romagna region (Italy): Strategies for hazard assessment and risk management. Landslides, 2, 302–312. https://doi.org/10.1007/s10346-005-0020-1

Bertolini, G., Corsini, A., & Tellini, C. (2017). Fingerprints of large-scale landslides in the landscape of the Emilia Apennines. In M. Soldati & M. Marchetti (Eds.), Landscapes and landforms of Italy. World geomorphological landscapes (pp. 215–224). Springer. https://doi.org/10.1007/978-3-319-26194-2_18

Bhuyan, K., Tanya¸s, H., Nava, L., Puliero, S., Meena, S. R., Floris, M., van Westen, C., & Catani, F. (2023). Generating multi-temporal landslide inventories through a general deep transfer learning strategy using HR EO data. Scientific Reports, 13, 162. https://doi.org/10.1038/s41598-022-27352-y

Brath, A., Casagli, N., Marani, M., Mercogliano, P., & Motta, R. (2023). Rapporto della Commissione tecnico-scientifica istituita con deliberazione della Giunta Regionale n. 984/2023 e determinazione dirigenziale 14641/2023, al fine di analizzare gli eventi meteorologici estremi del mese di maggio 2023 (147 pp.). Regione Emilia-Romagna. Retrieved March 3, 2025, from https://www.regione.emilia-romagna.it/alluvione/rapporto-della-commissione-tecnico-scientifica

Cardinali, M., Galli, M., Guzzetti, F., Ardizzone, F., Reichenbach, P., & Bartoccini, P. (2006). Rainfall-induced landslides in December 2004 in south-western Umbria, central Italy: Types,

extent, damage and risk assessment. Natural Hazards and Earth System Sciences, 6, 237–260. https://doi.org/10.5194/nhess-6-237-2006

Catani, F. (2021). Landslide detection by deep learning of non-nadiral and crowdsourced optical images. Landslides, 18, 1025–1044. https://doi.org/10.1007/s10346-020-01513-4

Ciccarese, G., Mulas, M., Alberoni, P. P., Truffelli, G., & Corsini, A. (2020). Debris flow rainfall thresholds in the Apennines of Emilia-Romagna (Italy) derived by the analysis of recent severe rainstorm events and regional meteorological data. Geomorphology, 358, 107097. https://doi.org/10.1016/j.geomorph.2020.107097

Corominas, J., van Westen, C., Frattini, P., Cascini, L., Malet, J., Fotopoulou, S., Catani, F., Van Den Eeckhaut, M., Mavrouli, O., Agliardi, F., Pitilakis, K., Winter, M., Pastor, M., Ferlisi, S., Tofani, V., Hervás, J., & Smith, J. (2014). Recommendations for the quantitative analysis of landslide risk. Bulletin of Engineering Geology and the Environment, 73, 209–263. https://doi.org/10.1007/s10064-013-0538-8

Corsini, A., Ciccarese, G., Diena, M., Alberoni, P. P., & Amorati, R. (2017). Debris flows in Val Parma and Val Baganza (Northern Apennines) during the 13 October 2014 alluvial event in Parma province (Italy). Italian Journal of Engineering Geology and Environment, 35, 29–38. https://doi.org/10.4408/IJEGE.2017-01.S-03

Crozier, M. (2005). Multiple-occurrence regional landslide events in New Zealand: Hazard management issues. Landslides, 2, 247–256. https://doi.org/10.1007/s10346-005-0019-7

Cruden, D. M., & Varnes, D. J. (1996). Landslide types and processes. In Transportation Research Board, U.S. National Academy of Sciences, Special Report (Vol. 247, pp. 36–75). https://doi.org/10.17226/11057

ESA (European Space Agency). (2024). Copernicus Global Digital Elevation Model. OpenTopography. https://doi.org/10.5069/G9028PQB

Fell, R., Corominas, J., Bonnard, C., Cascini, L., Leroi, E., & Savage, W. Z. (2008). Guidelines for landslide susceptibility, hazard and risk zoning for land use planning. Engineering Geology, 102, 85–98. https://doi.org/10.1016/j.enggeo.2008.03.022

Ferrario, M. F., & Livio, F. (2024). Rapid mapping of landslides induced by heavy rainfall in the Emilia-Romagna (Italy) region in May 2023. Remote Sensing, 16, 122. https://doi.org/10.3390/rs16010122

Gaidzik, K., & Ramírez-Herrera, M. T. (2021). The importance of input data on landslide susceptibility mapping. Scientific Reports, 11, 19334. https://doi.org/10.1038/s41598-021-98830-y

Galli, M., Ardizzone, F., Cardinali, M., Guzzetti, F., & Reichenbach, P. (2008). Comparing landslide inventory maps. Geomorphology, 94, 268–289. https://doi.org/10.1016/j.geomorph.2006.09.023

Gariano, S. L., & Guzzetti, F. (2016). Landslides in a changing climate. Earth-Science Reviews, 162, 227–252. https://doi.org/10.1016/j.earscirev.2016.08.011

Guzzetti, F., Cardinali, M., Reichenbach, P., Cipolla, F., Sebastiani, C., Galli, M., & Salvati, P. (2004). Landslides triggered by the 23 November 2000 rainfall event in the Imperia Province, Western Liguria, Italy. Engineering Geology, 73, 229–245. https://doi.org/10.1016/j.enggeo.2004.01.006

Guzzetti, F., Mondini, A. C., Cardinali, M., Fiorucci, F., Santangelo, M., & Chang, K. T. (2012). Landslide inventory maps: New tools for an old problem. Earth-Science Reviews, 112, 42–66. https://doi.org/10.1016/j.earscirev.2012.02.001

Handwerger, A. L., Fielding, E. J., Sangha, S. S., & Bekaert, D. P. S. (2022). Landslide sensitivity and response to precipitation changes in wet and dry climates. Geophysical Research Letters, 49, e2022GL099499. https://doi.org/10.1029/2022GL099499

Hao, L., Rajaneesh, A., van Westen, C., Sajinkumar, K. S., Martha, T. R., Jaiswal, P., & McAdoo, B. G. (2020). Constructing a complete landslide inventory dataset for the 2018 monsoon disaster in Kerala, India, for land use change analysis. Earth System Science Data, 12, 2899–2918. https://doi.org/10.5194/essd-12-2899-2020

Hungr, O., Leroueil, S., & Picarelli, L. (2014). The Varnes classification of landslide types, an update. Landslides, 11, 167–194. https://doi.org/10.1007/s10346-013-0436-y

Jaboyedoff, M., Oppikofer, T., Abellan, A., & Pedrazzini, A. (2012). Use of LiDAR in landslide investigations: A review. Natural Hazards, 61, 5–28. https://doi.org/10.1007/s11069-010-9634-2

Luetzenburg, G., Svennevig, K., Bjørk, A. A., Keiding, M., & Kroon, A. (2022). A national landslide inventory for Denmark. Earth System Science Data, 14, 3157–3165. https://doi.org/10.5194/essd-14-3157-2022

Malamud, B. D., Turcotte, D. L., Guzzetti, F., & Reichenbach, P. (2004). Landslide inventories and their statistical properties. Earth Surface Processes and Landforms, 29, 687–711. https://doi.org/10.1002/esp.1064

McColl, S. T., & Cook, S. J. (2024). A universal size classification system for landslides. Landslides, 21, 111–120. https://doi.org/10.1007/s10346-023-02131-6

Notti, D., Cignetti, M., Godone, D., Cardone, D., & Giordan, D. (2024). The unsupervised shallow landslide rapid mapping: PANDA method applied to severe rainfalls in northeastern Apennine (Italy). International Journal of Applied Earth Observation and Geoinformation, 129, 103806. https://doi.org/10.1016/j.jag.2024.103806

Piacentini, D., Troiani, F., Daniele, G., & Pizziolo, M. (2018). Historical geospatial database for landslide analysis: The Catalogue of Landslide Occurrences in the Emilia-Romagna Region (CLOCkER). Landslides, 15, 811–822. https://doi.org/10.1007/s10346-018-0962-8

Pittau, S., Rossi, M., Llena, M., & Brardinoni, F. (2024). Evaluating historical, basin-wide landslide activity in a context of land abandonment and climate change: Effects of landslide visibility and temporal resolution. Geomorphology, 452, 109122. https://doi.org/10.1016/j.geomorph.2024.109122

Pizziolo, M., Berti, M., Scaroni, M., Generali, M., Critelli, V., Mulas, M., Tondo, M., Lelli, F., Fabbiani, C., Ronchetti, F., Ciccarese, G., Dal Seno, N., Ioriatti, E., Rani, R., Zuccarini, A., Simonelli, T., & Corsini, A. (2024). RER2023: The landslide inventory dataset of the May 2023 Emilia-Romagna event – Version 1. Zenodo [Data set]. https://doi.org/10.5281/zenodo.13742643

Rosi, A., Tofani, V., Tanteri, L., Tacconi Stefanelli, C., Agostini, A., Catani, F., & Casagli, N. (2018). The new landslide inventory of Tuscany (Italy) updated with PS-InSAR: Geomorphological features and landslide distribution. Landslides, 15, 5–19. https://doi.org/10.1007/s10346-017-0861-4

Rossi, F., Witt, A., Guzzetti, F., Malamud, B. D., & Peruccacci, S. (2012). Analysis of historical landslide time series in the Emilia-Romagna region, northern Italy. Earth Surface Processes and Landforms, 35, 1123–1137. https://doi.org/10.1002/esp.1858

Santangelo, M., Althuwaynee, O., Alvioli, M., Ardizzone, F., Bianchi, C., Brunetti, M. T., Bucci, F., Cardinali, M., Donnini, M., Esposito, G., Gariano, S. L., Grita, S., Marchesini, I., Melillo, M., Peruccacci, S., Salvati, P., Yazdani, M., & Fiorucci, F. (2023). Inventory of landslides triggered by an extreme rainfall event in Marche-Umbria, Italy, on 15 September 2022. Scientific Data, 10, 1–11. https://doi.org/10.1038/s41597-023-02336-3

Scorpio, V., Crema, S., Marra, F., Righini, M., Ciccarese, G., Borga, M., Cavalli, M., Corsini, A., Marchi, L., Surian, N., & Comiti, F. (2018). Basin-scale analysis of the geomorphic effectiveness of flash floods: A study in the northern Apennines (Italy). Science of the Total Environment, 640–641, 337–351. https://doi.org/10.1016/j.scitotenv.2018.05.252

Soeters, R., & Van Westen, C. J. (1996). Slope instability recognition, analysis and zonation. In A. K. Turner & R. L. Schuster (Eds.), Landslide types and processes. Transportation Research Board, National Research Council, National Academy Press. https://doi.org/10.17226/11057

Steger, S., Brenning, A., Bell, R., & Glade, T. (2017). The influence of systematically incomplete shallow landslide inventories on statistical susceptibility models and suggestions for improvements. Landslides, 14, 1767–1781. https://doi.org/10.1007/s10346-017-0820-0

Van Den Eeckhaut, M., & Hervás, J. (2012). State of the art of national landslide databases in Europe and their potential for assessing susceptibility, hazard and risk. Geomorphology, 139–140, 545–558. https://doi.org/10.1016/j.geomorph.2011.12.006

Zêzere, J. L. (2002). Landslide susceptibility assessment considering landslide typology: A case study in the area north of Lisbon (Portugal). Natural Hazards and Earth System Sciences, 2, 73–82. https://doi.org/10.5194/nhess-2-73-2002

Zieher, T., Perzl, F., Rössel, M., Rutzinger, M., Meißl, G., Markart, G., & Geitner, C. (2016). A multi-annual landslide inventory for the assessment of shallow landslide susceptibility – Two test cases in Vorarlberg, Austria. Geomorphology, 259, 40–54. https://doi.org/10.1016/j.geomorph.2016.02.008

4. Chapter 4: Transformer, more than meets the eye - a deep learning approach to integrate rainfall time series in multi-type landslide probability modelling

4.1 Preface

The preceding Chapter 3 provided the landslide foundation for a detailed susceptibility analysis. In this chapter, in addition to the spatial information, the contribution of the May 2023 triggering events in Emilia-Romagna is incorporated, thereby moving from a static susceptibility framework to a dynamic landslide probability assessment.

The chapter focuses on estimating the contribution of rainfall as dynamic, continuous time-series predictors to landslide probability, using a Transformer Neural Network (TNN) coupled with a Dense Neural Network (DNN) for static predictors. The modeling is performed separately for the major landslide categories identified in Chapter 3, producing landslide-type-specific probabilities and enabling a more comprehensive analysis of the region's hydrological hazard.

To interpret model behavior, a detailed evaluation is conducted using multiple performance metrics and the SHAP-based Expected Gradients algorithm, which reveals the temporal and spatial influence of rainfall on different landslide types.

Although the current analysis allows retrieval of event-specific probabilities only, it establishes a foundation for generalized, spatiotemporal landslide forecasting model if trained with multi-landslide inventories and associated rainfall time series. Such future model could simulate different rainfall scenarios, providing landslide probability for specific rainfall conditions supporting the land use planning, alongside with a regional, detailed, real-time early-warning system.

The chapter is based on a paper currently under review on *Geoscience Frontiers* journal:

Rani R.[1*], Dahal A.[2], Berti M.[1], Lombardo L.[2], 2025. Transformer, more than meets the eye: A deep learning approach to integrate rainfall time-series in multi-type landslide probability modelling.

[1] Department of Biological, Geological, and Environmental Sciences, University of Bologna, Bologna, Italy.

[2]University of Twente, ITC, Applied Earth Sciences, Enschede, Netherlands

4.2 Abstract

Mountainous and hilly regions are frequently affected by landslides, leading to severe damage and loss of life. Different failure mechanisms have distinct geomorphological characteristics and impacts, requiring type-specific treatments. Moreover, incorporating triggering factors, particularly rainfall, is essential to shift from static susceptibility assessments to dynamic, event-specific probability models. While conventional data-driven methods typically use scalar representations of rainfall (e.g., cumulative values over fixed windows), We propose a more comprehensive approach using rainfall as a continuous time series. In this study, we model landslide-type-specific probabilities in the Emilia-Romagna region, which experienced two extreme rainfall events in May 2023, 14 days apart, with a combined return period exceeding 500 years. These events triggered over 80,000 mapped landslides, classified into five types. The 11,670 km² study area was divided into slope units (SUs), with landslide presence labeled by initiation area–SU intersections. Rainfall time series were generated for each SU via spatial interpolation using a linear radial basis function of 188 rain gauges at daily (31 days) and hourly (744 hours) intervals. We embedded the rainfall sequences into a Transformer Neural Network (TNN), coupled with a Dense Neural Network (DNN) for static predictors, and tested various model configurations. Performance evaluation on a 30% test set and stratified cross-validation showed strong results (AUC > 0.90; F1 from 0.75 to 0.20, depending on landslide type), with daily time series consistently performing best. To interpret model behavior, we applied the SHAP-based Expected Gradients algorithm, revealing the temporal and spatial influence of rainfall. Though event-specific, the study sets the foundation for generalized, spatiotemporal landslide forecasting models.

4.3 Introduction

Landslides are a threat in any mountainous landscape, being regularly responsible for large financial (Marín-Rodríguez et al., 2024; Fang et al., 2025) and life (Froude and Petley, 2018; Fidan et al., 2024) losses. They inherently constitute a multi-hazard phenomenon, being often associated to triggers that give raise to a wide range of processes, each of which can simultaneously affect the same area (e.g., Görüm et al., 2025) or cascade through time in a chain of inter-dependent events (Fan et al., 2019). The instability affecting a slope can be the result of natural forcings such as rainfall (Crosta and Frattini, 2008), earthquakes (Wasowski et al., 2011), snowmelt (Kawagoe et al., 2009) or river undercutting (Leshchinsky et al., 2019), whose effects can act independently or in combination with anthropic effects such as land use changes (Persichillo et al., 2017) or road cuts (Pradhan et al., 2022). In this overall picture, an added level of complexity is brought by the fact that landslides occurring within the same area and in response to the same trigger can exhibit different failure mechanisms, and each one of these can even evolve or change as the moving mass travels downhill (Hungr et al., 2014). Because of this multi-faceted nature, predicting where landslides may manifest across a given landscape represents a topic that has been investigated over thousands of scientific experiments since 1970^{ies} (Caleca et al., 2025).

The notion of landslide spatial prediction (Brenning, 2005) is commonly referred to as susceptibility (Reichenbach et al., 2018). This numerically translates into the probability of a

given unit partitioning the landscape to host at least one landslide, conditional on landscape characteristics (Zêzere et al., 2017). This relatively simple definition further includes more than one aspect of fundamental importance. A unit, or more specifically a mapping unit, is the smallest spatial element upon which an occurrence probability value is ultimately assigned to. In this sense, the relevant literature has seen two divergent approaches been tested over the years. The first is to partition the landscape into regular mapping units or objects whose geometry is consistent across any study area. These usually correspond to squared grid cells (Frattini et al., 2010), or much more rarely, to hexagons (Sabokbar et al., 2014). Conversely, the second approach subdivides any landscape of interest into irregular mapping units, with various objects experimented on in the literature depending on specific goals. For instance, Slope Units, or half-sub-basins have been used to fractally delineate areas bound between ridges and streamlines, with the aim of capturing the geomorphological response of a slope to a landslide (Carrara et al., 2003). In a similar manner, watersheds have also been used, for instance when the landslide class under consideration falls under the flow type (Hungr et al., 2014), thus being influenced by broader effects due to catchment hydrology (Wang et al., 2022). Aside from these irregular units tailored to physiographic characteristics, administrative units have been employed when the prediction is sought to match municipal boundaries due to financial reasons (Fang et al., 2025).

Irrespective of the mapping unit of choice, another important element in the standard susceptibility notion is the dependence of the probability on landscape characteristics whose spatial patterns are assumed to be constant over time (Günther et al., 2013). This has contributed to interpret the susceptibility as an intrinsic property of any given landscape (van Westen et al., 2008), thus mostly considering it independently of preparatory factors such as land use changes (Pacheco Quevedo et al., 2023) or wildfires (Di Napoli et al., 2024), and triggering ones such as storms (Minder et al., 2009) or earthquakes (Shao and Xu, 2022). Yet, even taking aside temporal aspects, the spatial distribution of landslides is still closely tied to the nature of the triggering factor (Tanyas and Lombardo, 2019), and with the type of trigger being further correlated to the resulting type of landslides. For instance, falls and topples are frequently triggered by seismic events (Koukouvelas et al., 2015), while deep-seated, slow-moving landslides in fine-grained materials are typically controlled by prolonged rainfall (Bayer et al., 2018) or river erosion at the slope toe (Ahmad et al., 2022). In contrast, debris slides (Tao and Barros, 2014) and flows (Segoni et al., 2016) are usually triggered by intense, short-duration rainfall (Abbate et al., 2021). This is why an increasing number of contributions have challenged the assumption that landslide susceptibility should be estimated independently of the trigger (Ozturk et al., 2021), mostly because statistically this would affect the estimation of the predictors' effects in a model that, being blind to the trigger contribution, would over-rely on terrain characteristics alone (Opitz et al., 2022).

In such a complex literature overview, a further diversification has emerged when modelling earthquake or rainfall induced landslides. In the first case, the shaking pattern is usually integrated in the model tasked with estimating the coseismic susceptibility (Gupta and Satyam, 2024), from which near-real-time estimates are generated for specific seismic scenarios

(Lombardo and Tanyas, 2022). As for the second case, rainfall thresholds are derived by matching rainfall against records of landslide occurrences (Segoni et al., 2018). The main difference here is that earthquakes are impulsive phenomena, whereas precipitation can be distributed over longer periods. In this sense, the standard approach is to pass the rainfall signal to the given model by cumulating the antecedent rain for fixed time windows (Monsieur et al., 2019). This implies that predictions of rainfall induced landslides are obtained by compressing continuous rainfall time series into single scalar values for specific aggregation periods (Berti et al., 2012).

The inclusion of entire time series as an integral part of geospatial problems is an aspect that has built up a large literature in geoscience over the years (Schuster, 2024). However, in the case of landslide applications, few contributions have explored the use of complex techniques that would bypass the need for any temporal aggregation of the rainfall signal and its consequent simplification into scalar values. Fang et al. (2023) published a first example where speech-recognition tools are used to regress landslide occurrences against continuous precipitation signals, this being still applied in a somewhat traditional sense, with the rainfall constituting the only predictor in their model. Lim et al. (2024) further added to this experimental design by including terrain predictors to generate a more comprehensive susceptibility output. Interestingly, the nature of both architectures imposes to the continuous rainfall a temporal decay in their contribution to the model (Che et al., 2018). In other words, the models tend to "forget" past rainfall contributions and rather weigh the ones closer to the landslide occurrence date. This could be a reasonable assumption, as most landslide studies distinguish preparatory and triggering precipitation effects. However, in case where two rainfall events shortly affect the same area, the same reasoning may not hold.

In May 2023, in the Emilia-Romagna region of Italy, two extreme rainfall events occurred few days apart from each other (Berti et al., 2025). With this idea in mind, we thought that if passing a continuous time series to a proper landslide prediction model would have been the target of an experiment, then architectures that inherently "forget" past effects would likely misrepresent the interaction between the two storms and introduce a potential bias in the final landslide probability estimates. Therefore, we initially sought to create a model capable of handling both static predictors linked to terrain and geological characteristics of the affected area and add to those the rainfall as time-series, encompassing the two extreme events in the process. To further add complexity to an already intricated modelling design, the combined effect of the two storms mentioned above, gave rise to nearly 80,000 landslides, including various failure mechanisms (Berti et al., 2025). Therefore, we also replicated the initial design for each mapped landslide type and implemented an explainable AI (Angelov et ala, 2021) routine to study the continuous rainfall contribution in each model configuration.

The structure of the manuscript presents an overview of the landslide inventory and the weather system that led to it in Section 2. The method and experimental design are detailed in Section 3. As for Section 4 presents the results, whereas Section 5 allows for an open discussion on the advantages and limitations of our modelling protocol. Finally, Section 6 concludes the manuscript with remarks for future directions.

4.4 Study area and triggering events

The Emilia-Romagna region, located in northern Italy, is bordered by the Apennine Mountains to the southwest, the Po River to the north, and the Adriatic Sea to the east. Geomorphologically, the region is divided into two main domains: the hilly and mountainous Apennine ridge and the flat Po Valley. For the purposes of this study on landslide probability, we focused on the mountainous area (11670 km^2), as outlined in Figure 4.1, which was identified using a 5 m-resolution Digital Elevation Model (DEM) freely available from the regional web-geoportal of the Emilia-Romagna (https://geoportale.regione.emilia-romagna.it/download/download-data?type=raster).

The geological formations of the selected area were previously classified by Berti et al. (2025) into eight units, based on lithological characteristics and structural domains. In Figure 4.1, the distribution of these lithotypes is shown, revealing a significant difference between the western and eastern portions of the area, which correspond to the Ligurian and Tuscan-Umbrian geological domains, respectively. The flysch rocks and clays (unit maps5) in the Ligurian domain are generally more fractured, less resistant, and more prone to deep-seated landslides compared to those in the Tuscan-Umbrian domain (unit 7), due to the extensive tectonic stress they have experienced in the accretionary wedge (Conti et al. 2020).

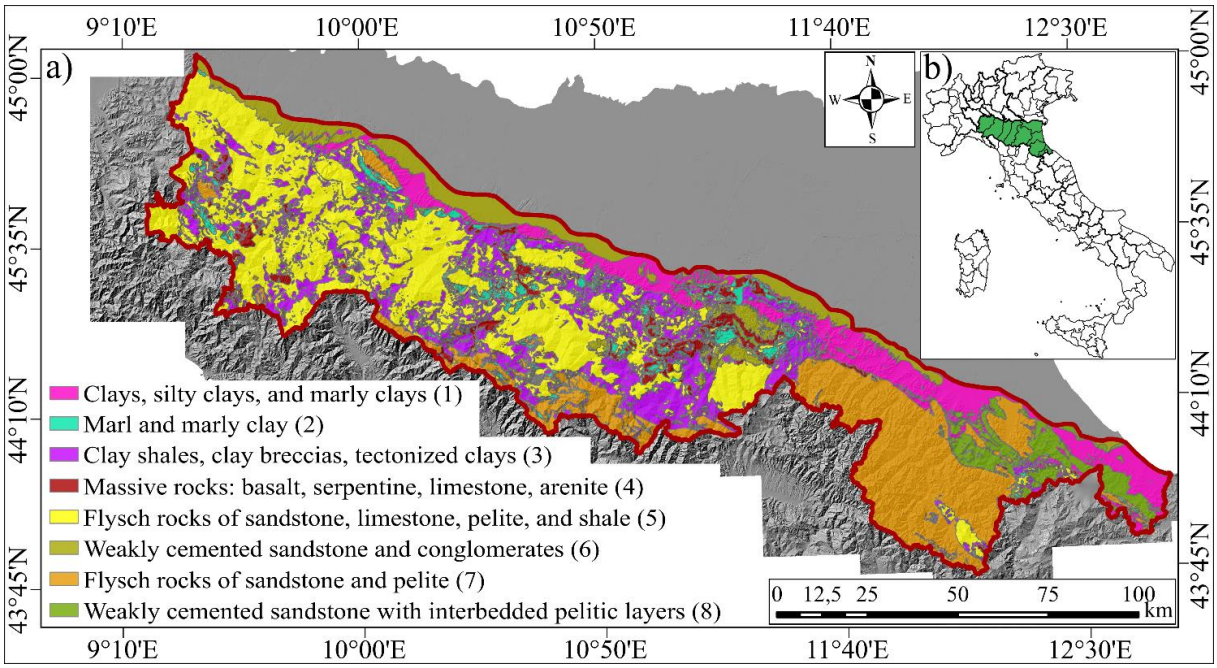


Figure 4.1. a) Geology of the study area (mountains zone of the Emilia-Romagna region), hillshade effect of the entire region in the background. b) Location of the Emilia-Romagna region highlighted in green inside Italy.

In May 2023, Emilia-Romagna experienced two severe meteorological events that impacted approximately 7,000-8,000 km² in the southeastern part of the region. According to Foraci et al. (2023), the first event occurred from May 1-3, delivering nearly 200 mm of rainfall within 48 hours, with a peak of 274 mm. A second event followed on May 16-17, bringing a similar intensity of rainfall, with totals ranging from 200 to 250 mm over 48 hours. Between these two

major events, additional rainfall episodes occurred, notably on May 9-10 and May 12-14. The recurrence interval for each event was estimated at 100-300 years, yet the combined effect of these closely spaced storms exceeded a 500-year return period (Brath et al., 2023).

The resulting rainfalls of combined events triggered widespread flooding across the Po Plain and thousands (>80000) of landslides in the Apennine mountains. Total damage estimates surpass €9 billion, with extensive impacts on roads, railways, buildings, and cultural heritage sites. Additionally, bridges, power infrastructures, and communication lines were severely affected. Agricultural activities and farming operations over approximately 1,000 km² were also disrupted. Tragically, 15 fatalities were caused by flooding, and two more were attributed to landslides.

4.5 Methodology

4.5.1 Slope Unit partitioning

We chose to implement a slope unit (SU) partitioning method instead of a grid-cell approach, capturing broader geomorphological conditions. Slope units are polygons defined by ridges and streamlines, reflecting homogeneous hydrological influences within each unit (Carrara et al., 1995; Ba et al., 2018). Therefore, we assume that if a landslide occurs and starts (triggering area) inside a SU, then it propagates within it and not beyond. Alvioli et al. (2016) introduced a fast SU generator, "*r.slopeunits*", which has demonstrated its reliability in subsequent studies (Alvioli et al., 2020; Lombardo and Tanyas, 2021). This tool generates SUs based solely on slope aspect homogeneity, using only a DEM as input, along with a few parameters to refine the output: *i) area_min*, the minimum SU area for the algorithm to converge to; *ii) circular variance*, a parameter controlling the flexibility of the aspect criterion, ranging between 0 and 1, whereby the former imposes a rigid criterion leading to a very dissected landscape, and the latter which allows for a greater variability, thus leading to less SU overall; *iii) cleansize*, a parameter through which the algorithm controls the merging process between spurious and neighbouring SUs; *iv) thresh*, the reference spatial extent from which the SU generation starts. To capture the variability of the terrain and landslide distribution, we aimed for a high-resolution partitioning into small SUs. After iterative tuning and visual checks, we settled on the following parameter values: *area_min* = 2500, *circular variance* = 0.25, *cleansize* = 2500, and *thresh* = 200000. This configuration generated 69159 SUs, with a mean area of 0.17 km^2 and a standard deviation of 0.14 km^2, balancing resolution, morphological accuracy, and computational efficiency. Visualisation of the SUs and their area distribution in Figure 4.2.

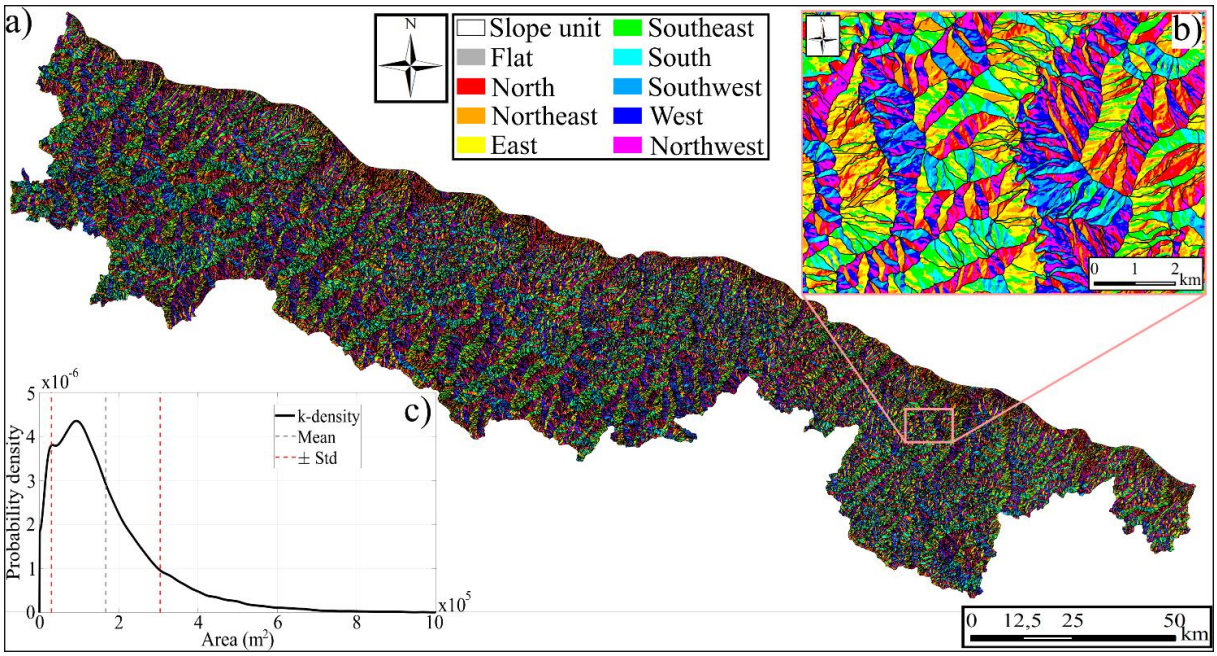


Figure 4.2. a) Slope units partitioning with the underlying slope aspect. b) Zoomed area of the slope units with aspect. c) Probability kernel density distribution of the SU areas in squared meters, plus the two-sigma deviation band.

4.5.2 Landslide inventory

The landslide dataset was created by Berti et al. (2025) through manual mapping using high-resolution aerial imagery (RGB and near-infrared) with a 0.2 m resolution. Post-event imagery from May 23 and pre-event imagery from 2020 were utilized. The mapping process was supported by field investigations, which aided in the accurate identification of landslide features. The dataset comprises 80,997 landslides, without differentiation between the two main rainfall events, classified into the following types: debris slides (DS), debris flows (DF), earth slides (ES), earth flows (EF), and rock slides (RS). The classification follows the system of Cruden and Varnes (1996).

Debris slides (DS), small to very small in area, typically occurred on steep slopes (>25°). They represent 66% of the total landslides by number and 49% by affected area. Debris flows (DF), frequently initiated by debris slides, usually travelled without following predefined channels and cleared vegetation in straight, elongated, rectangular shapes (small to medium areas). DF account for 15% of the total landslides by number and 29% by area. Earth slides (ES) and earth flows (EF) predominantly occurred in the clay-rich substrates (lithotypes 1 and 2 in Fig. 4.1), where fine-grained material controls landslide mechanisms and morphology. ES represent 12% of the total landslides by number and 7% by area, while EF account for 5% of the total landslides but cover 10% of the affected area. Both types are concentrated in badlands areas, with gentler slopes largely unaffected. Rock slides (RS) occurred on homoclinal slopes within the Marnoso-Arenacea Formation (unit 7) and developed as planar slides along bedding planes aligned with the slope. The displaced rock mass ranged from 2 to over 30 meters in thickness, and several slides extended over areas larger than 10 hectares. RS affected gentler slopes (typically less than 15°) and, although constituting less than 2% of landslides by number and 5% by area, they

posed a significant hazard due to their large volume, high velocity, and their occurrence on urbanized and cultivated lands.

Landslide processes can be considered stochastic for modelling purposes (Campforts et al., 2022). Therefore, we shifted from a density-based representation to a Boolean classification (0, 1), converting the landslide inventory polygons into binary presence/absence labels at the Slope Unit (SU) level. SUs were labelled as 1 (presence) if they contained the centroid of the triggering area of at least one landslide of a given type, in line with landslide susceptibility definitions (Reichenbach et al., 2018; Lima et al., 2022). The triggering area (shown in Figure 4.3) was defined as the continuous portion of the landslide polygon located above the 75th percentile of its internal elevation distribution. This threshold was selected based on careful visual inspection of the release zones across different landslide types. As shown, the triggering polygons consistently encompass the source area (i.e., the broader scarp) of the landslide body. The method also allows for distinguishing multiple source areas in coalescent earth flows (see Figure 4.3b). The resulting centroids of these identified triggering areas were then used to determine whether they fall within SU polygons. This simplification enables clean attribution of landslide presence while respecting the spatial resolution differences between landslide boundaries and SUs, as well as the SU generation process. Figure 4.3 further illustrates the slope units (blue contours) to show how they capture landslide propagation, while accounting for the mismatch in spatial resolution. Notably, we intersected only the triggering zones with the SUs, rather than the full landslide bodies, to ensure a clean and accurate assignment of landslide presence. This avoids potential misattribution caused by marginal overlaps between landslide bodies and SU boundaries, especially near low-resolution SU borders or shared boundaries along drainage divides.

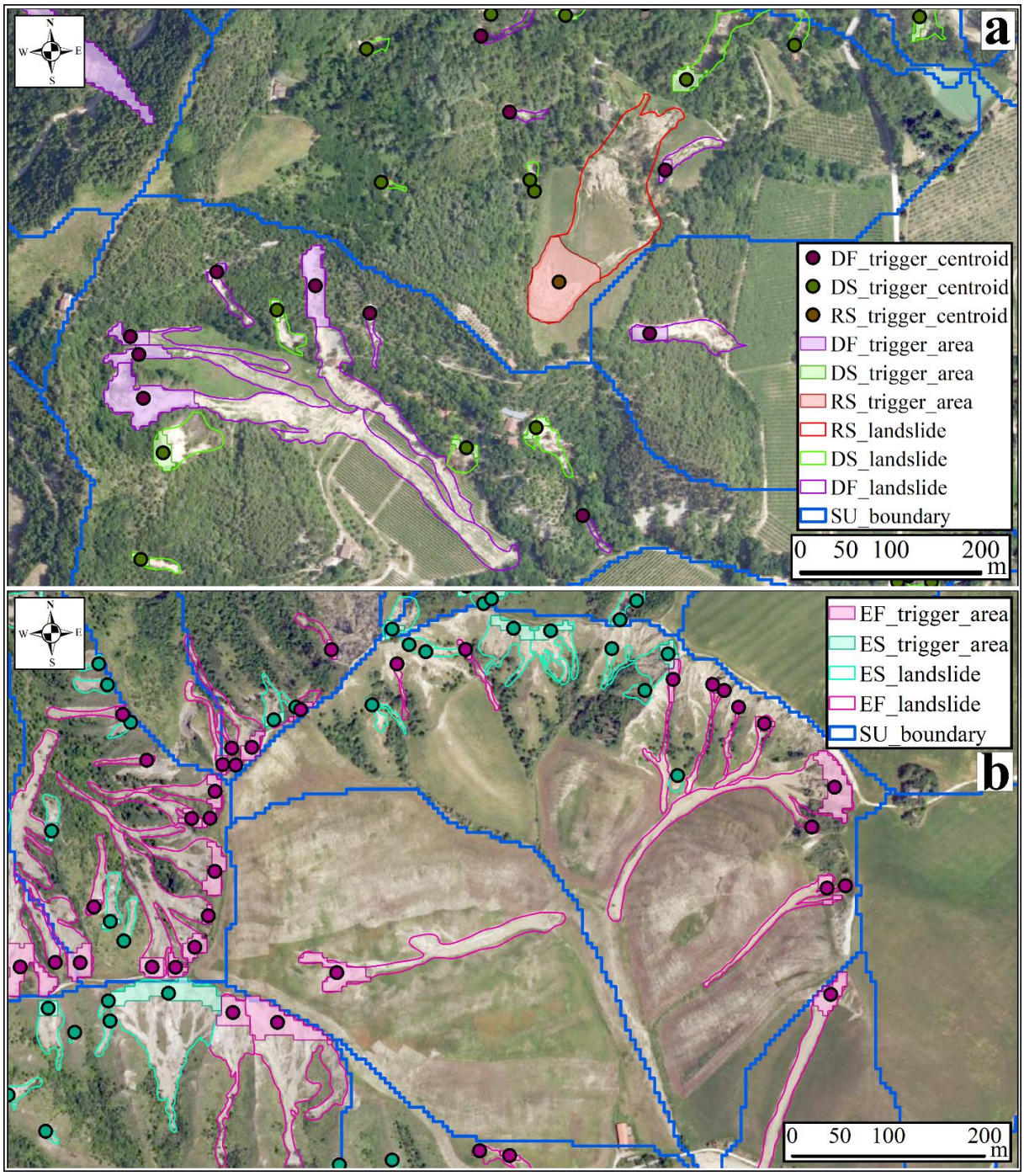


Figure 4.3. Areas details of landslide triggering polygons (5 m resolution) and centroids identification for all landslide type: Debris slide, debris flow, rock slide above (a); and earth slide and earth flow below (b). Visualized also the slope unit boundaries and RGB images at 1m resolution.

This methodology results in 8502 SUs with DS presence (11.65%), 4670 SUs with DF presence (6.75%), 2002 SUs with ES presence (2.89 %), 1247 SUs with EF presence (1.80%), and 221 SUs with RS presence (0.32%). This method was necessary because the toes of some landslide deposits crossed multiple SU boundaries, potentially distorting the true location of the landslide. As a result, our analysis reflects the probability of landslide triggering within each SU. For a general overview, we calculated the normalized landslide density for each SU (Figure

4.4). The distribution of landslide types across the SUs corresponds to the geological setting (Fig. 4.1), as previously discussed.

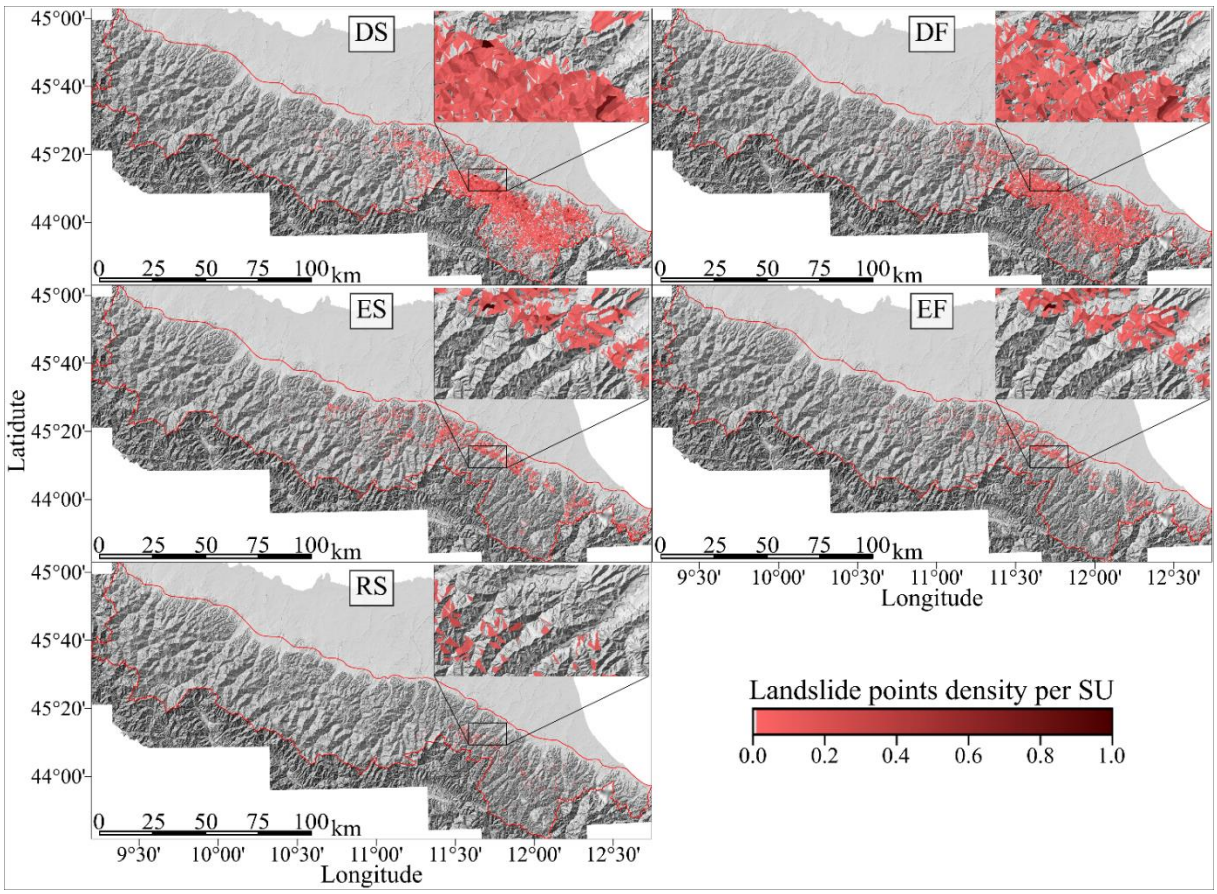


Figure 4.4. Normalized landslide density distribution per Slope Unit (SU), per each separate type of landslide (DS, DF, ES, EF, and RS). Hill shade effect in the background ad zoomed panels in the top right position per each landslide type panels.

As discussed earlier, different types of landslides originated under distinct geomorphological conditions. It is therefore essential to maintain this differentiation in the analysis. We treated the five landslide types as separate training datasets, producing specific probability maps for each type. This approach not only improves the accuracy of predictions but also offers practical benefits for early warning systems and territorial planning. Different types of landslides pose varying levels of hazard in terms of speed and volume, each requiring tailored mitigation strategies (Dai et al., 2002). For instance, EF, while typically slower, are less hazardous in the short term and can be managed over the long term. In contrast, DS are extremely rapid but involve smaller volumes, making them easier to repair. RS, on the other hand, present a severe hazard due to their speed and large volumes, making them difficult to manage.

4.5.3 Predictors

The predictors used in our model are divided into two categories: static (Table 1) and time-dependent variables. The static ones can be further classified into three main categories: geology-related, land cover, and topography. We included geology, land cover, slope, and

aspect as the basic and most significant factors (Budimir et al., 2015) because they represent the primary controls on landslide occurrence. For the geology (Fig. 4.1) and land cover, we calculated the percentage of each lithological unit and land cover type within each SU. Thus, each lithological/land cover unit becomes a new predictor with values ranging from 0 to 100 per SU, which is treated by the model as a continuous variable, preserving the information of different geological formations and land cover types within the SU. This approach is similar to what was implemented by Dahal & Lombardo (2025) using soil type information, which is often neglected in the literature by selecting only the dominant lithological/land cover unit to represent the entire SU, as expressed by Amato et al. (2019) and He et al. (2025). We incorporated plan (horizontal) and profile (vertical) curvature to distinguish concave and convex portions of slopes, and we added the bedding plane–slope orientation relationship, which is mechanistically linked to translational slides on cataclinal slopes. The Topographic Wetness Index (TWI) was used as an integrated hydrological predictor because it combines slope and flow accumulation, making the addition of distance to streams or other flow metrics largely redundant. Bedrock dip and dip direction information were derived by interpolating point measurements available from the geological map (provided by the regional geoportal). The inclusion of bedding information was inspired by field observations reported by Berti et al. (2025), who highlighted that strata attitude largely controls the failure mechanism, especially in the case of rock slides.

We intentionally excluded several predictors to maintain a more straightforward modelling interpretation in relation to the dynamic predictors (preparatory and triggering rainfall). Topographic Position Index (TPI) and related roughness measures were not used because their signal overlaps with what is already captured by plan/profile curvature and slope; distance to faults (and analogously, seismic PGA proxies) was omitted as it is not related to the triggering mechanism; structural predisposition is already represented by bedding–slope orientation and geology, while fracture information, although valuable for rock slide type, requires detailed, high-resolution, site-specific investigations that are not feasible at the regional scale. NDVI largely reflects land cover (already included), even though its temporary, seasonal trends could be included as dynamic predictors along with rainfall, though at this stage of research it was not incorporated. Soil maps and soil thickness were not used due to their heterogeneous scale and quality, making them unsuitable for regional-scale analysis. Road network density would introduce anthropogenic bias that is not representative of natural predisposition. From a future risk analysis perspective, including elements at risk directly in the landslide probability model would introduce redundant information, as the final hazard outputs would later be intersected with road networks and vulnerable assets to compute risk.

Covariate	Categories/Values	Source/Description
Geology	See Figure 4.1	Litho-technical units classification - Berti et al. (2025)

Land cover	*i)* *ii)* *iii)* *iv)* *v)* *vi)* *vii)* *viii)* *ix)* *x)* *xi)* *xii)*	Urban and Transport Areas Extractive Areas Artificial Green Areas, Arboriculture Arable Land Mixed Agricultural Lands Permanent Grasslands Forests Shrub and Herbaceous Lands Rare or Absent Vegetation Wetlands Water Environments	Regional geoportal, referencing data from the year 2020 at a 1:10,000 scale. https://geoportale.regione.emilia-romagna.it/download/dati-e-prodotti-cartografici-preconfezionati/pianificazione-e-catasto/uso-del-suolo/2020-coperture-vettoriali-uso-del-suolo-di-dettaglio-edizione-2023
Mean Slope		0° – 90°	Average slope value inside the SU, derived from DTM 5x5 m
Std Slope		0° – 90°	Standard deviation of slope values inside the SU, derived from DTM 5x5 m
Mean Aspect		0° – 360°	Average aspect value inside the SU, derived from DTM 5x5 m
Std Aspect		0° – 360°	Standard deviation of aspect values inside the SU, derived from DTM 5x5 m
Planar curvature		-682 – 496	Average planar curvature value inside the SU, derived from DTM 5x5 m
Profile Curvature		-469 – 686	Average profile curvature value inside the SU, derived from DTM 5x5 m
Bedrock Dip Direction		0° – 360°	Average bedrock dip direction value inside the SU, derived from point measurements interpolation
Bedrock Dip		0° – 90°	Average bedrock dip value inside the SU, derived from derived from point measurements interpolation

Table 4.1. Summary of the static predictors

The time-dependent variable in our analysis consists of the rainfall signal. The relevant data was processed and incorporated into our models following two criteria. In the first, rainfall is treated as a static map, expressing the cumulative precipitation over multiple time windows for which we tested our model. As for the second criterion, the rainfall was kept in its original form, and treated as time-series to be fed to our model (Dahal et al., 2024). These different representations of rainfall data form the basis for the various configurations of our analysis.

We sourced rainfall data from the Emilia-Romagna region's meteorological network, which consists of 300 stations distributed almost uniformly across the region (https://simc.arpae.it/dext3r/). Out of these, 188 stations are located within our study's mountainous area, providing an approximate station density of one per 62 km². Each station records rainfall amounts at three intervals: daily, hourly, and up to a 15-minute frequency. To capture the rainfall conditions leading up to the landslides, we focused on the month (31 days, or 744 hours) preceding the second landslide-triggering event. This timeframe includes both

major rainfall events as well as relevant wet and dry periods. Notably, we opted to test the rainfall signal brought by the daily and hourly records, leaving aside the 15-minute information due to the large computational resources it would require as well as the likely overdetermined problems it would introduce.

We interpolated the daily and hourly rainfall data across the study area, using the centroids of the SUs as the interpolation grid. Figure 4.5 shows the cumulative of the interpolated rainfall day by day. For interpolation, we employed the Radial Basis Function (RBF) with a linear kernel (Franke, 1982). This method was chosen because it allows for the use of an unstructured, irregular grid (based on the centroids of the SUs) and the linear function ensures that the interpolated surface passes exactly through the measured values from the rain gauge stations, thereby preserving the integrity of the original data.

To assess accuracy and validate our choice, we performed a leave-one-out cross-validation (LOOCV) using the available rain-gauge network. For each day, one station within the study area was iteratively excluded, and the remaining stations were used to interpolate the rainfall at the excluded station's location. Prediction errors were recorded via the Mean Absolute Error (MAE), the Root Mean Square Error (RMSE), and the mean error (Bias). We compared three interpolation methods using their standard, out-of-the-box configurations: Inverse Distance Weighting (IDW) with power = 2, no max-distance, and epsilon = 1.5 km snap; Ordinary Kriging with a spherical variogram (automatic fitting, 10 lags, exact snapping); and RBF with a linear kernel and zero smoothing (exact interpolant). The LOOCV error metrics are reported in Table 2. Based on its lowest MAE and near-zero bias, along with the desirable property of exactly honoring measured values, we selected the RBF (linear) interpolation for the analyses presented in this study.

The resulting interpolated rainfall values across all SU centroids for a representative day (May 17, during the second rainfall event peak) were plotted, along with rain gauge locations throughout the Emilia-Romagna region. These comparisons across the three methods are included in Supplementary Material S 4.1. The visualizations also confirm the relatively homogeneous spatial distribution, thus well representative, of rain gauges across the region.

Method:	**MAE (mm)**	**RMSE (mm)**	**Bias (mm)**
IDW	2.253	6.472	−0.267
Kriging	1.953	**5.761**	−0.143
RBF (linear)	**1.881**	5.856	**−0.063**

Table 4.2. LOOCV metrics (MAE, RMSE, Bias) comparing IDW, Kriging, and RBF interpolation methods for daily rainfall.

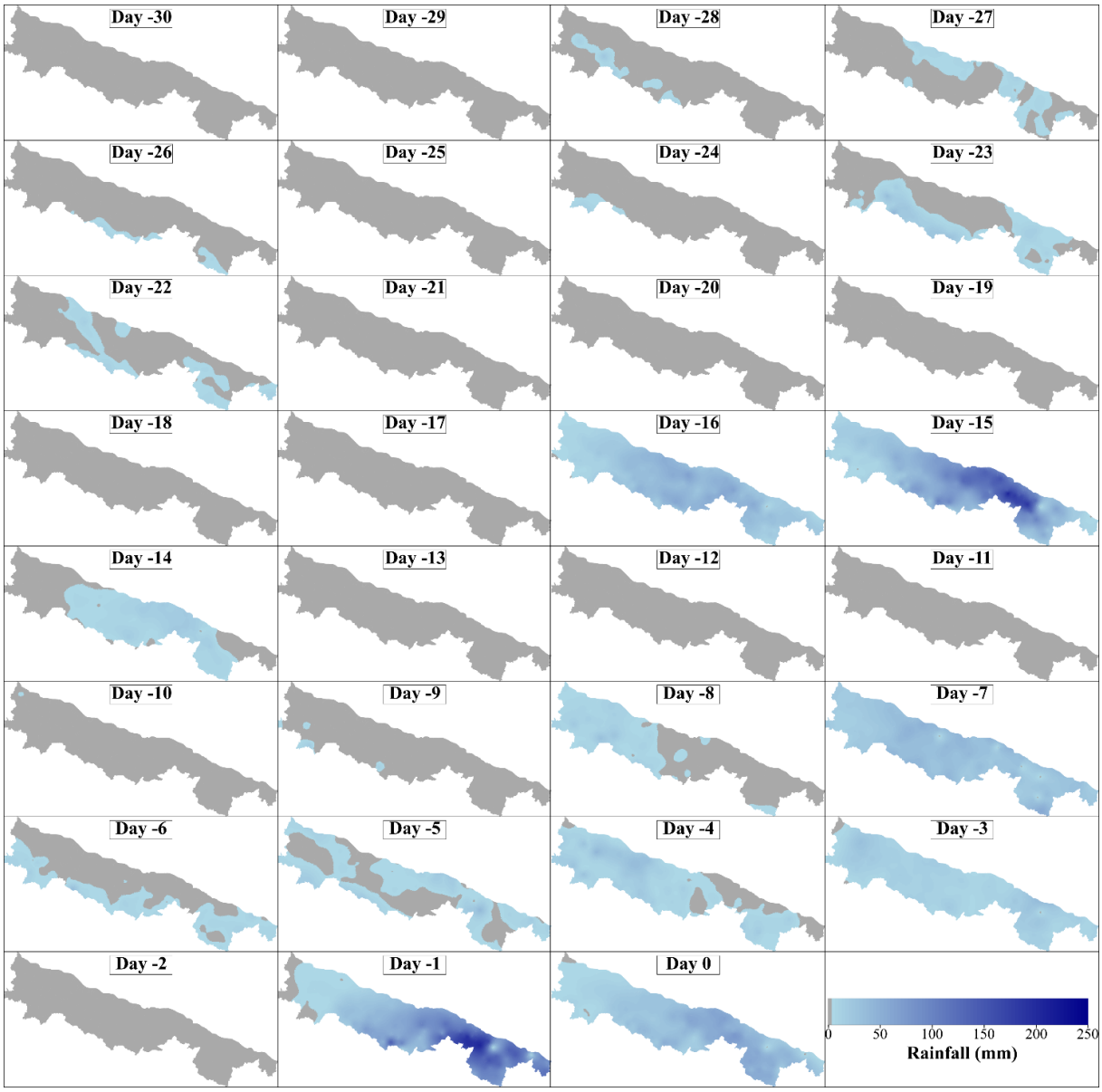


Figure 4.5. Rainfall per day from the daily interpolation in the study area, from 31 days before the end of the second event (day -30 = 16/04/2023) to the ending of the second event (day 0 = 17/05/2023).

4.5.4 Transformer Neural Network: architecture and experimental design

Building upon the foundation laid by Dahal et al. (2024), we re-engineered the model architecture to shift its focus from seismic waveforms to rainfall time series. The architecture is a modified version of the transformer neural network (TNN) by Vaswani et al. (2017), adapted to integrate static covariates, such as slope and geology.

The model estimates the pseudo-probability of a landslide occurring at a location, $p(s)$, based on static $(X_s(s))$ and dynamic $(X_D(s,t))$ predictors. This pseudo-probability follows a Bernoulli distribution, representing landslide occurrence as a binary event [0-1], where 0 represents the absence of landslide and 1 represents the presence of landslide. The architecture is briefly described below:

- **Dynamic Covariate Encoding**: Multiple transformer layers process the dynamic input (rainfall time series, at daily and hourly intervals). Each layer includes normalisation, multi-head attention, dropout, and feed-forward operations. A global average pooling layer is used to compress the temporal information and emphasize the spatial aspects.

$$f1_{(s)} = \tau(X_D(s,t))$$

- **Static Covariate Processing:** The static predictors are passed through a series of dense layers with batch normalisation, dropout, and activation functions, refining the static feature space.

$$f2_{(s)} = \delta(X_s(s))$$

- **Feature Concatenation and Decoding**: The outputs from the dynamic and static layers are concatenated into a unified feature tensor.

$$f3_{(s)} = f1_{(s)} \oplus f2_{(s)}$$

  This combined feature tensor is then processed through a decoding neural network, $\eta_{(s)}$, with a sigmoid activation function at the final layer to generate the $p(s)$.

The model's components $\tau_{(s)}$ and $\delta_{(s)}$ can be selectively turned on or off for benchmarking, allowing the evaluation of different input configurations. So, we tested several configurations for incorporating rainfall data into our models, varying the amount and type of rainfall information used. The configurations were designed to capture different aspects of the rainfall-triggering conditions, allowing us to evaluate the influence of both short-term (weekly) and long-term (monthly) rainfall on landslide occurrence. These configurations are detailed below:

1. **Terrain-only**: No rainfall data included; the model uses only static terrain predictors (geology, topography, land cover).
2. **1-week cumulative rainfall**: The cumulative rainfall per SU over the week preceding the second event is used as input.
3. **2-week cumulative rainfall**: The cumulative rainfall per SU over the two weeks preceding the second event is used as input.
4. **3-week cumulative rainfall**: The cumulative rainfall per SU over the three weeks preceding the second event is used as input.
5. **1-month cumulative rainfall**: The cumulative rainfall per SU over the entire month preceding the second event (encompassing both events) is used as input.
6. **All scalar cumulative values**: This configuration includes all cumulative rainfall values (1 week, 2 weeks, 3 weeks, and 1 month) as input variables simultaneously.
7. **Time series (daily)**: The entire daily rainfall time series (waveform day-by-day, not cumulative) per SU for the month preceding the event is used as input.
8. **Time series (hourly)**: The entire hourly rainfall time series (waveform hour-by-hour, not cumulative) per SU for the month preceding the event is used as input.

Our method specifically investigates the potential of the proposed TNN model, which incorporates time series of rainfall data (waveforms) instead of the scalar cumulative values more commonly used in the literature (Ahmed et al., 2023; Steger et al., 2023; Huang et al.,

2022; Chikalamo et al., 2020; Wu and Chen, 2009). While some authors focus on identifying the optimal cumulative rainfall time window preceding an event to explore the influence of rainfall on landslide occurrence, our model goes beyond this approach. Instead of relying on a single optimal time window at the end of the analysis, our method incorporates the complete rainfall time series, capturing the influence of rainfall over the entire period. This allows us to account for both dry and rainy intervals and their cumulative effect on landslide probability, offering a more comprehensive evaluation of the rainfall's impact.

4.5.4.1 Architecture parameters tuning

The model architecture was initially settled by literature basis (Vaswani et al. 2017; Dahal et al. 2024, Zhou et al. 2021), then finished, aligning with our specific aims and preliminary results. The dataset is divided into 70% for training and 30% for testing, out of which 30% of the training data used for validation to monitor performance and prevent overfitting during the training itself. The sampling strategy for the training dataset is set to random with a fixed random seed (random state = 420) was used across all types, to ensure reproducibility and consistency. This allowed us to guarantee comparability between landslide-type models while avoiding variability due to random selection.

The model is designed to handle an imbalanced dataset, such as in our case (see Tab. 3), by assigning greater weight to the loss contributions of slope units (SUs) with label 1 (i.e., landslide occurrence). This weighting is implemented via the class_weight parameter in TensorFlow/Keras, which adjusts the importance of each class during training without altering the dataset itself. The weights were not derived directly from the raw class ratios but were instead optimised through a trial-and-error tuning process. This process involved iteratively training the model using different class weights and evaluating its performance on a fixed test set (30%) across several metrics (AUC, F1-score, accuracy, MCC, etc.). The tuning was conducted independently for each of the five landslide types due to the significant variation in their frequency distributions. This approach ensured that the model appropriately learned from the minority class without artificially altering the dataset composition through over- or under sampling.

The model was so configured with the following parameter settings: attention head size = 4; dropout rate (all layers) = 0.3; activation function = ReLU; initial learning rate =$9*10^{-4}$; learning rate decay steps = 40,000; learning rate decay factor = 0.95; landslide class weights (see Tab. 3); number of attention heads = 4; feed-forward units = 8; static layer depth = 8; static layer width = 32; number of epochs = 200; and batch size = 32. The model is trained using the Adam optimizer (Kingma and Ba, 2014) with a binary cross-entropy loss function.

The internal model architecture (attention heads number, dropout, static layers depth, etc…) was arbitrary fixed for all landslide types, after an internal tuning to find the best compromise between predictive capacity and efficiency. This choice was guided by three main reasons: (i) ensuring practical usability and scalability for future operational applications such as early warning systems, where keeping a single architecture simplifies retraining and use by operators; (ii) reducing the risk of overfitting, especially for underrepresented classes such as RS (0.32%

of samples); and (iii) achieving a good balance between performance and computational cost, given the high cost of the TNN modelling (hours per landslide types).

Landslide Type	Positive SUs (1)	Negative SUs (0)	Original Ratio (0/1)	Training Ratio (0/1)	Testing Ratio (0/1)	Best class weight (0:1)
Debris Slide	8502	60657	7.1	6.3	6.4	1:2
Debris Flow	4670	64489	13.8	12.2	12.9	1:3
Earth Slide	2002	67157	33.5	30.1	31.5	1:5
Earth Flow	1247	67912	54.5	48.4	51.0	1:6
Rock Slide	221	68938	311.9	268.7	316.3	1:22

Table 4.3. Original, train and test splitted positive (1) and negative (0) slope units and their ratios, and the best class weight for all landslide types.

4.5.4.2 Rainfall filtering process

We finally applied a filtering process during the training phase to reduce the potential bias caused by minor rainfall events on landslide presence. Given the large number of "0" SUs in the dataset, we recognized the risk of the model learning a spurious relationship that even minimal rainfall could lead to landslides. To mitigate this, we calculated the mean rainfall for each day and identified the best percentile of these daily means as a threshold. Days with rainfall values exceeding this threshold were classified as relevant rainfall days**.** Then, we exclude the SUs where there is at least one zero rainfall day in the relevant rainfall day, see Fig. 4.6 for clarification. This ensured that the model focused on the relationship between rainfall intensity (low to high) and landslide presence, rather than simply identifying rainfall occurrence (yes or no) as a predictor. The tested daily means rainfall thresholds were the percentiles value of 50^{th}, 60^{th}, 70^{th}, 75^{th}, and compared with the complete dataset (all) without filtering process. The percentile thresholds do not go higher than 75^{th} percentile, because at already 80^{th} percentile the totality of the SUs is included, resulting in a no-filter situation (all).

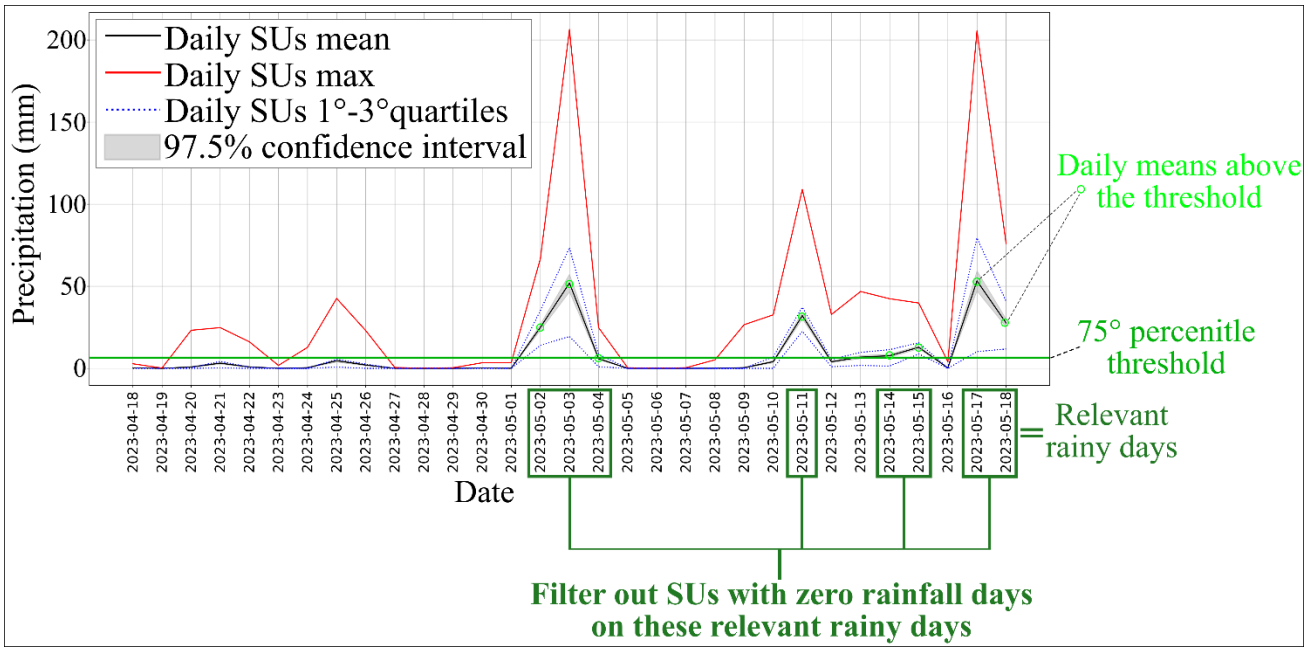


Figure 4.6. Filtering scheme for excluding SUs with zero rainfall during relevant rainy days. The base plot shows daily precipitation statistics across all slope units (SUs) for the 31-day analysis window. Relevant rainy days (green markers) are identified as those with daily mean rainfall above the 75th percentile threshold (green line).

The application of these percentile thresholds results in the training datasets in Table 4, and the relative performance metrics on prediction phase (include the whole dataset) are reported in the supplementary material (S 4.2 to S 4.6). The 70th and 75th percentile thresholds exhibit the best results, with the latter being slightly better for DS, ES, and RS. The filtering process demonstrate that the model learns a better rainfall-landslide relationships if we did not include all SUs, providing an efficient reduction of the bias caused by minor rainfall events on landslide presence. Finally, we adopted the 75th percentile threshold for all landslide types to ensure a better and faster reproducibility of the modelling solution.

Rainfall threshold	Total filtered SUs	Filtered training SUs	Filtered testing SUs
50th percentile	21178	14824	6354
60th percentile	46918	32842	14076
70th percentile	61751	43225	18526
75th percentile	62409	43686	18723
None	69159	48411	20748

Table 4.4. Resulting filtered dataset by rainfall thresholds.

Evaluation

We evaluated our model using a wide suite of performance metrics typical of classification problems (Taha and Hanbury, 2015). Each metric emphasizes different aspects of the confusion matrix, made of True Positives (TP), True Negatives (TN), False Positives (FP), and False Negatives (FN), thus offering different insights into the model functioning (Power, 2010).

Notably, calculating the confusion matrix requires the definition of a probability cutoff, which we set here to 0.5. Thus, TP indicates where pseudo-probability is higher than 0.5 and observed landslide is present, TN is where pseudo-probability is below 0.5 and no observed landslide, FP is where pseudo-probability above 0.5 but without observed landslide, FN is where pseudo-probability is below 0.5 but there is a observed landslide.

- Area Under the Curve (AUC) [0 - 1] represents the area under the Receiver Operating Characteristic (ROC) curve (Fawcett, 2006), which plots the true positive rate (TPR) $[{}^{TP}/_{(TP+FN)}]$ against the false positive rate (FPR) $[{}^{FP}/_{(FP+TN)}]$. A higher AUC value indicates better discrimination between classes. AUC is often considered one of the most useful metrics because it does not require a predefined binary threshold to build the confusion matrix. Instead, it evaluates the model over a range of thresholds, which are used to construct the ROC curve, making it a robust measure of performance across different decision boundaries (Hanley and McNeil, 1982).
- Accuracy (ACC) [0 - 1] measures the proportion of correct predictions (both TP and TN) out of all predictions. It is straightforward but can be misleading in imbalanced datasets, as a model may appear to perform well by predicting the majority class accurately (Menditto et al., 2006).

$$Acc = \frac{TP + TN}{TP + TN + FP + FN}$$

- Recall (REC) [0 - 1], or Sensitivity, focuses on how well the model identifies true positives (landslides), minimizing false negatives (Altman and Bland, 1994). It is critical when missing a landslide event (FN) is more costly than false alarms (FP).

$$Rec = \frac{TP}{TP + FN}$$

- Matthew's Correlation Coefficient (MCC) [-1 - 1] provides a balanced evaluation by considering all four elements of the confusion matrix (Matthews, 1975). It is especially useful for imbalanced datasets (Chicco and Jurman, 2020).

$$MCC = \frac{(TP \times TN) - (FP \times FN)}{\sqrt{(TP + FP)(TP + FP)(TN + FP)(TN + FN)}}$$

- Kohen's Kappa (*k*) [-1 - 1] measures the agreement between predicted and actual classifications, adjusting for chance agreement. It is particularly useful when dealing with imbalanced datasets, as it highlights performance beyond random chance (Fatourechi et al., 2008).

$$k = \frac{Acc - Expected\ Acc}{1 - Expected\ Acc}$$

- F1 Score [0 - 1] balances TP, FP, and FN. It is the harmonic mean of precision and recall (Van Rijsbergen, 1979). The F1 score emphasizes both false positives and false negatives equally, making it useful when both types of errors are of similar concern.

$$F1 = \frac{2TP}{2TP + FP + FN}$$

- F2 Score [0 - 1] gives more weight to Recall, prioritizing the minimization of FN, and is especially important when missing a landslide is more critical than falsely predicting one (false alarm).

$$F2 = \frac{5TP}{5TP + 4FN + FP}$$

To clarify the role of various metrics, ACC may usually appear high in imbalanced datasets but fails to capture the model's performance in detecting actual landslides, which Rec emphasizes. F1 and F2 are particularly important for understanding trade-offs between different kinds of errors: while F1 treats false positives and false negatives equally, F2 focuses more on reducing false negatives, making it more suitable for contexts where missing a landslide is more problematic than predicting false alarms. F-scores also do not take TN into account (Power, 2015), an aspect which is otherwise addressed by MCC and *k*, and for which they are known to be useful specifically in unbalanced data (Chicco and Jurman, 2020; Fatourechi et al., 2008).

4.5.5 Explainability

The contribution of predictors to landslide pseudo-probability is crucial for understanding both the model's functioning and the phenomenon itself (Gunning et al., 2019; Dahal & Lombardo, 2023), particularly when considering different landslide types. In this study, we focused on exploring the relationship between rainfall as a dynamic input and landslide pseudo-probability, aiming to gain insights into how the TNN model operates and makes decisions. We excluded other static predictors from this analysis, as their contributions to landsliding are generally well-documented and fall outside the primary objectives of this study. Several methods have been proposed in the literature to evaluate the contribution of predictors to model decisions (Dahal & Lombardo 2023, Das & Rad, 2020; Molnar, 2019). We opted for the Expected Gradients method, which integrates SHAP (SHapley Additive exPlanations) (Shapley, 1953; Lundberg and Lee, 2017) with concepts from Integrated Gradients (IG) (Sundararajan et al., 2017) and SmoothGrad (Smilkov et al., 2017). This approach was implemented using the Python package "GradientExplainer" from the SHAP repository (Lundberg, 2022: https://github.com/slundberg/shap). The method satisfies the axioms of SHAP:

- Local Accuracy: the total contribution of all features should exactly match the model's output for the given input.
- Missingness: features that are missing (or irrelevant) in the input should have zero contribution to the model's prediction.
- Consistency: if the contribution of a feature increases (or stays the same) in a new version of the model, its SHAP value should not decrease.

In addition, the method incorporates the IG concept of calculating gradients for each feature value along the path from a baseline to the actual input value and averaging these gradients at multiple points. This ensures the model satisfies the following axioms:

- Sensitivity: If the input and baseline differ in one feature, and the model's prediction changes because of this, the differing feature must have a non-zero attribution.
- Completeness: The sum of all feature attributions must equal the difference between the model's output at the input and the baseline.

The GradientExplainer method extends these principles by using multiple baseline samples from the input dataset rather than a single reference point, aligning with SHAP's principle of leveraging a background distribution to better incorporate feature interactions. Additionally, SmoothGrad introduces Gaussian noise to input values, generating perturbed inputs that approximate the input's neighborhood. This effectively smooths the gradients over nearby regions in the input space, stabilizing SHAP value estimates and reducing sensitivity to small input variations. For our analysis, we used the following parameters in the GradientExplainer package: batchsize = 150, local_smoothing (SmoothGrad effect**) =** 0.01, nsamples (number of baselines) = 200. This setting is a good balance between efficiency and results, we deliberate choose a small smoothing effect to not introduce to much noise.

4.6 Results

In this section, we present the model's performance across the different configurations of the dynamic rainfall predictor. We then showcase the resulting pseudo-probability maps for the best-performing case, followed by an analysis of the influence of rainfall over time and across spatial distributions.

4.6.1 Model evaluation

The performance of the model is evaluated on a random 30% SU subset of the test area. The corresponding details are shown in Figure 4.7 for each landslide type, as well as a case where the metrics have been computed irrespective of the failure mechanism. A general trend can be observed: the performance metrics decrease from DS to RS**.** These landslide types also decrease in landslide numbers, or more precisely in unstable labels assigned to the SU partition. This suggests that the model performs better when it has more data to be calibrated on, an aspect consistent with a model archetype such a TNN, built for large datasets.

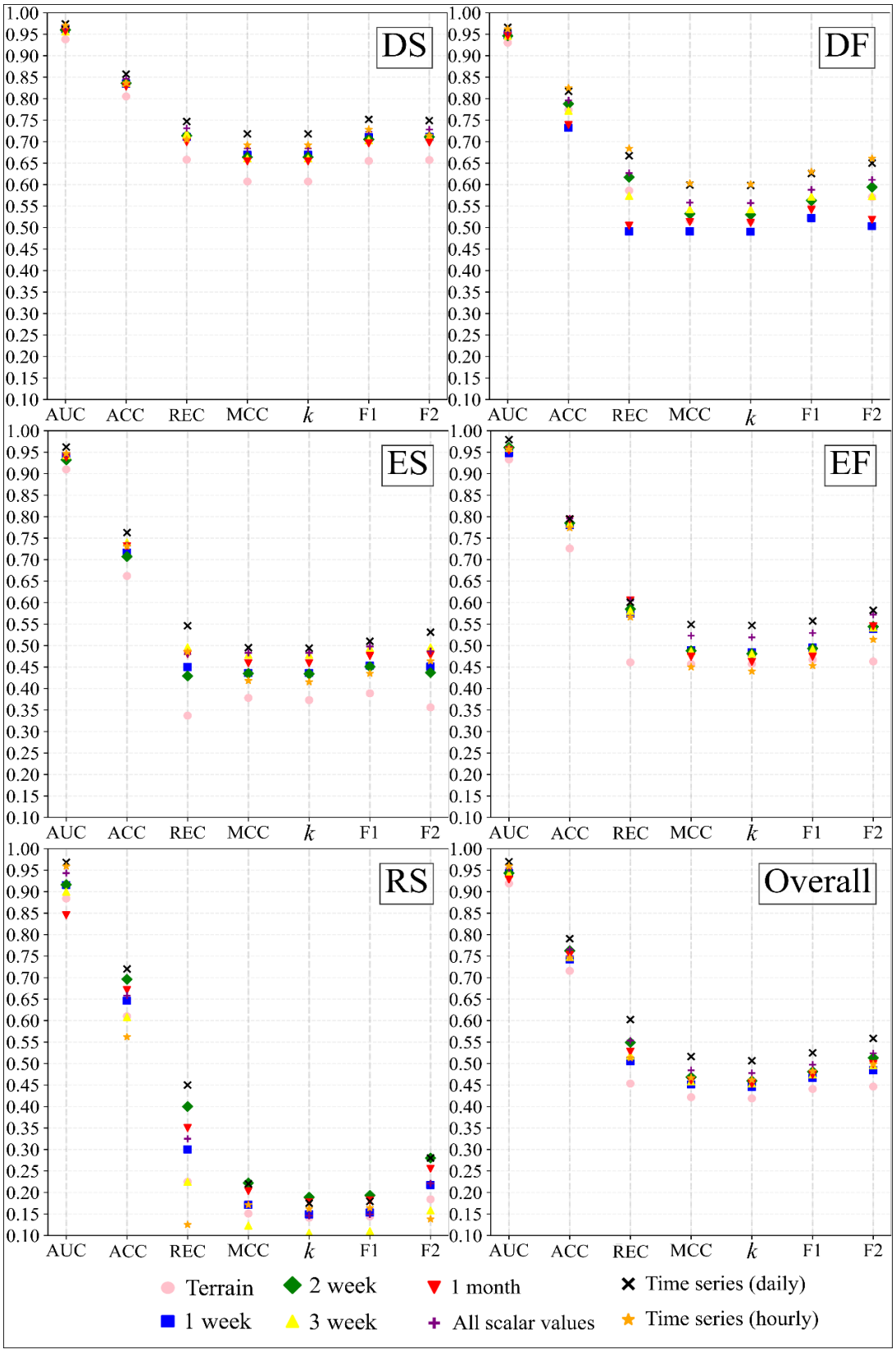


Figure 4.7. Performance evaluation for different landslide types (DS, DF, ES, EF, RS) and the overall. All exanimated metrics are Area Under the Curve (AUC), Accuracy (ACC), Recall (REC), Mathew Correlation Coefficient (MCC), Kohen's Kappa (k), F-score (F1 and F2).

For debris slides (DS), the model achieved the highest performance, with AUC values between 0.938 and 0.974, MCC ranging from 0.607 to 0.718, and F1 scores between 0.655 and 0.752. The configuration using only terrain predictors yielded the lowest performance across all metrics, while the time series (daily interval) configuration provided the best results. The hourly time series also performed quite well, ranking second among all configurations for MCC, F1, and *k*. Notably, F2 scores were similar to F1, indicating that the model does not disproportionately emphasise reducing false negatives (FN) for DS but instead maintains a balanced approach between minimising false alarms and missed events. ACC was high, indicating the model effectively identified true positives (TP) and true negatives (TN), although it was not overly skewed toward perfect classification, reflecting a balanced handling of presence and absence in the weighted data.

For debris flows (DF), the AUC ranged from 0.930 (Terrain) to 0.966 (Time series – daily interval), MCC from 0.491 (1 week) to 0.603 (Time series – hourly interval), and F1 scores from 0.522 (1 week) to 0.630 (Time series – hourly interval). Interestingly, the model equipped with the hourly time series outperformed the daily time series one (although by a slight margin). This may suggest that debris flows may be more sensitive to short-duration, high-intensity rainfall events, which are better captured by finer temporal resolution. The model with DF showed large differences across configurations for REC and F2 and particularly low for the 1 week and 1 month cumulative rainfall predictors, indicating a higher number of FN for these configurations. This reflects the model's lower ability to predict landslide occurrences with cumulative rainfall predictors for DF. ACC remained high, confirming that the model still managed to maintain a good balance in correctly classifying both TP and TN.

For earth slides (ES), the AUC ranged from 0.910 (Terrain) to 0.962 (Time series – daily), MCC from 0.378 (Terrain) to 0.495 (Time series – daily), and F1 scores from 0.389 (Terrain) to 0.510 (Time series – daily). The time series (daily) configuration outperformed the others, closely followed by the "all scalar values" configuration. The time series (hourly) configuration performed worse, though still better than the terrain-only model. These results suggest that scalar values help with ES prediction, while finer temporal rainfall resolution (hourly) may not necessarily carrying valuable information for failure mechanism that may therefore be less affected by short-duration, high-intensity rainfall. In other words, a scalar aggregation of the rainfall may already be enough and a much richer information may actually end up acting as noise. REC and F2 showed larger differences across configurations, reflecting variability in false negatives. ACC remained sufficiently high, confirming that the model balanced TP and TN classifications well for ES.

For earth flows (EF), the AUC ranged from 0.933 (Terrain) to 0.979 (Time series – daily), MCC from 0.450 (Time series – hourly) to 0.549 (Time series – daily), and F1 scores from 0.453 (Time series – hourly) to 0.557 (Time series – daily). In this case, the time series (hourly) configuration performed the worst, indicating that finer rainfall resolution introduces noise rather than improving prediction for EF. Similar to ES, EF are less influenced by short-duration rainfall. The best configuration was time series (daily), followed by "all scalar values. "REC and F2 did not show significant differences across configurations, suggesting consistency in

false negative management. ACC remained high, reflecting a well-balanced model for EF. Interestingly, for both flow type landslides, the only terrain configuration is not the worst, suggesting that geomorphology and topography played a more important role than in the slide types, or at least compared to scalar rainfall configurations.

For rock slides (RS), the model performed poorly overall, likely due to the limited (just 221) sample of SU labelled as unstable. AUC ranged from 0.845 (1 month) to 0.968 (Time series – daily), MCC from 0.123 (3 weeks) to 0.222 (2 weeks), and F1 scores from 0.110 (3 weeks) to 0.193 (2 weeks). Despite these challenges, the model still sufficiently captured the spatial distribution of RS (see the next section). The best configuration in this case corresponded to the model equipped with the daily rainfall time series, an observation valid across AUC, ACC, and REC. However, the 2-week cumulative rainfall performed best for more comprehensive metrics like MCC, *k*, and F1. Interestingly, the 3-week cumulative rainfall configuration performed the worst, suggesting that the first event (incorporate in the 3 weeks cumulative but not in the 2 weeks) was less relevant for RS, while the rainy period before the second event was critical. RS results align more closely with ES and EF, as the finer rainfall resolution (hourly) was less relevant. Large differences in REC and F2 across configurations indicate that some configurations struggled with false negatives, particularly the 3-week and time series (hourly) configurations. Nonetheless, ACC remained relatively good, showing a reasonable balance between TP and TN classifications, even with poorer overall performance.

Due to the reflections shared so far on performance assessment, we want to emphasize that relying solely on AUC as a performance metric can be misleading. AUC often appears overly positive and unrepresentative of actual model performance. For instance, when comparing DS and RS, AUC suggests RS performs relatively well. However, a closer look at other metrics (like MCC, F1, and *k*) reveals significant differences between the two, showing that RS does not perform nearly as well as DS.

Finally, we computed average metrics across all landslide types for a general overview, shown in the bottom-right plot (Overall). The time series (daily) configuration stood out as the best, improving AUC by 1.66%, ACC by 3.21%, REC by 8.78%, MCC by 6.61%, *k* by 5.99%, F1 by 5.51%, and F2 by 6.61% compared to the best scalar configuration (all scalar values combined). Interestingly, the time series (hourly) configuration performed worse than using all scalar values overall.

4.6.2 Pseudo-probability maps

The resulting pseudo-probability maps from different configurations exhibit both similarities and distinct patterns. For conciseness, we present the best-performing scalar configuration (all scalar values), the two time series configurations (daily and hourly), and the terrain-only configuration as a baseline for comparison. For the readers interested enough to explore all the results, the outputs are here included as supplementary materials, whereas data and codes are available at the following GitHub repository: https://github.com/rodolforani/TNN_landslide_GESHAP.git.

For the cases presented here, we also generated confusion maps (CF) to compare the model outputs with the ground truth (presence/absence of landslides in the slope units). Figure 4.8 shows the results for DS. In the terrain-only configuration (Fig. 4.8a.1), the pseudo-probability map shows a more scattered and dispersed spatial pattern compared to the other configurations (Fig. 4.8b.1, 4.8c.1, 4.8d.1). This is further highlighted in the corresponding CF (Fig. 4.8a.2), where the presence of False Positives in the northwestern part of the region suggests that the terrain-only model struggles to capture the spatial distribution driven by rainfall as a triggering factor. However, the terrain configuration does show good accuracy in the southeastern portion of the region, likely due to the influence of static factors, in particular geology (presence of flysch rocks of sandstone and pelite). In the "all scalar values" configuration (Fig. 4.8b.1), the inclusion of rainfall data is evident, with higher overall pseudo-probability values compared to the other configurations. The rainfall contribution is particularly notable, improving the model's ability to predict landslide-prone areas. The time series configurations (daily and hourly, shown in Fig. 4.8c.1 and 4.8d.1) produce smoother pseudo-probability maps, with a more gradual spatial transition, especially in areas farther from the rainfall event epicentres. This smoothing effect can be attributed to the influence of the continuous rainfall data, as the time series model captures the temporal distribution of rainfall more effectively than scalar values. The interpolation of rainfall data (see Fig. 4.4) is inherently smoother, and the TNN model with time series inputs seems to give more weight to the rainfall distribution, enhancing the model's ability to identify spatial patterns more precisely than when using scalar rainfall values.

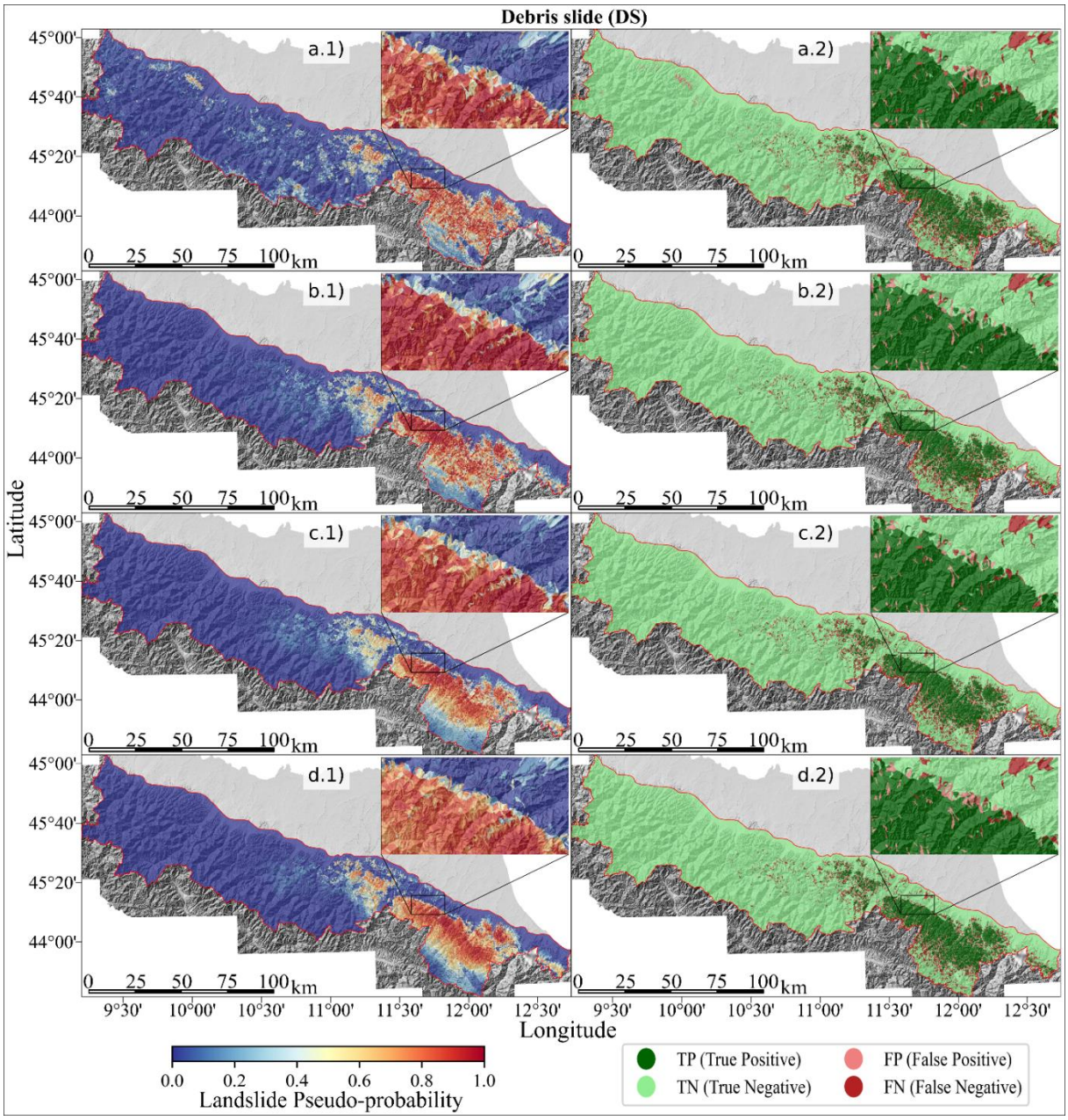


Figure 4.8. Spatial pseudo-probability (.1) and Confusion Matrix (*.2) Debris slide (DS) for the May 2023 Emilia Romagna events. a) Shown the "Only terrain" configuration. b) Shown the "All cumulative scalar rainfall" configurations. c) Shown the "Time series with daily intervals rainfall" configuration. d) Shown the "Time series with hourly intervals rainfall" configuration.*

Similar observations can be made for DF, as shown in S 4.7. The general trends between configurations remain consistent with those for DS, but key differences arise between the daily and hourly intervals. For DS, the daily interval produced slightly higher pseudo-probability values compared to the hourly interval. In contrast, for DF, the daily interval configuration (S 4.7c) shows lower maximum pseudo-probability values, with a smoother and more evenly dispersed distribution across the region. This results in a narrower range of pseudo-probability values, which are more homogeneously spread. On the other hand, the hourly interval configuration for DF (S 4.7d) shows more concentrated areas with higher pseudo-probability

values in specific locations. This tighter spatial concentration aligns with the results from the evaluation metrics in Fig. 4.7, where the hourly interval configuration performed slightly better for DF. This suggests that DF events are more sensitive to short-duration, high-intensity rainfall, which is better captured by the finer temporal resolution of hourly rainfall data.

For ES and EF (shown in S 4.8 and S 4.9, respectively in the supplementary material), we observed similar behaviours of DS and DF in the time series configurations, particularly with the daily interval. These configurations produced smoother pseudo-probability maps, though without narrowing the range of pseudo-probability values in ES and EF. The spatial distribution of landslides for these types is largely controlled by the underlying geology (Fig. 4.1), which is evident from the affected areas highlighted in the pseudo-probability maps, according to their density distribution (Fig. 4.4). The terrain-only configuration, again, struggles to accurately capture the true extent of landslide distribution, generating FP in the north-western part of the region. Additionally, moving from ES to EF maps more clustered distribution of pseudo-probability areas appears, partially notable from DS to DF. We believe these differences may be related to the distinct topographic features of flow-type landslides, which are more easily identifiable and better modelled by static factors alone. This aligns with the performance observed in the evaluation metrics, where flow types (DF and EF) showed relatively better modelling outcomes based on static terrain predictors.

Finally, for RS, as shown in S4.10 in the supplementary material, we observe similar trends to the other landslide types. The terrain-only and scalar rainfall configurations produce more spatially scattered pseudo-probability values, whereas the time series configurations, especially with the daily interval, result in a smoother and wider spatial distribution. Due to the smaller number of positive SUs, the model using the time series (daily) (S 4.10c.1), which is more influenced by the rainfall distribution, leads to a higher number of FP (S 4.10c.2). At the same time, this configuration also has fewer FN compared to the others, especially the time series (hourly) (S 4.10d.2). For RS, the daily interval appears to be more cautious, which explains why it has the highest accuracy (ACC) and recall (REC) values, even though it does not perform best in more comprehensive metrics like MCC, *k*, and F-scores. In contrast, the time series (hourly) configuration underestimates RS pseudo-probability, leading to a higher number of FN. This result indicates that while the daily interval may overestimate some areas, the hourly interval tends to underestimate RS pseudo-probability, resulting in a higher number of missed events, which is more problematic and should be avoided. Overall, of RS the terrain-only configuration results in a more reliable and realistic simulation, given more weight on the static factors that for this type of landslide is extremely important as experienced in the field. Especially the relationship between the aspect of the slope and the immersion of the geological stratification, so change abruptly from slope to adjacent ones, a condition that is not well captured including rainfall effect cause is spread across the area more uniformly.

4.6.3 Model explainability

The SHAP – GradientExplainer (SHAP_GE) results allow for understanding how the TNN works, this being particularly supported by looking into the spatiotemporal patterns of the SHAP_GE scores. As the most innovative part of the model we tested here is the use of the

rainfall as a continuous space-time signal, we focused the explainability on the precipitation, although the scores of the static factors on different landslide types are reported in the supplementary material in the form of a beeswarm plots (from S.15 to S.19). Being the precipitation contribution expressed in two dimensions, visualizing the SHAP_GE scores as a full data cube is a complex task. For this reason, we decoupled the contribution of the precipitation first looking into its temporal role and later in it spatial one.

Figure 4.9 corresponds to the first case, with the average SU scores being plotted as solid lines and bounded by the associate SU variability. Notably, the SHAP_GE score are intrinsically sums up to zero. Therefore, for visualization purposes we subsetted and plotted the scores only for SUs with pseudo-probability higher than 0.5 (arbitrary threshold). This representation captures the influence of two significant rainfall peaks, on days t_{-15} and t_{-1}, which correspond to two peaks in the SHAP_GE score mean and 95^{th} percentile range. In general, the mean SHAP_GE scores for all landslide types follow the rainfall trend, with increased score variability during rainy days. This variability arises because rainfall is not uniformly distributed across the region on any given day, leading to both positive and negative contributions to the scores depending on the spatial distribution of rainfall. We recall the reader that for the model, landslides presence/absence data are only expressed spatially, whereas the rainfall both spans over space and time. Therefore, it is particularly interesting to see how the model captures the role of a continuous precipitation and assigns scores that actually match positive rainfall discharge values. Another relevant consideration relates to three peaks, at t_{-16}, t_{-8} and t_{-7}, where rainfall that exhibits positive values is not necessarily associated to positive explainability scores.

Comparing the SHAP_GE across all failure modes, we can observe (Fig. 4.9) that the highest score variability is obtained for DS, and decreases progressively for other types, reaching a minimum for RS. In other words, this can be used to infer that rainfall exhibits a higher control on DS as compared to the other types. However, this could also be due to purely numerical reasons, because the amount of SUs with RS pseudo-probability > 0.5 is much smaller than for other landslide types, while the number of SUs with DS pseudo-probability > 0.5 is the largest. Additionally, ES and EF types appear less influenced by the first rainfall event (day t_{-15}), particularly EF (S 4.13). By contrast, RS seems equally affected by both rainfall events (S 4.14), highlighting its distinct response pattern.

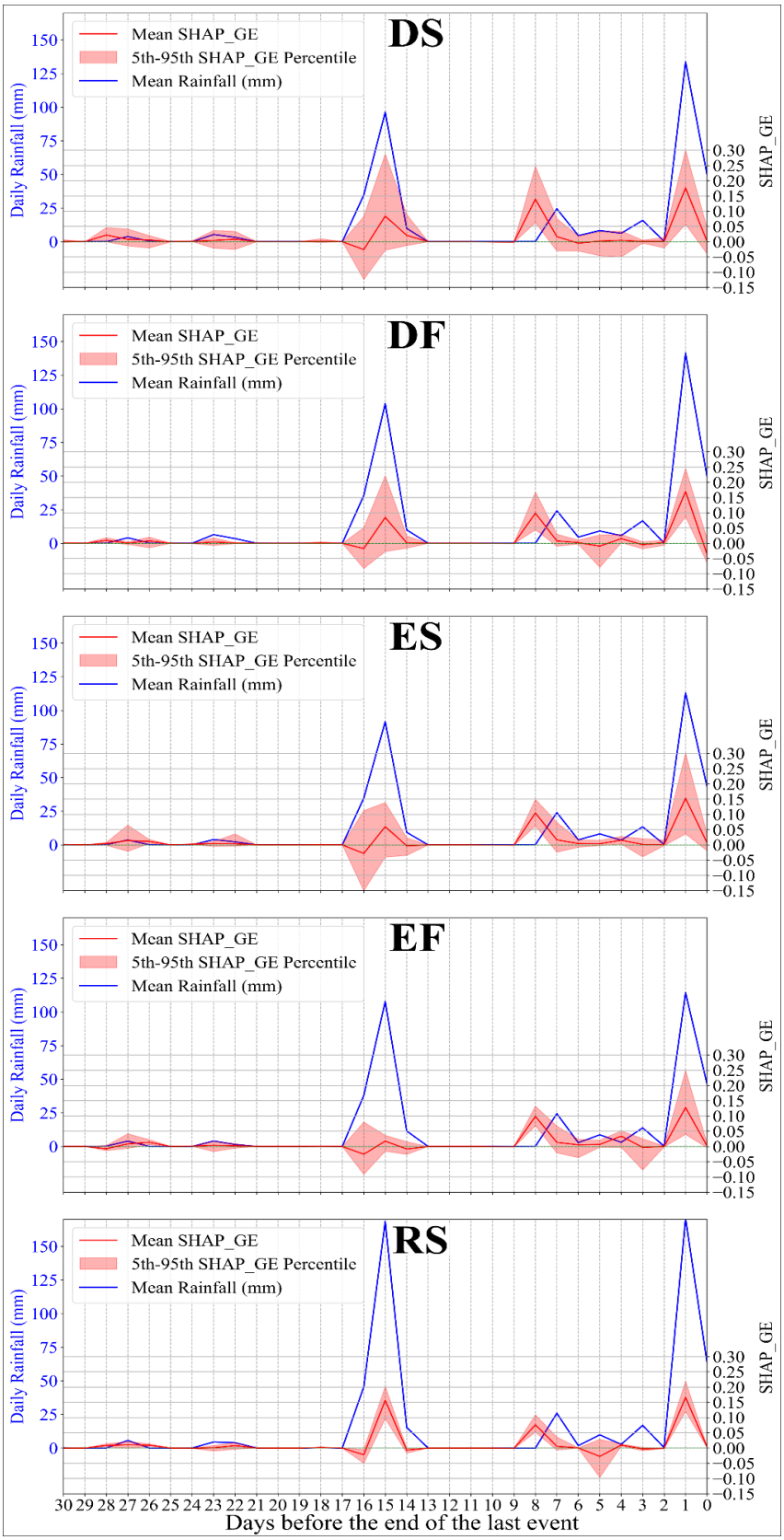


Figure 4.9. SHAP_GE scores (mean and 5-95 percentile) and mean rainfall of the SUs with pseudo-probability landslide > 0.5 (per each type) throughout time (daily).

As for studying the spatial score patterns, these are shown in Figure 4.10 for DS (and S 4.5, S 4.6, S 4.7, and S 4.8 for the other landslide types) for each day or the considered time series.

The same observation corresponding to the rainfall and score mismatch, mentioned above, is spatially highlighted here by comparing the scores to the rainfall maps shown in Figure 4.5. The first mismatch (16 days before the second event), rainfall is distributed evenly across the region at low intensity, without being concentrated in areas associated with landslide presence. Consequently, the model assigns an opposite contribution for this day compared to day t_{-15}, when rainfall intensity increased significantly in a specific area (Fig. 4.5) corresponding to landslide presence (Fig. 4.4). This results in higher SHAP_GE scores in that sector (Fig. 4.10). The mismatch on day's t_{-8} and t_{-7} show a different pattern. On t_{-8}, the model assigns high SHAP_GE scores despite extremely low rainfall, while on t_{-7}, it assigns low SHAP_GE scores despite high rainfall. At first glance, this behaviour appears contradictory. However, examining the rainfall and score distributions more closely (Figures 4.5 and 4.9), it becomes evident that on t_{-8}, areas with low rainfall receive low negative SHAP_GE scores, while areas with no rainfall receive high positive scores. This occurs because landslide presence (Fig. 4.4) corresponds to regions with no rainfall on that day. Mathematically, the model correctly interprets the absence of rainfall as contributing to landslide occurrence, as the areas without landslides, on that specific day, coincided with regions where rainfall was present. Even though this may seem contradictory to our physical assumptions on slope stability, we need to recall that this type of architectures works in relative terms, which is likely why the model may ends up assigns scores in such a way. Further reflections on these aspects will be presented in the Discussion section.

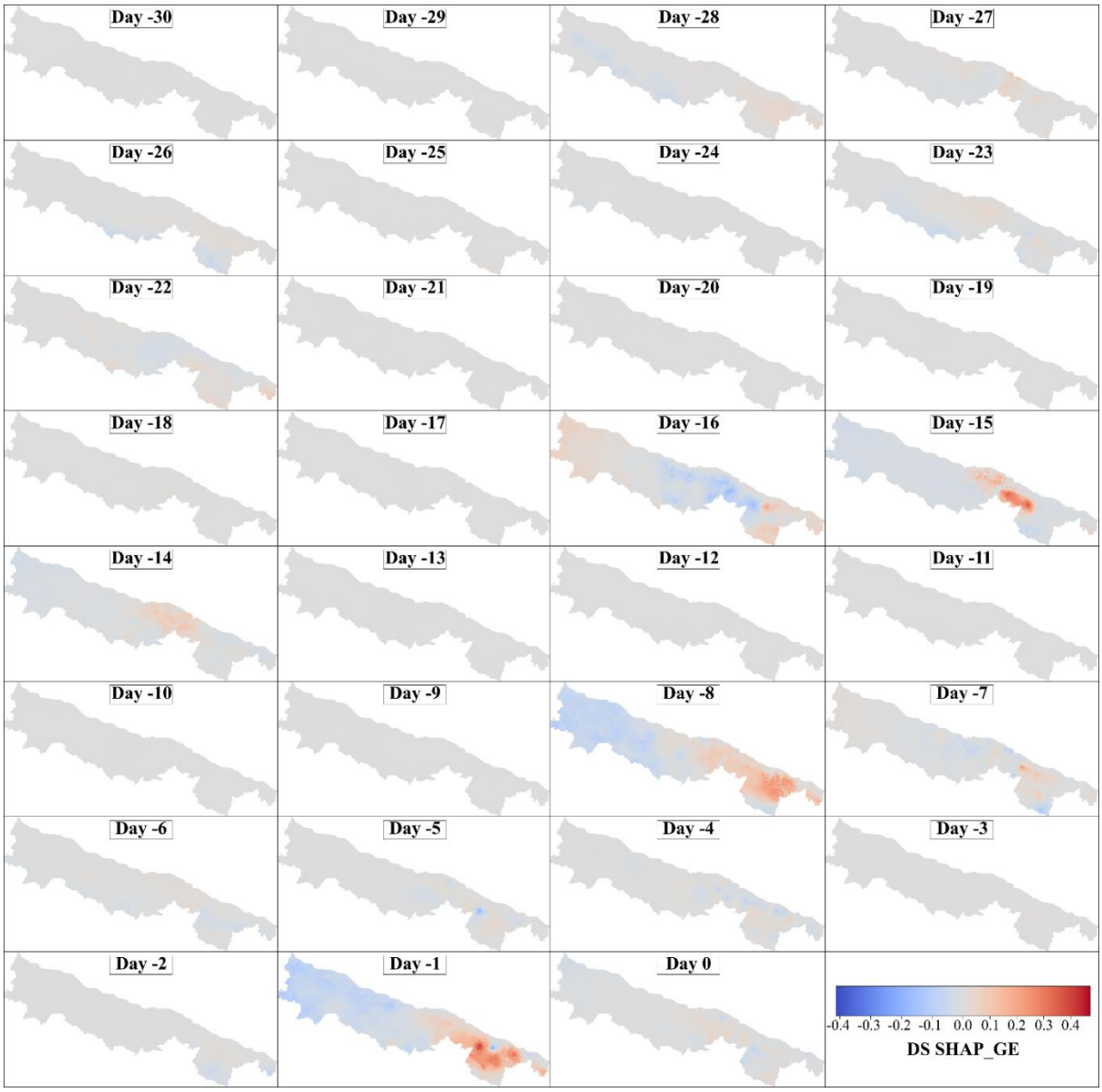


Figure 4.10. Spatial visualization of the SHAP_GE scores per day for the DS landslide type. Colour bar shared among all landslide types.

4.7 Discussion and future application

4.7.1 Insights from the Stratified Cross-Validation

The large imbalance between positive (1) and negative (0) samples can introduce bias in model performance depending on how the training, testing, and validation sets are split. To address this, we implemented a stratified cross-validation (SCV) strategy, which offers a more robust evaluation framework for our unbalanced dataset. We opted for a 5-fold SCV, maintaining the 0/1 class ratio across all folds and splits: training (80% of the total dataset), test (20% of the total), and internal validation (30% of the selected training set). Stratification was applied separately for each landslide type, preserving their respective class distributions across the folds. The analysis was conducted using the same model configuration described in Section 3.4,

including class weights per landslide type. No additional fold-specific tuning was applied, in order to maintain consistency and comparability.

The resulting performance metrics across the five test folds are summarized in Figure 4.11, showing: Mean ROC curves with ± standard deviation (SD), and Box plots for binary classification metrics. Overall, the SCV metrics are consistent with the evaluation results from the original 70/30 train/test split (with 30% validation on the training set), as reported in Figure 4.7. The standard deviation in ROC curves and the interquartile range (IQR) in box plots increase slightly as the number of positive SUs decreases, with the highest variability observed for the Rock Slide. Nonetheless, this variability remains limited, indicating no extreme outliers in model performance across folds. Therefore, the SCV results confirm the robustness of our original 70/30 train/test split and support its use in the modelling framework.

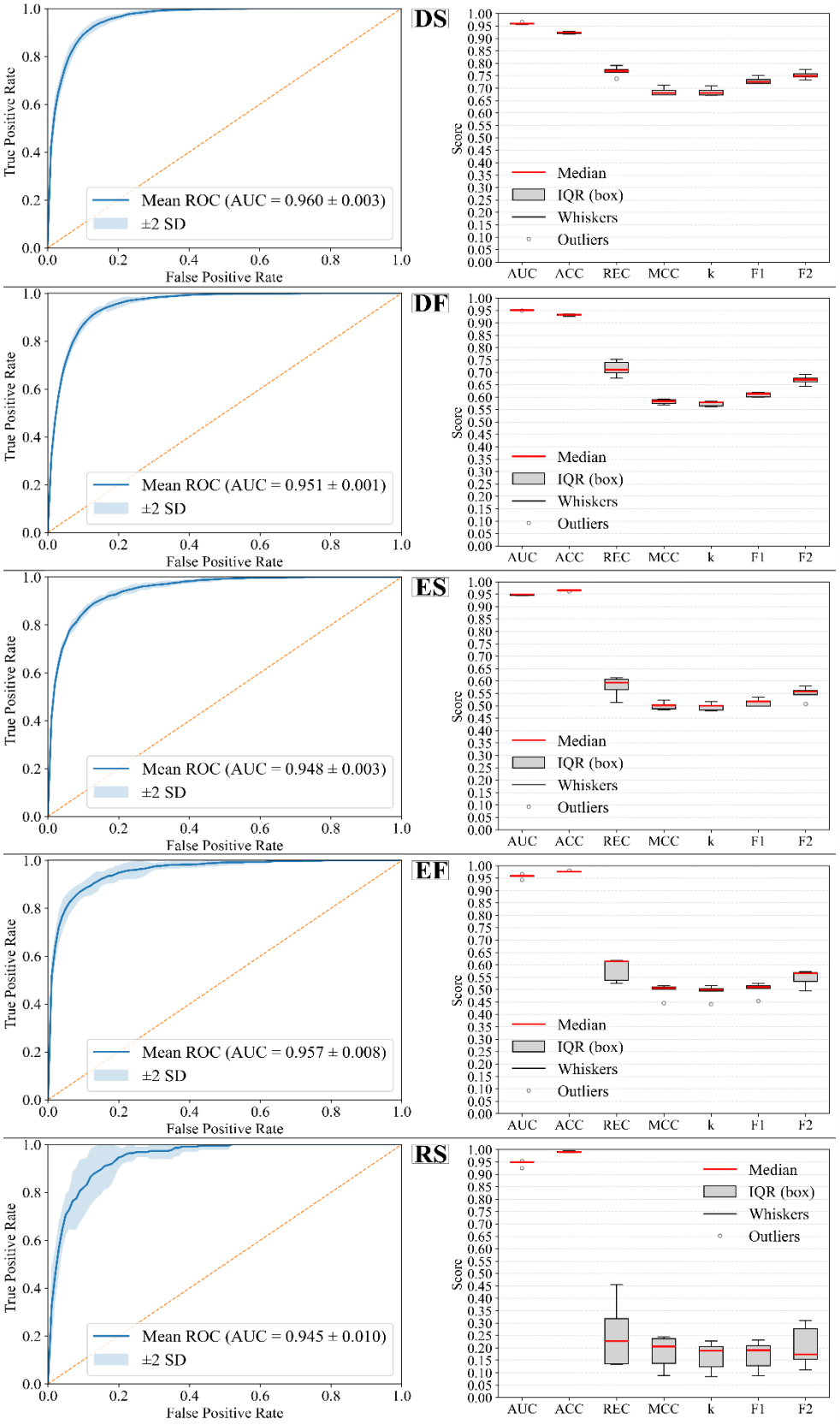


Figure 4.11. Stratified cross-validation performance metrics across five test folds: Mean ROC curves with standard deviation shaded areas; Box plots showing median, IQR (25th–75th percentile), whiskers (25th–1.5×IQR to 75th+1.5×IQR), and individual outliers.

4.7.2 Spatial and time model dimensionality literature comparison

The architecture we present here has several characteristics worth highlighting both with respect to the current literature trends as well as the potential for opening new directions. The first worth stressing has to do with the dimensionality that our model definition entails. In fact, the response variable only varies in space, thus implying a model structure intrinsically defined over 2D. However, the dynamic predictor we employed exhibits rainfall variations defined both over space and time. As a result, our model classifies the presence-absence labels against a 3D rainfall dataset, to which the information carried by 2D static predictors gets added. As a result, our model does not fall in a standard 2D spatial case, yet it does not fall in a standard 3D space-time case (Fang et al. 2024, Nocentini et al. 2024) with scalar dynamic predictors. For the sake of simplicity, we could refer to our model as a 2.5D one, whose extension to a full space-time case mainly requires the availability of landslide information over multiple times, explored only by (Fang et al. 2023, Lim et al 2024, and Moreno et al. 2025), summarized in Table 5. In this sense, upon collecting multiple landslide event inventories, or multiple landslide occurrences in space and time, our model could also be trained over multiple rainfall scenarios. This would have large implications not only for the dimensionality of the problem at hand but also for the potential applications of our model. In relation to the aspect dimensions, we should stress here that our model would still correspond to an exceptional case. In fact, a standard space-time model typically operates by explaining the variability of the response over slope and temporal units, with a temporal frequency matched between the response and the predictors. Interestingly, as our model looks into continuous time series, by modeling a data cube of presence absence instances, the architecture would demand a data cube inside a data cube, where each response would be associated to a precipitation time series. A second and even more relevant consideration related to the potential to use our model for operational uses. In the current setting, we tested our model on a single event. Even if our model proved to perform extremely well, and even as interesting as it is to showcase its use for a disastrous event such as the one happened in Emilia Romagna 2023, it is important to keep in mind that its architecture has only "seen" one landslide event. As a result, we expect the generalization of our model to be quite limited. This expectation is supported by examining the explainability scores, whose space-time pattern is intrinsically linked to the rainfall patterns that have led to the 80,000 landslides we modelled. By extending the dimension of the response variable to that of a multi-temporal landslide inventory, our model would inevitably learn how multiple examples of rainfall events leave a signature on the geomorphological response of a given landscape. Along the same lines, we should also mention the need not only to extend the temporal dimension of the response, but also the spatial one. In fact, the current setting makes our model tailored to a hilly landscape (such as the Apennine belt), thus making its use likely limited to such terrains and much less suited to be applied to much rougher topographies (for instance, Alpine slopes).

Model Type	*Dynamic Factor Form*	*Model/Architecture*	*Output Dimension*	*Notes*
Traditional 2D	One scalar summary factor of long-period climatic/PGA factor	Logistic Regression, RF, SVM, GAM, CNN etc…	Spatial (2D)	Widely used; ignores event-scale (triggering and preparatory factors)
Scalar 3D *(Lombardo et al. 2020; Steger et al. 2022; Fang et al. 2024; Nocentini et al. 2024)*	Multiple scalar summary factors over multiple time windows (e.g., 7-day cum., max rainfall, PGA)	Bayesian + LGCP, GAMM, Dynamic RF	Space-Time (3D) scenario	Scalar simplification of dynamic predictors; window choice affects sensitivity
Dynamic 3D *(Fang et al. 2023; Lim et al. 2024; Moreno et al. 2025)*	Rainfall as a continuous time series (e.g. 120 hours/14 days pre-landslide events per 20 years)	LSTM, GRD + NN, Functional Regression GAM	Full Space-time (3D)	LSTM/GRU introduce memory decay; Functional GAMM avoids this but lacks NN depth
2.5D proposed model	Full-time series (31-day daily/ 744-hour hourly)	TNN with attention + DNN	Pseudo-spatiotemporal (2.5D)	No memory decay with the attention mechanism, but currently lacking in multiple inventory training

Table 4.5. Comparative overview of model categories used in spatiotemporal landslide probability modelling. Terminology: RF=Random Forest, SVM=Support Vector Machine, GAMM=Generalized Additive Mixed Models, CNN=Convolutional Neural Network, LGCP=Log-Gaussian Cox Processes, LSTM=Long-Short term memory, GRD=Gated Recurrent Unit.

4.7.3 Limitations ad future improvements

Only upon satisfying the needs for a rich dataset could our model potentially be used as the basis for space-time predictions typical of early warnings. However, even in this case, the typical setting of landslide early warnings may require some more in-depth reflections. For instance, our model has been trained over a very large population of landslides, whose constituting types mainly encompass DS and DF, with the remaining types being much more underrepresented in comparison. This means that for a model such as ours to become generalizable in space and time while preserving its flexibility towards landslide types,

classified inventories should also become available. This is, unfortunately, not often the case, as discussed at length by Guzzetti et al. (2012). In fact, most of the landslide mapping practices still only report landslide identification points (e.g., Trigilia et al. 2010), without even necessarily referring to the failed locations but rather to the damaged ones (SWISS database: https://www.wsl.ch/en/natural-hazards/understanding-and-forecasting-floods/flood-and-landslide-damage-database/damage-map/) or even reporting the polygonal extent (Santangelo et al. 2015) or type (Ardizzone et al. 2023) and dates (Santangelo et al. 2023). Fortunately, the recent progress in automated mapping procedures is largely addressing this issue by providing structured, consistent and temporally distributed landslide inventories. As for the specific considerations on the failure types, this is still mostly a niche even among the automated mapping tools in the literature, though we few examples have started to appear (see the work of Bhuyan et al. 2024).

By combining the reflections on the operational use and the flexibility of our model towards different landslide types, here we stress applications that could potentially involve tailor-made mitigation strategies. By assuming that our model could be trained over multiple events, thus learning the generic role of rainfall and rainfall extremes (even sequential extremes such as the present one), not only, it could be used to simulate on time-scales typical of early warnings but also of master plans. We recall here that our model has a spatial dimension expressed as SU and a temporal one that potentially allows for daily predictions. Therefore, instead of simulating pseudo-probability patterns for weather forecast data, one could use climate data instead. In short, we potentially see applications where climate scenarios could be used to generate long-term predictions, and in light of predictions tailored towards landslide type, decision makers could be informed about what failure mechanism would likely affect specific locations at specific years of interest.

Aside from these generic considerations and future plans, below we would like to reflect on the added values (or not) of using time series as compared to using scalar aggregated rainfall values. The latter is, by definition, a simplification of the continuous rainfall signal. However, it does not necessarily imply a large improvement in modelling performance. For instance, Moreno et al. (2025) tested a similar setting by implementing a space-time functional regression, which the authors compared to a scalar use of the rainfall alone, and in addition to terrain characteristics. What emerged was that a model equipped with terrain factors and scalar rainfall essentially performed very close to the one employing the whole time series. This could be due to two reasons. First, the model that Moreno and co-authors tested was defined for a province in the Italian Alps. Therefore, the terrain was extremely steep and the role of rainfall, whether expressed as scalar aggregated values or time series, only defined a triggering threshold, without requiring the need to differentiate between the two ways of incorporating the precipitation into the model. Second, the model only examined the hourly rainfall time series over a time window of just five days. In this sense, the overall difference in rainfall scalar or continuous amounts over time must be less as compared to a model such as ours trained over 31 days. As a result, the idea of using space-time models and including predictor time-series in them is at such an embryonal stage that requires many more experimentation before finding a

suitable setting that would support pushing those for operational purposes. Nevertheless, it has been proven to produce reliable results, and it fits perfectly in a climate change environment, where we expect an increase in intense rainstorms or very wet conditions over short periods, as well as prolonged drought conditions (Myhre et al., 2019) and their influence on landsliding (Ma et al., 2025; Gariano & Rianna, 2025). Therefore, our model allows us to avoid assumptions and directly understand the influence of both the preparatory factors (wet or drought) and the triggering factors (intense or persistent rainfall).

Another element of novelty in this contribution relates to the way the model architecture can be explored to understand why certain pseudo-probability values are assigned for each slope unit. The scientific area pertaining to explainable AI has seen a drastic increase in recent years, with hundreds of contributions published so far since the inception of the concept in 2017 with the XAI program (Gunning & Aha, 2019). However, the implementation of explainable machine learning models has mostly (across many scientific branches, e.g., Jacovi 2023; Saraswat et al., 2022) if not exclusively (in landslide science, e.g. Youssef et al. 2022; Dahal et al. 2023) applied to the spatial case. Here, we extend the explainable framework to the space-time domain, thus deriving scores for each SU as well as each day of the time series. In some way, this is both a blessing and a curse. In fact, the effect of a predictor whose contribution is based solely on spatial patterns is relatively easy to explain (Nan Wang et al. 2024). However, our explainability scores are linked through time, or better, their influence on the model is not independent from another day of the time series. Therefore, this relative estimation makes it so that the interpretation becomes less straightforward. More experimentation is certainly required to evaluate how informative these tools can be. This being said, neural network architectures have recently been proven to be adaptive to dynamic cases when the loss function includes physical constraints. This branch of computer science corresponds to physics-informed neural networks whose implementation and use would be natively interpretable rather than being so in a post-processing phase. To conclude our discussions, we would like to stress a potential direction to take where the modeling architecture could be the same as the one presented here, but whose explainability could even be better. In fact, rainfall is not necessarily a physical parameter, as its units do not appear in any geotechnical slope stability model. Conversely, a parameter whose spatio-temporal expression is much closer to the mechanics of slope stability is the soil moisture, a type of data that can be consistently estimated in space and time through satellite products and that may be seen as a proxy of the water infiltrating into the slope. This is for instance the direction that has been recently explored by Peiro et al. (2024). In their work, the soil moisture was used instead of rainfall, proving high performance and interpretability. This could also be a venue to test, by introducing soil moisture time series as part of our TNN. Interestingly, this would not limit the simulation capacity of our model in the short term, as global soil moisture products do exist and therefore the early warning potential of our tool could still hold. However, long-term simulations would definitely be limited as the current state-of-the-art has not reached a level where reliable soil-moisture scenarios are available for the community at large.

4.8 Conclusions

Our "2.5D" model explained the distribution of landslides mapped after the two sequential rainfall extreme events occurred in Emilia Romagna in May 2023 as a function of rainfall time series. This is a task achieved thanks to a complex Transformer architecture, tailored not only to predict landslide occurrence per se but rather diversifies per landslide type. This experiment further adds on the recent literature on space-time landslide modeling even though our architecture does not formally satisfy the space-time definition. Instead, it offers a viewpoint on the use of rainfall time series as part of landslide data-driven solutions whereby the inclusion of the whole signal implies dropping the need to cumulate rain over specific windows. This idea of using time-series is something that has recently been tested. However, the previous cases adopted architectures whose rainfall influence is conveyed by sequential models. In this sense, previous examples, by definitions, operated by "forgetting" rainfall amounts distributed far away in time from the landslide occurrence date. This is why we opted for an TNN architecture here that employs an attention mechanism through which the rainfall contribution does not forcefully decays over time. In our case, if a decay would indeed been shown in the distribution of the explainability scores, this would be a characteristic that our model would return free of a priori modeling constraints. Indeed, the resulting explainability in time and space-time do not show particularly dependency from the previous rainfall events, instead, but is more influenced by the rainfall on the day itself. We stress this aspect here specifically because climate change is surprisingly leading to weather extremes with hundreds of years of return periods happening even shortly one after the other. Therefore, recurrent or "forgetful" architectures are not suitable and would likely become less suitable as the effect of global warming would exacerbate landslide extremes. Future experiments will feature a proper application in a data cube whose distribution of landslides changes in space and time. Upon completing these tests, we would have more insight into whether the model presented here holds the potential to become operational.

4.9 References chapter 4

Abbate, A., Papini, M., & Longoni, L. (2021). Analysis of meteorological parameters triggering rainfall-induced landslide: A review of 70 years in Valtellina. *Natural Hazards and Earth System Sciences, 21*(7), 2041–2058. https://doi.org/10.5194/nhess-21-2041-2021

Ahmad, S. M., Sadhasivam, N., Lisa, M., Lombardo, L., Emil, M. K., Fadel, I., van Westen, C. J., Ahmed, A., & Tanyas, H. (2022). Standing on the shoulder of a giant landslide: An InSAR look at a slow-moving hillslope under melting glaciers in the western Karakoram. *EarthArXiv*. https://doi.org/10.31223/X5P63T

Ahmed, M., Tanyas, H., Huser, R., Dahal, A., Titti, G., Borgatti, L., & Lombardo, L. (2023). Dynamic rainfall-induced landslide susceptibility: A step towards a unified forecasting system. *International Journal of Applied Earth Observation and Geoinformation, 125*, 103593. https://doi.org/10.1016/j.jag.2023.103593

Amato, G., Eisank, C., Castro-Camilo, D., & Lombardo, L. (2019). Accounting for covariate distributions in slope-unit-based landslide susceptibility models. A case study in the alpine environment. Engineering geology, 260, 105237. https://doi.org/10.1016/j.enggeo.2019.105237

Altman, D. G., & Bland, J. M. (1994). Diagnostic tests. 1: Sensitivity and specificity. *BMJ, 308*(6943), 1552. https://doi.org/10.1136/bmj.308.6943.1552

Alvioli, M., Marchesini, I., Reichenbach, P., Rossi, M., Ardizzone, F., Fiorucci, F., & Guzzetti, F. (2016). Automatic delineation of geomorphological slope units with r.slopeunits v1.0 and their optimization for landslide susceptibility modeling. *Geoscientific Model Development, 9*(11), 3975–3991. https://doi.org/10.5194/gmd-9-3975-2016

Angelov, P. P., Soares, E. A., Jiang, R., Arnold, N. I., & Atkinson, P. M. (2021). Explainable artificial intelligence: An analytical review. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 11*(5), e1424. https://doi.org/10.1002/widm.1424

Ardizzone, F., Bucci, F., Cardinali, M., Fiorucci, F., Pisano, L., Santangelo, M., & Zumpano, V. (2023). Geomorphological landslide inventory map of the Daunia Apennines, southern Italy. *Earth System Science Data, 15*, 753–767. https://doi.org/10.5194/essd-15-753-2023

Ba, Q., Chen, Y., Deng, S., Yang, J., & Li, H. (2018). A comparison of slope units and grid cells as mapping units for landslide susceptibility assessment. *Earth Science Informatics, 11*, 373–388. https://doi.org/10.1007/s12145-018-0335-9

Bayer, B., Simoni, A., Mulas, M., Corsini, A., & Schmidt, D. (2018). Deformation responses of slow moving landslides to seasonal rainfall in the Northern Apennines, measured by InSAR. *Geomorphology, 308*, 293–306. https://doi.org/10.1016/j.geomorph.2018.02.020

Berti, M., Martina, M. L. V., Franceschini, S., Pignone, S., Simoni, A., & Pizziolo, M. (2012). Probabilistic rainfall thresholds for landslide occurrence using a Bayesian approach. *Journal of Geophysical Research: Earth Surface, 117*(F4). https://doi.org/10.1029/2012JF002367

Berti, M., Pizziolo, M., Scaroni, M., Generali, M., Critelli, V., Mulas, M., Tondo, M., Lelli, F., Fabbiani, C., Ronchetti, F., Ciccarese, G., Dal Seno, N., Ioriatti, E., Rani, R., Zuccarini, A., Simonelli, T., & Corsini, A. (2024). RER2023: The landslide inventory dataset of the May 2023 Emilia-Romagna event. *Earth System Science Data Discussions*, 1–24. https://doi.org/10.5194/essd-17-1055-2025

Bhuyan, K., Rana, K., Ferrer, J. V., Cotton, F., Ozturk, U., Catani, F., & Malik, N. (2024). Landslide topology uncovers failure movements. *Nature Communications, 15*, 2633. https://doi.org/10.1038/s41467-024-46741-7

Brath, A., Casagli, N., Marani, M., Mercogliano, P., & Motta, R. (2023). *Rapporto della Commissione tecnico-scientifica istituita con deliberazione della Giunta Regionale n. 984/2023 e determinazione dirigenziale 14641/2023, al fine di analizzare gli eventi meteorologici estremi del mese di maggio 2023*. Technical Report, Regione Emilia-Romagna, 147 pp. Retrieved from https://www.regione.emilia-romagna.it/alluvione/rapporto-della-commissione-tecnico-scientifica

Brenning, A. (2005). Spatial prediction models for landslide hazards: Review, comparison and evaluation. *Natural Hazards and Earth System Sciences, 5*(6), 853–862. https://doi.org/10.5194/nhess-5-853-2005

Caleca, F., Lombardo, L., Steger, S., Tanyas, H., Raspini, F., Dahal, A., Nefros, C., Mărgărint, M. C., Drouin, V., Jemec-Auflič, M., Novellino, A., Tonini, M., Loche, M., Casagli, N., & Tofani, V. (2025). Pan-European landslide risk assessment: From theory to practice. *Reviews of Geophysics, 63*(1), e2023RG000825. https://doi.org/10.1029/2023RG000825

Campforts, B., Shobe, C. M., Overeem, I., & Tucker, G. E. (2022). The art of landslides: How stochastic mass wasting shapes topography and influences landscape dynamics. *Journal of Geophysical Research: Earth Surface, 127*(8), e2022JF006745. https://doi.org/10.1029/2022JF006745

Carrara, A., Crosta, G., & Frattini, P. (2003). Geomorphological and historical data in assessing landslide hazard. *Earth Surface Processes and Landforms, 28*(10), 1125–1142. https://doi.org/10.1002/esp.545

Che, Z., Purushotham, S., Cho, K., Sontag, D., & Liu, Y. (2018). Recurrent neural networks for multivariate time series with missing values. *Scientific Reports, 8*(1), 6085. https://doi.org/10.1038/s41598-018-24271-9

Chicco, D., & Jurman, G. (2020). The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation. *BMC Genomics, 21*, 6. https://doi.org/10.1186/s12864-019-6413-7

Chikalamo, E. E., Mavrouli, O. C., Ettema, J., van Westen, C. J., Muntohar, A. S., & Mustofa, A. (2020). Satellite-derived rainfall thresholds for landslide early warning in Bogowonto Catchment, Central Java, Indonesia. *International Journal of Applied Earth Observation and Geoinformation, 89*, 102093. https://doi.org/10.1016/j.jag.2020.102093

Conti, P., Cornamusini, G., & Carmignani, L. (2020). An outline of the geology of the Northern Apennines (Italy), with geological map at 1: 250,000 scale. *Italian Journal of Geosciences, 139*(2), 149-194. https://doi.org/10.3301/IJG.2019.25

Crosta, G. B., & Frattini, P. (2008). Rainfall-induced landslides and debris flows. *Hydrological Processes, 22*(4), 473–477. https://doi.org/10.1002/hyp.6885

Dahal, A., & Lombardo, L. (2025). Towards physics-informed neural networks for landslide prediction. *Engineering Geology, 344*, 107852. https://doi.org/10.1016/j.enggeo.2024.107852

Dahal, A., Tanyaş, H., & Lombardo, L. (2024). Full seismic waveform analysis combined with transformer neural networks improves coseismic landslide prediction. *Communications Earth & Environment, 5*(1), 75. https://doi.org/10.1038/s43247-024-01243-8

Dahal, A., & Lombardo, L. (2023). Explainable artificial intelligence in geoscience: A glimpse into the future of landslide susceptibility modeling. *Computers & Geosciences, 176*, 105364. https://doi.org/10.1016/j.cageo.2023.105364

Dai, F. C., Lee, C. F., & Ngai, Y. Y. (2002). Landslide risk assessment and management: An overview. *Engineering Geology, 64*(1), 65–87. https://doi.org/10.1016/S0013-7952(01)00093-X

Das, A., & Rad, P. (2020). Opportunities and challenges in explainable artificial intelligence (XAI): A survey. *arXiv preprint arXiv:2006.11371*.

Di Napoli, M., Eroglu, C., van den Bout, B., Di Martire, D., Tanyas, H., & Lombardo, L. (2024). Space-time modeling of cascading hazards: Chaining wildfires, rainfall and landslide events through machine learning. *Catena, 246*, 108452. https://doi.org/10.1016/j.catena.2024.108452

Fan, X., Scaringi, G., Korup, O., West, A. J., van Westen, C. J., Tanyas, H., Hovius, N., Hales, T. C., Jibson, R. W., Allstadt, K. E., & Zhang, L. (2019). Earthquake-induced chains of geologic hazards: Patterns, mechanisms, and impacts. *Reviews of Geophysics, 57*(2), 421–503. https://doi.org/10.1029/2018RG000626

Fang, Z., Tanyas, H., Gorum, T., Dahal, A., Wang, Y., & Lombardo, L. (2023). Speech-recognition in landslide predictive modelling: A case for a next generation early warning system. *Environmental Modelling & Software, 170*, 105833. https://doi.org/10.1016/j.envsoft.2023.105833

Fang, Z., Wang, Y., van Westen, C., & Lombardo, L. (2024). Space–time landslide susceptibility modeling based on data-driven methods. *Mathematical Geosciences, 56*(6), 1335–1354. https://doi.org/10.1007/s11004-023-10105-6

Fang, Z., Morales, A. B., Wang, Y., & Lombardo, L. (2025). Climate change has increased rainfall-induced landslide damages in central China. *International Journal of Disaster Risk Reduction*, 105320. https://doi.org/10.1016/j.ijdrr.2025.105320

Fatourechi, M., Ward, R. K., Mason, S. G., Huggins, J., Schlögl, A., & Birch, G. E. (2008, December). Comparison of evaluation metrics in classification applications with imbalanced datasets. In *2008 Seventh International Conference on Machine Learning and Applications* (pp. 777–782). IEEE. https://doi.org/10.1109/ICMLA.2008.131

Fawcett, T. (2006). An introduction to ROC analysis. *Pattern Recognition Letters, 27*(8), 861–874. https://doi.org/10.1016/j.patrec.2005.10.010

Fidan, S., Tanyaş, H., Akbaş, A., Lombardo, L., Petley, D. N., & Görüm, T. (2024). Understanding fatal landslides at global scales: A summary of topographic, climatic, and anthropogenic perspectives. *Natural Hazards, 120*(7), 6437–6455. https://doi.org/10.1007/s11069-024-06487-3

Foraci, R., Tesini, M. S., Nanni, S., Antolini, G., & Pavan, V. (2023). L'inquadramento meteo e idrologico degli eventi. *Ecoscienza, ARPAE Emilia-Romagna, XIV*(5), 20–24.

Franke, R. (1982). Scattered data interpolation: Tests of some methods. *Mathematics of Computation, 38*(157), 181–200. https://doi.org/10.1090/S0025-5718-1982-0637296-4

Frattini, P., Crosta, G., & Carrara, A. (2010). Techniques for evaluating the performance of landslide susceptibility models. *Engineering Geology, 111*(1–4), 62–72. https://doi.org/10.1016/j.enggeo.2009.12.004

Froude, M. J., & Petley, D. N. (2018). Global fatal landslide occurrence from 2004 to 2016. *Natural Hazards and Earth System Sciences, 18*(8), 2161–2181. https://doi.org/10.5194/nhess-18-2161-2018

Gariano, S. L., & Rianna, G. (2025). How will the projected climate change influence rainfall-induced landslides in Europe? A review of modelling approaches. *Landslides*, 1-17. https://doi.org/10.1007/s10346-025-02550-7

Gunning, D., & Aha, D. (2019). DARPA's explainable artificial intelligence (XAI) program. *AI Magazine, 40*(2), 44–58. https://doi.org/10.1609/aimag.v40i2.2850

Gunning, D., Stefik, M., Choi, J., Miller, T., Stumpf, S., & Yang, G. Z. (2019). XAI—Explainable artificial intelligence. *Science Robotics*, *4*(37), eaay7120. https://doi.org/10.1126/scirobotics.aay7120

Günther, A., Reichenbach, P., Malet, J. P., Van Den Eeckhaut, M., Hervás, J., Dashwood, C., & Guzzetti, F. (2013). Tier-based approaches for landslide susceptibility assessment in Europe. *Landslides, 10*(5), 529–546. https://doi.org/10.1007/s10346-013-0435-z

Gupta, K., & Satyam, N. (2024). Optimizing seismic hazard inputs for co-seismic landslide susceptibility mapping: A probabilistic analysis. *Natural Hazards*, *120*(9), 8459–8481. https://doi.org/10.1007/s11069-024-06517-0

Guzzetti, F., Mondini, A. C., Cardinali, M., Fiorucci, F., Santangelo, M., & Chang, K. T. (2012). Landslide inventory maps: New tools for an old problem. *Earth-Science Reviews, 112*(1–2), 42–66. https://doi.org/10.1016/j.earscirev.2012.02.001

Hanley, J. A., & McNeil, B. J. (1982). The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology, 143*(1), 29–36. https://doi.org/10.1148/radiology.143.1.7063747

He, J., Tanyas, H., Dahal, A., Huang, D., & Lombardo, L. (2025). Spatial prediction of InSAR-derived hillslope velocities via deep learning. *Bulletin of Engineering Geology and the Environment*, *84*(3), 1-14. https://doi.org/10.1007/s10064-025-04161-x

Hungr, O., Leroueil, S., & Picarelli, L. (2014). The Varnes classification of landslide types, an update. *Landslides, 11*(2), 167–194. https://doi.org/10.1007/s10346-013-0436-y

Jacovi, A. (2023). Trends in explainable AI (XAI) literature. *arXiv preprint arXiv:2301.05433*.

Kawagoe, S., Kazama, S., & Sarukkalige, P. R. (2009). Assessment of snowmelt triggered landslide hazard and risk in Japan. *Cold Regions Science and Technology, 58*(3), 120–129. https://doi.org/10.1016/j.coldregions.2009.05.004

Kingma, D. P., & Ba, J. (2014). Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*.

Koukouvelas, I., Litoseliti, A., Nikolakopoulos, K., & Zygouri, V. (2015). Earthquake triggered rock falls and their role in the development of a rock slope: The case of Skolis Mountain, Greece. *Engineering Geology, 191,* 71–85. https://doi.org/10.1016/j.enggeo.2015.03.010

Leshchinsky, B., Olsen, M. J., Mohney, C., O'Banion, M., Bunn, M., Allan, J., & McClung, R. (2019). Quantifying the sensitivity of progressive landslide movements to failure geometry, undercutting processes and hydrological changes. *Journal of Geophysical Research: Earth Surface, 124*(2), 616–638. https://doi.org/10.1029/2018JF004833

Lim, J., Santinelli, G., Dahal, A., Vrieling, A., & Lombardo, L. (2024). An ensemble neural network approach for space–time landslide predictive modelling. *International Journal of Applied Earth Observation and Geoinformation, 132*, 104037. https://doi.org/10.1016/j.jag.2024.104037

Lima, P., Steger, S., Glade, T., & Mergili, M. (2023). Conventional data-driven landslide susceptibility models may only tell us half of the story: Potential underestimation of landslide impact areas depending on the modeling design. *Geomorphology*, 430, 108638. https://doi.org/10.1016/j.geomorph.2023.108638

Lombardo, L., & Tanyas, H. (2022). From scenario-based seismic hazard to scenario-based landslide hazard: Fast-forwarding to the future via statistical simulations. *Stochastic Environmental Research and Risk Assessment, 36*(8), 2229–2242. https://doi.org/10.1007/s00477-021-02020-1

Lundberg, S. M., & Lee, S. I. (2017). A unified approach to interpreting model predictions. *arXiv preprint arXiv:1705.07874*.

Lundberg, S. (2022). SHAP (SHapley Additive exPlanations). Retrieved October 2022, from https://github.com/slundberg/shap

Ma, P., Chen, L., Yu, C., Zhu, Q., Ding, Y., Wu, Z., Li, H., Tian, C. & Fan, X. (2025). Dynamic landslide susceptibility mapping over last three decades to uncover variations in landslide causation in subtropical urban mountainous areas. *Remote Sensing of Environment*, *326*, 114800. https://doi.org/10.1016/j.rse.2025.114800

Marín-Rodríguez, N. J., Vega, J., Zanabria, O. B., González-Ruiz, J. D., & Botero, S. (2024). Towards an understanding of landslide risk assessment and its economic losses: A scientometric analysis. *Landslides, 21*(8), 1865–1881. https://doi.org/10.1007/s10346-024-02272-2

Matthews, B. W. (1975). Comparison of the predicted and observed secondary structure of T4 phage lysozyme. *Biochimica et Biophysica Acta (BBA) - Protein Structure, 405*(2), 442–451. https://doi.org/10.1016/0005-2795(75)90109-9

Menditto, A., Patriarca, M., & Magnusson, B. (2007). Understanding the meaning of accuracy, trueness and precision. *Accreditation and Quality Assurance*, *12*(1), 45–47. https://doi.org/10.1007/s00769-006-0191-z

Minder, J. R., Roe, G. H., & Montgomery, D. R. (2009). Spatial patterns of rainfall and shallow landslide susceptibility. *Water Resources Research, 45*(4). https://doi.org/10.1029/2008WR007027

Molnar, C. (2019). *Interpretable Machine Learning: A Guide for Making Black Box Models Explainable*. https://christophm.github.io/interpretable-ml-book/

Monsieurs, E., Dewitte, O., & Demoulin, A. (2019). A susceptibility-based rainfall threshold approach for landslide occurrence. *Natural Hazards and Earth System Sciences, 19*(4), 775–789. https://doi.org/10.5194/nhess-19-775-2019

Myhre, G., Alterskjær, K., Stjern, C. W., Hodnebrog, Ø., Marelle, L., Samset, B. H., Sillman, J., Schaller, N., Fisher, E., Schulz, M. & Stohl, A. (2019). Frequency of extreme precipitation increases extensively with event rareness under global warming. *Scientific reports*, *9*(1), 16063. https://doi.org/10.1038/s41598-019-52277-4

Nocentini, N., Rosi, A., Piciullo, L., Liu, Z., Segoni, S., & Fanti, R. (2024). Regional-scale spatiotemporal landslide probability assessment through machine learning and potential applications for operational warning systems: A case study in Kvam (Norway). *Landslides, 21*(10), 2369–2387. https://doi.org/10.1007/s10346-024-02287-9

Opitz, T., Bakka, H., Huser, R., & Lombardo, L. (2022). High-resolution Bayesian mapping of landslide hazard with unobserved trigger event. *The Annals of Applied Statistics, 16*(3), 1653–1675. https://doi.org/10.1214/21-AOAS1561

Ozturk, U., Saito, H., Matsushi, Y., Crisologo, I., & Schwanghart, W. (2021). Can global rainfall estimates (satellite and reanalysis) aid landslide hindcasting? *Landslides, 18*(9), 3119–3133. https://doi.org/10.1007/s10346-021-01689-3

Pacheco Quevedo, R., Velastegui-Montoya, A., Montalván-Burbano, N., Morante-Carballo, F., Korup, O., & Daleles Rennó, C. (2023). Land use and land cover as a conditioning factor in landslide susceptibility: A literature review. *Landslides, 20*(5), 967–982. https://doi.org/10.1007/s10346-022-02020-4

Peiro, Y., Volpe, E., Ciabatta, L., & Cattoni, E. (2024). High resolution precipitation and soil moisture data integration for landslide susceptibility mapping. *Geosciences, 14*(12), 330. https://doi.org/10.3390/geosciences14120330

Powers, D. M. (2020). Evaluation: From precision, recall and F-measure to ROC, informedness, markedness and correlation. *arXiv preprint arXiv:2010.16061*. https://arxiv.org/abs/2010.16061

Pradhan, S., Toll, D. G., Rosser, N. J., & Brain, M. J. (2022). An investigation of the combined effect of rainfall and road cut on landsliding. *Engineering Geology, 307*, 106787. https://doi.org/10.1016/j.enggeo.2022.106787

Reichenbach, P., Rossi, M., Malamud, B. D., Mihir, M., & Guzzetti, F. (2018). A review of statistically-based landslide susceptibility models. *Earth-Science Reviews, 180*, 60–91. https://doi.org/10.1016/j.earscirev.2018.03.001

Sabokbar, H. F., Roodposhti, M. S., & Tazik, E. (2014). Landslide susceptibility mapping using geographically-weighted principal component analysis. *Geomorphology, 226*, 15–24. https://doi.org/10.1016/j.geomorph.2014.07.028

Saha, S., Roy, J., Pradhan, B., & Hembram, T. K. (2021). Hybrid ensemble machine learning approaches for landslide susceptibility mapping using different sampling ratios at East Sikkim Himalayan, India. *Advances in Space Research*, *68*(7), 2819-2840. https://doi.org/10.1016/j.asr.2021.05.018

Santangelo, M., Marchesini, I., Bucci, F., Cardinali, M., Fiorucci, F., & Guzzetti, F. (2015). An approach to reduce mapping errors in the production of landslide inventory maps. *Natural Hazards and Earth System Sciences*, *15*(9), 2111–2126. https://doi.org/10.5194/nhess-15-2111-2015

Santangelo, M., Althuwaynee, O., Alvioli, M., Bucci, F., Cardinali, M., Fiorucci, F., & Guzzetti, F. (2023). Inventory of landslides triggered by an extreme rainfall event in Marche-Umbria, Italy, on 15 September 2022. *Scientific Data*, *10*, 427. https://doi.org/10.1038/s41597-023-02336-3

Saraswat, D., Bhattacharya, P., Verma, A., Prasad, V. K., Tanwar, S., Sharma, G., Bokoro, P. N., & Sharma, R. (2022). Explainable AI for Healthcare 5.0: Opportunities and Challenges. *IEEE Access, 10*, 84486–84517. https://doi.org/10.1109/ACCESS.2022.3197671

Schuster, G. (2024). *Machine Learning Methods in Geoscience*. Society of Exploration Geophysicists. https://seg.org/shop/product/?id=fe5a3cd3-77b2-ef11-b8e8-6045bda82e05

Segoni, S., Rosi, A., Lagomarsino, D., Fanti, R., & Casagli, N. (2016). The impact of rainfall time series with different length in a landslide warning system, in the framework of changing precipitation trends. *Geoenvironmental Disasters, 3*, Article 1. https://doi.org/10.1186/s40677-016-0057-6

Segoni, S., Piciullo, L., & Gariano, S. L. (2018). A review of the recent literature on rainfall thresholds for landslide occurrence. *Landslides, 15*(8), 1483–1501. https://doi.org/10.1007/s10346-018-0966-4

Shao, X., & Xu, C. (2022). Earthquake-induced landslides susceptibility assessment: A review of the state-of-the-art. *Natural Hazards Research, 2*(3), 172–182. https://doi.org/10.1016/j.nhres.2022.03.002

Shapley, L. S. (1953). A value for n-person games. In H. W. Kuhn & A. W. Tucker (Eds.), *Contributions to the Theory of Games* (Vol. 2, pp. 307–317). Princeton University Press. https://doi.org/10.1515/9781400881970-018

Smilkov, D., Thorat, N., Kim, B., Viégas, F., & Wattenberg, M. (2017). SmoothGrad: Removing noise by adding noise. *arXiv preprint arXiv:1706.03825*. https://arxiv.org/abs/1706.03825

Steger, S., Moreno, M., Crespi, A., Zellner, P. J., Gariano, S. L., Brunetti, M. T., Melillo, M., Peruccacci, S., Marra, F., Kohrs, R., Goetz, J., Mair, V., & Pittore, M. (2023). Deciphering seasonal effects of triggering and preparatory precipitation for improved shallow landslide prediction using generalized additive mixed models. *Natural Hazards and Earth System Sciences, 23*, 1483–1506. https://doi.org/10.5194/nhess-23-1483-2023

Sundararajan, M., Taly, A., & Yan, Q. (2017). Axiomatic attribution for deep networks. In D. Precup & Y. W. Teh (Eds.), *Proceedings of the 34th International Conference on Machine Learning* (Vol. 70, pp. 3319–3328). PMLR. https://proceedings.mlr.press/v70/sundararajan17a.html

Taha, A. A., & Hanbury, A. (2015). Metrics for evaluating 3D medical image segmentation: Analysis, selection, and tool. *BMC Medical Imaging, 15*, Article 29. https://doi.org/10.1186/s12880-015-0068-x

Tanyaş, H., & Lombardo, L. (2019). Variation in landslide-affected area under the control of ground motion and topography. *Engineering Geology, 260*, 105229. https://doi.org/10.1016/j.enggeo.2019.105229

Tao, J., & Barros, A. P. (2014). Coupled prediction of flood response and debris flow initiation during warm- and cold-season events in the Southern Appalachians, USA. *Hydrology and Earth System Sciences, 18*(1), 367–388. https://doi.org/10.5194/hess-18-367-2014

Trigila, A., Iadanza, C., & Spizzichino, D. (2010). Quality assessment of the Italian Landslide Inventory using GIS processing. *Landslides, 7*, 455–470. https://doi.org/10.1007/s10346-010-0213-0

Van Rijsbergen, C. J. (1979). Information retrieval: Theory and practice. In *Proceedings of the Joint IBM/University of Newcastle upon Tyne Seminar on Data Base Systems* (pp. 1–14).

Van Westen, C. J., Castellanos, E., & Kuriakose, S. L. (2008). Spatial data for landslide susceptibility, hazard, and vulnerability assessment: An overview. *Engineering Geology, 102*(3–4), 112–131. https://doi.org/10.1016/j.enggeo.2008.03.010

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention is all you need. In *Advances in Neural Information Processing Systems* (Vol. 30).

Wang, N., Cheng, W., Marconcini, M., Bachofer, F., Liu, C., Xiong, J., & Lombardo, L. (2022). Space-time susceptibility modeling of hydro-morphological processes at the Chinese national scale. *Engineering Geology, 301*, 106586. https://doi.org/10.1016/j.enggeo.2022.106586

Wasowski, J., Keefer, D. K., & Lee, C.-T. (2011). Toward the next generation of research on earthquake-induced landslides: Current issues and future challenges. *Engineering Geology, 122*(1–2), 1–8. https://doi.org/10.1016/j.enggeo.2011.06.001

Wu, C.-H., & Chen, S.-C. (2009). Determining landslide susceptibility in Central Taiwan from rainfall and six site factors using the analytical hierarchy process method. *Geomorphology, 112*(3–4), 190–204. https://doi.org/10.1016/j.geomorph.2009.06.002

Xu, K., Zhao, Z., Chen, W., Ma, J., Liu, F., Zhang, Y., & Ren, Z. (2024). Comparative study on landslide susceptibility mapping based on different ratios of training samples and testing samples by using RF and FR-RF models. *Natural Hazards Research*, *4*(1), 62-74. https://doi.org/10.1016/j.nhres.2023.07.004

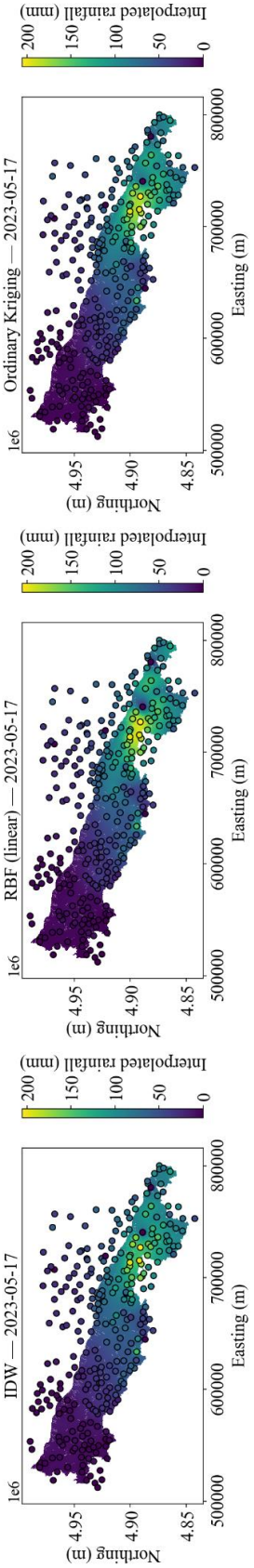
Youssef, K., Shao, K., Moon, S., & Bouchard, L. S. (2022). Landslide susceptibility modeling by interpretable neural network. *arXiv preprint arXiv:2201.06837*. https://arxiv.org/abs/2201.06837

Zêzere, J. L., Pereira, S., Melo, R., Oliveira, S. C., & Garcia, R. A. (2017). Mapping landslide susceptibility using data-driven methods. *Science of the Total Environment, 589*, 250–267. https://doi.org/10.1016/j.scitotenv.2017.02.188

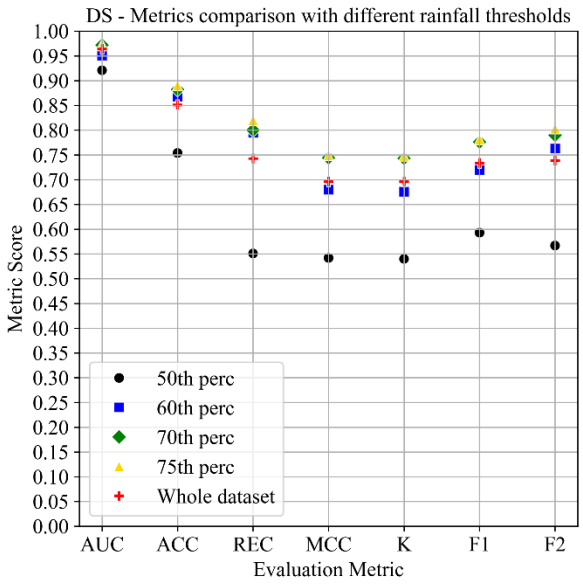
Zhou, H., Zhang, S., Peng, J., Zhang, S., Li, J., Xiong, H., & Zhang, W. (2021). Informer: Beyond efficient transformer for long sequence time-series forecasting. In *Proceedings of the AAAI conference on artificial intelligence* (Vol. 35, No. 12, pp. 11106-11115). https://doi.org/10.1609/aaai.v35i12.17325

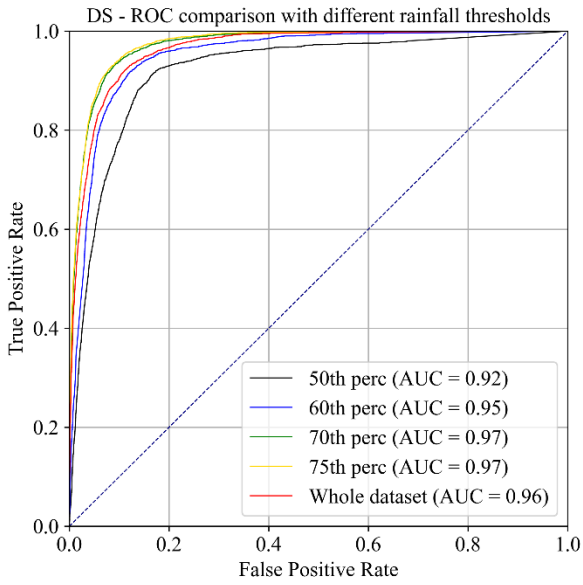
Zou, K. H., Wells III, W. M., Kikinis, R., & Warfield, S. K. (2004). Three validation metrics for automated probabilistic image segmentation of brain tumours. *Statistics in Medicine, 23*(8), 1259–1282. https://doi.org/10.1002/sim.1723

4.10 Supplementary material chapter 4

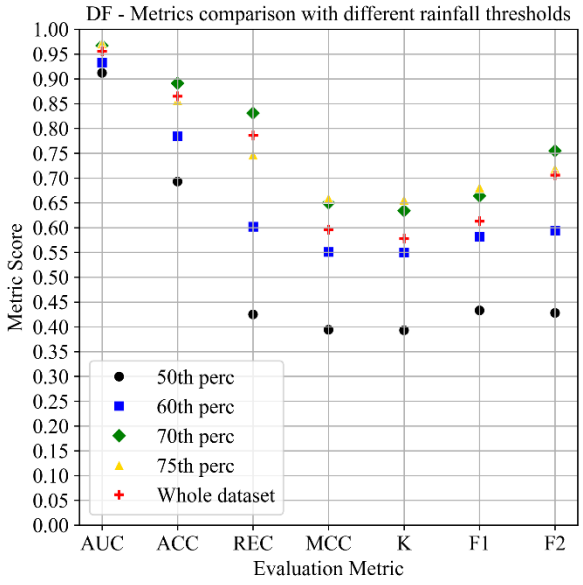


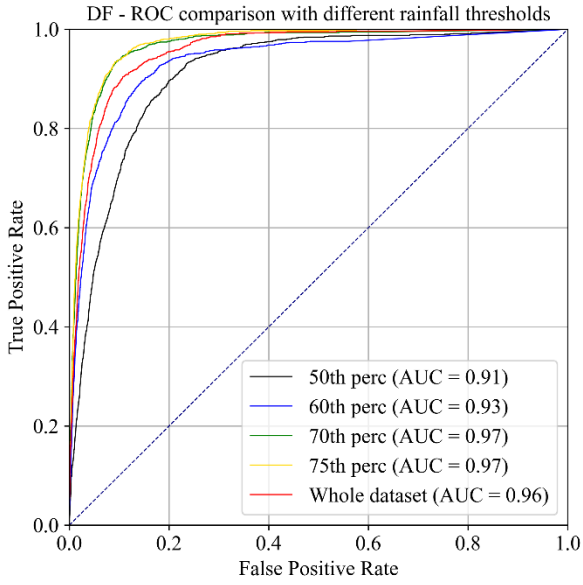
S 4.1. Rainfall Interpolations comparison between IDW, RBF (linear), and Ordinary kriging on the day 17-May 2023, the coloured dots represent the original pluviometry stations.



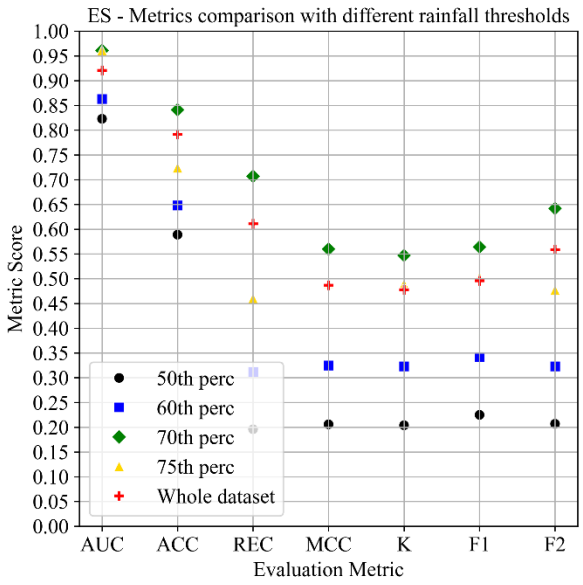


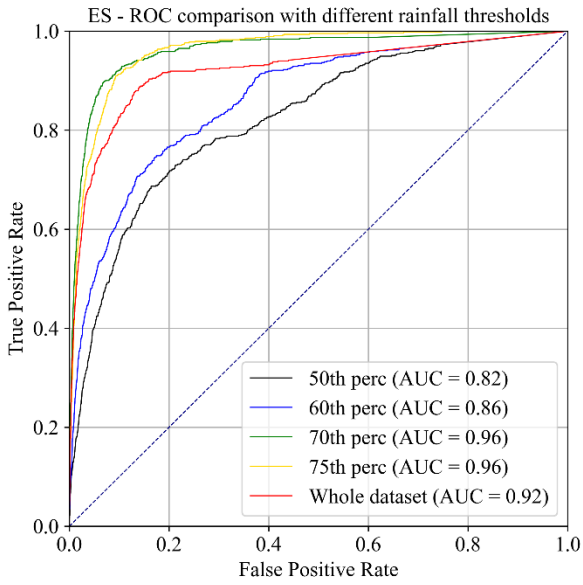
S 4.2. Debrs slide performing metrics on test dataset at prediction stage (inference on) for different SUs filtering by rainfall thresholds and the whole dataset.



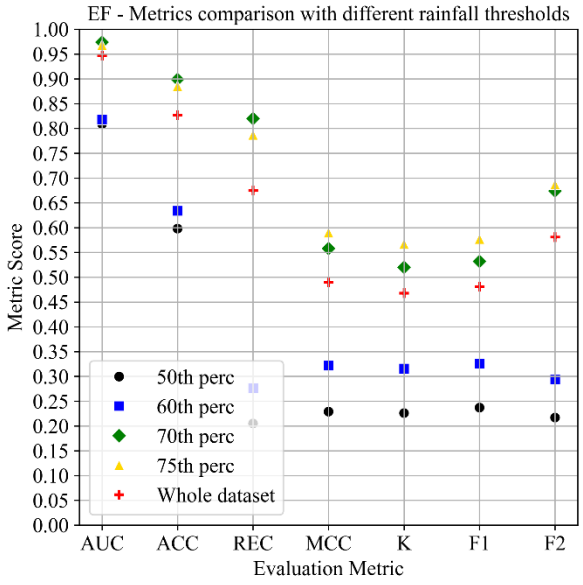


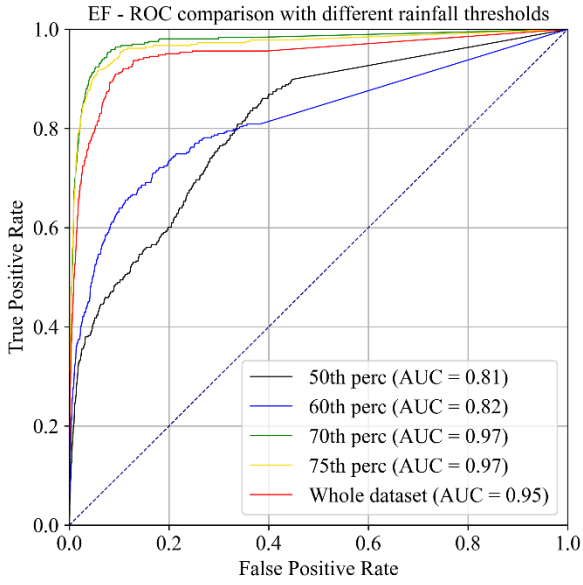
S 4.3. Debris flow performing metrics on test dataset at prediction stage (inference on) for different SUs filtering by rainfall thresholds and the whole dataset.



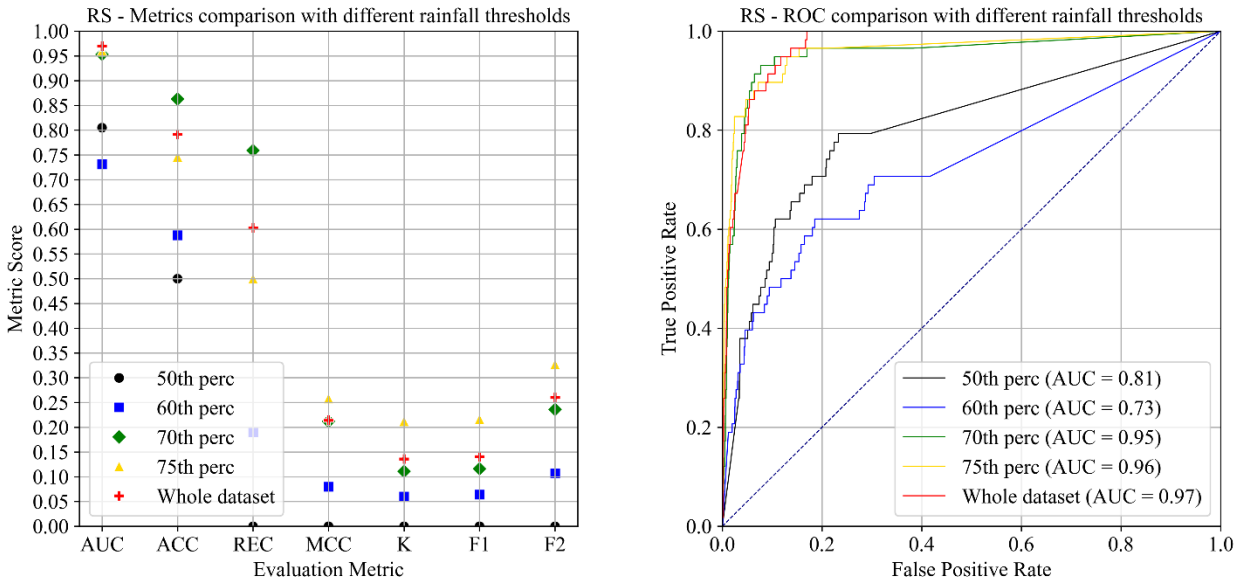


S 4.4. Earth slide performing metrics on test dataset at prediction stage (inference on) for different SUs filtering by rainfall thresholds and the whole dataset.

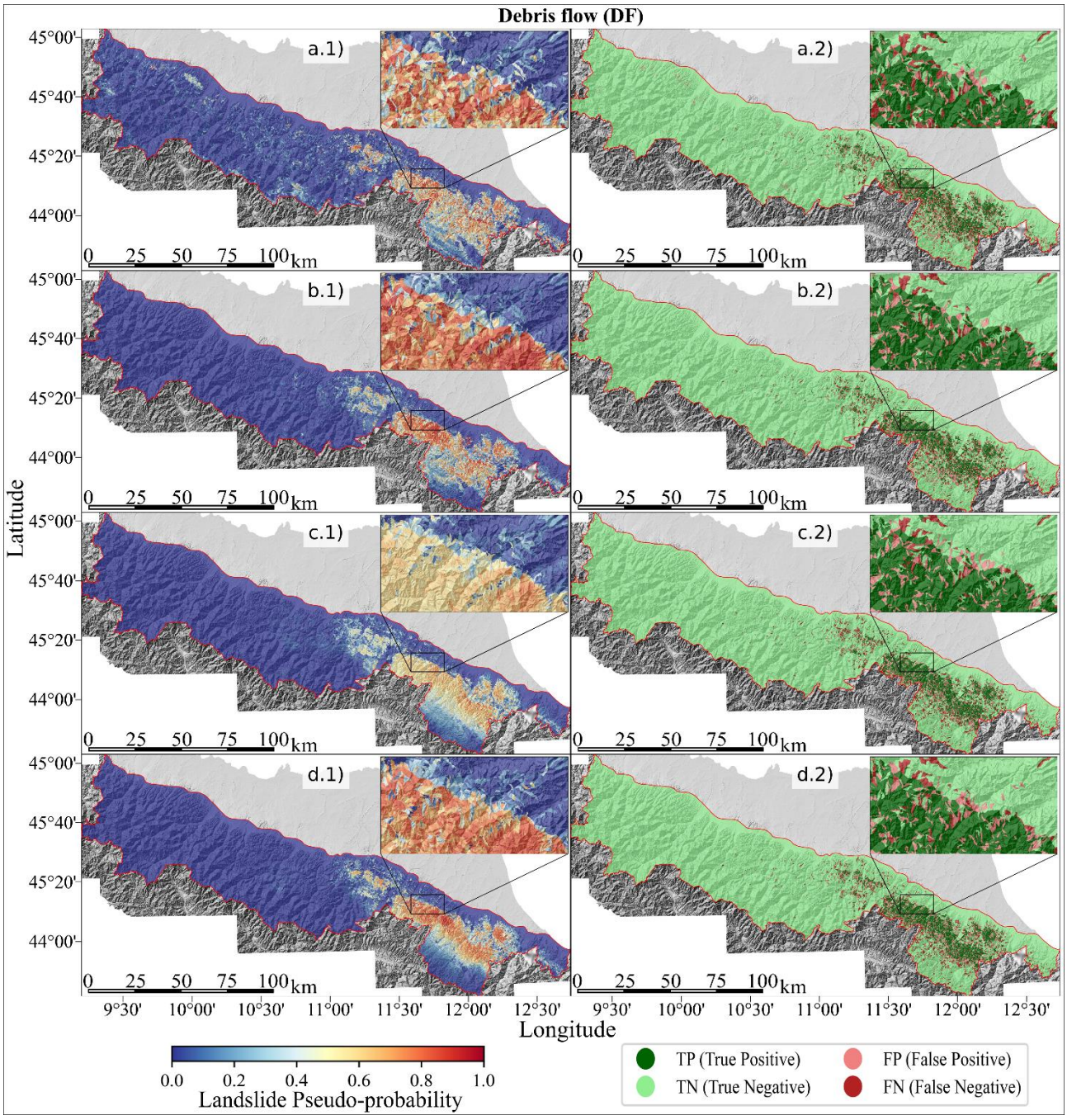




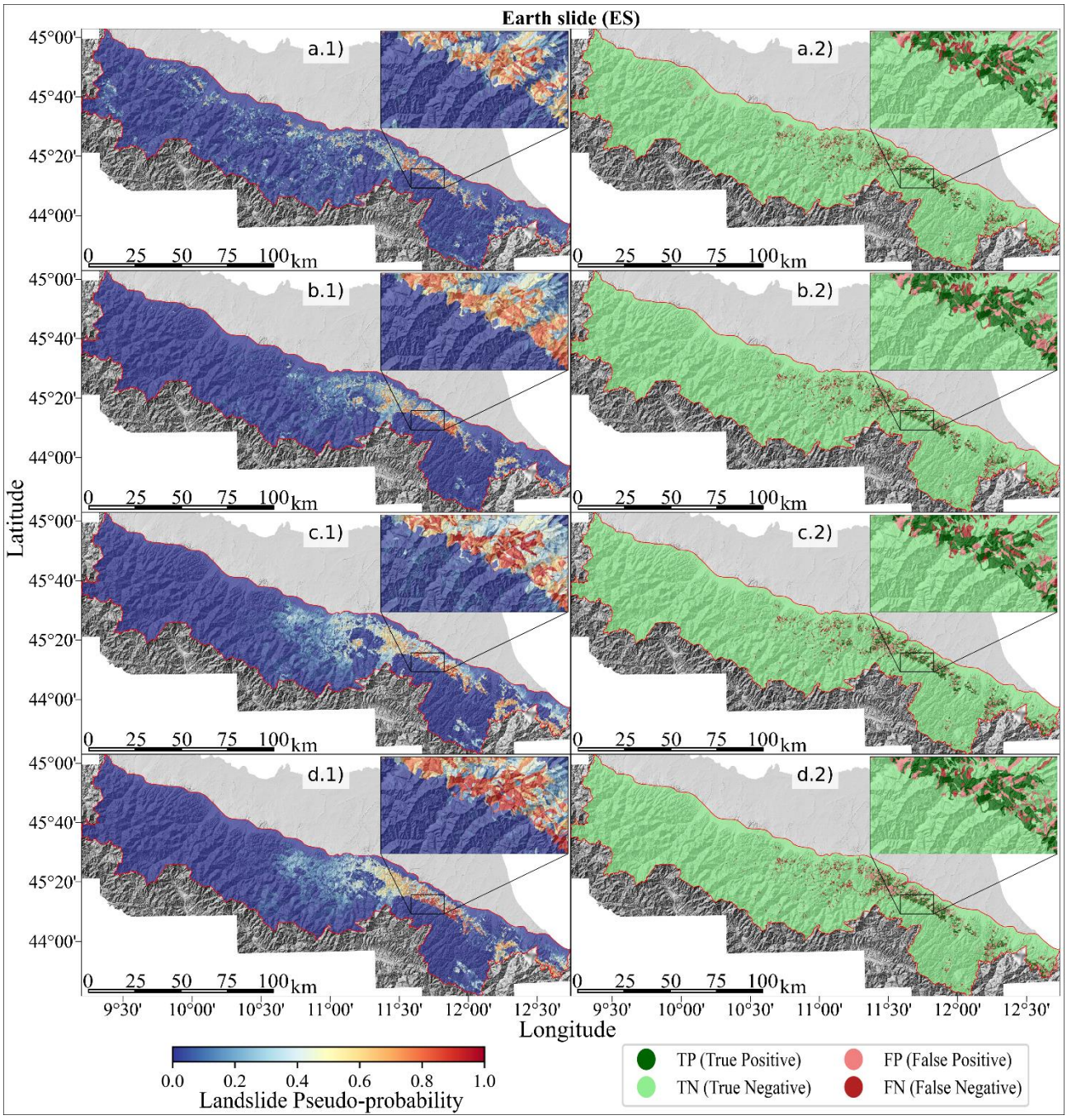
S 4.5. Earth flow performing metrics on test dataset at prediction stage (inference on) for different SUs filtering by rainfall thresholds and the whole dataset.



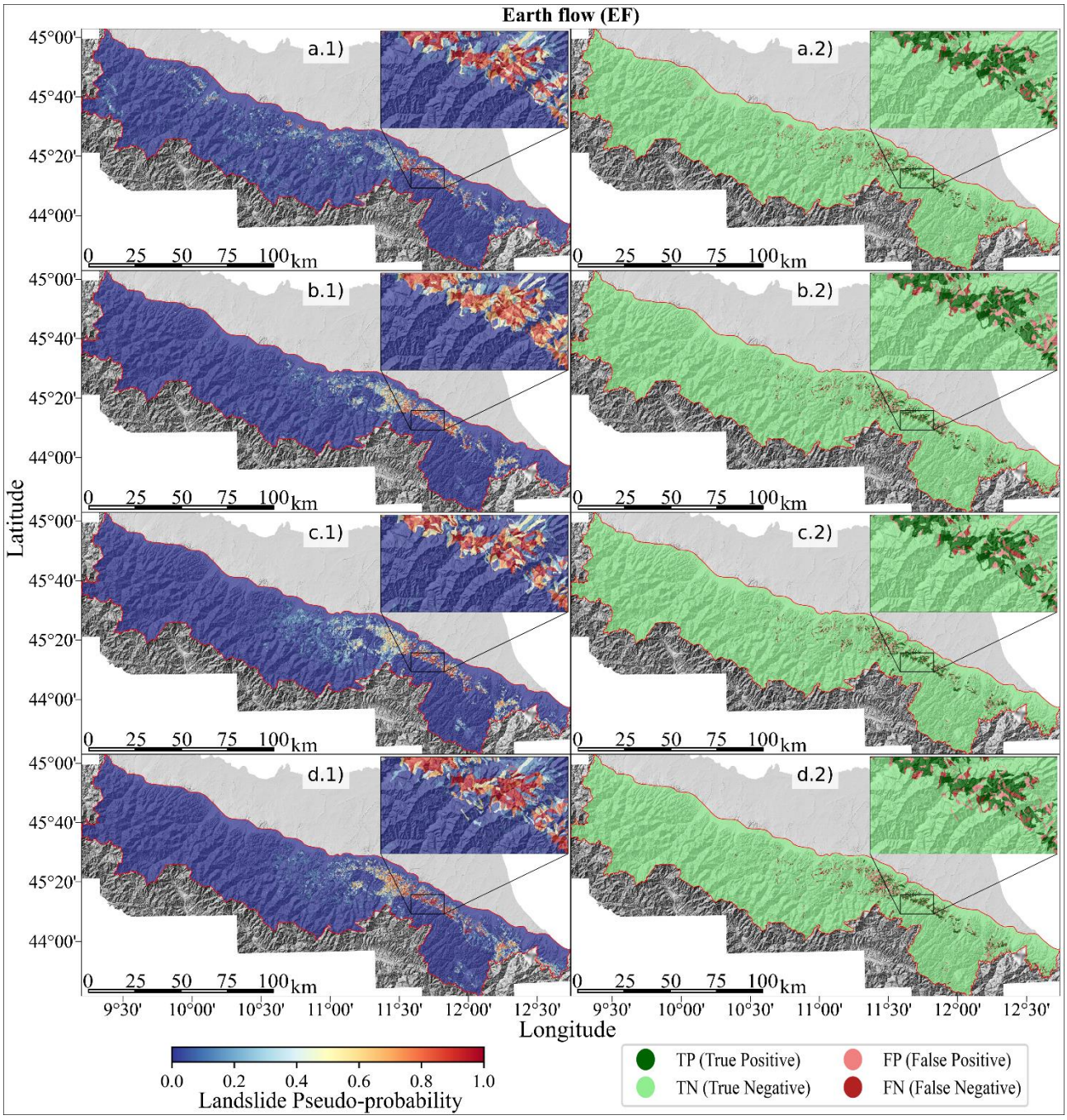
S 4.6. Rock slide performing metrics on test dataset at prediction stage (inference on) for different SUs filtering by rainfall thresholds and the whole dataset.



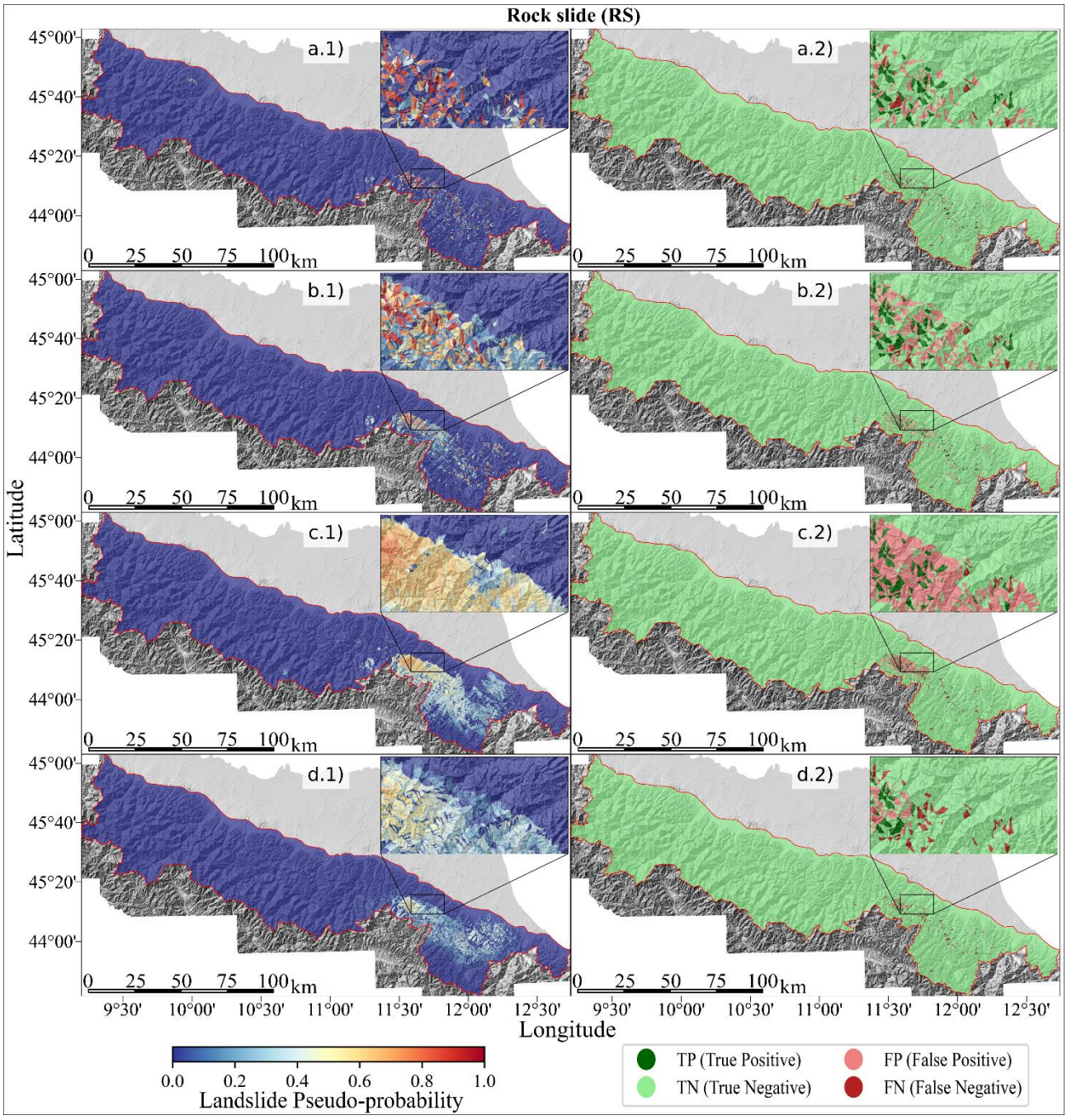
S 4.7. Spatial pseudo-probability (.1) and Confusion Matrix (*.2) Debris flow (DF) for the May 2023 Emilia Romagna events. a) Shown the "Only terrain" configuration. b) Shown the "All cumulative scalar rainfall" configurations. c) Shown the "Time series with daily intervals rainfall" configuration. d) Shown the "Time series with hourly intervals rainfall" configuration.*



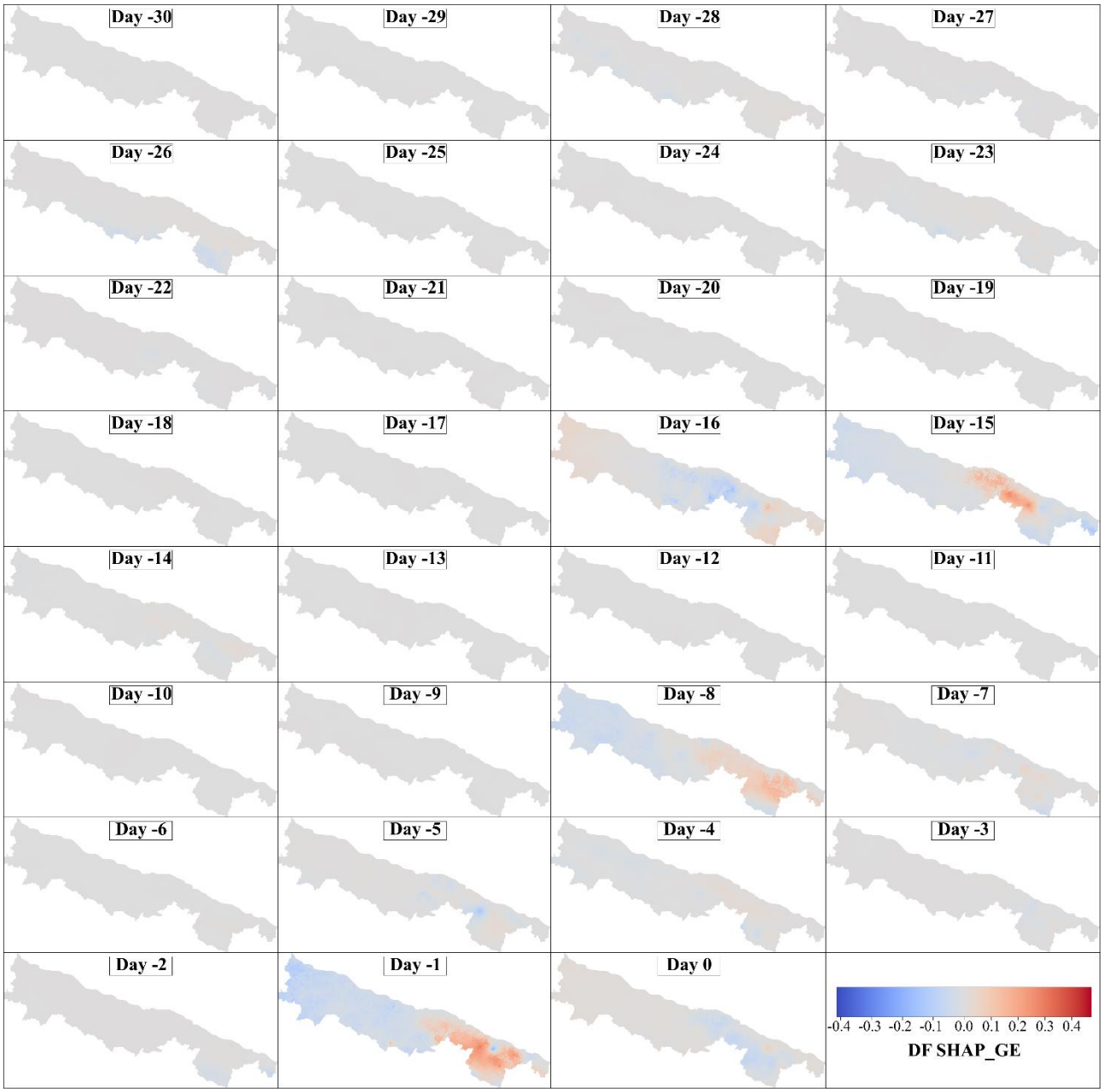
S 4.8. Spatial pseudo-probability (.1) and Confusion Matrix (*.2) Earth slide (ES) for the May 2023 Emilia Romagna events. a) Shown the "Only terrain" configuration. b) Shown the "All cumulative scalar rainfall" configurations. c) Shown the "Time series with daily intervals rainfall" configuration. d) Shown the "Time series with hourly intervals rainfall" configuration.*



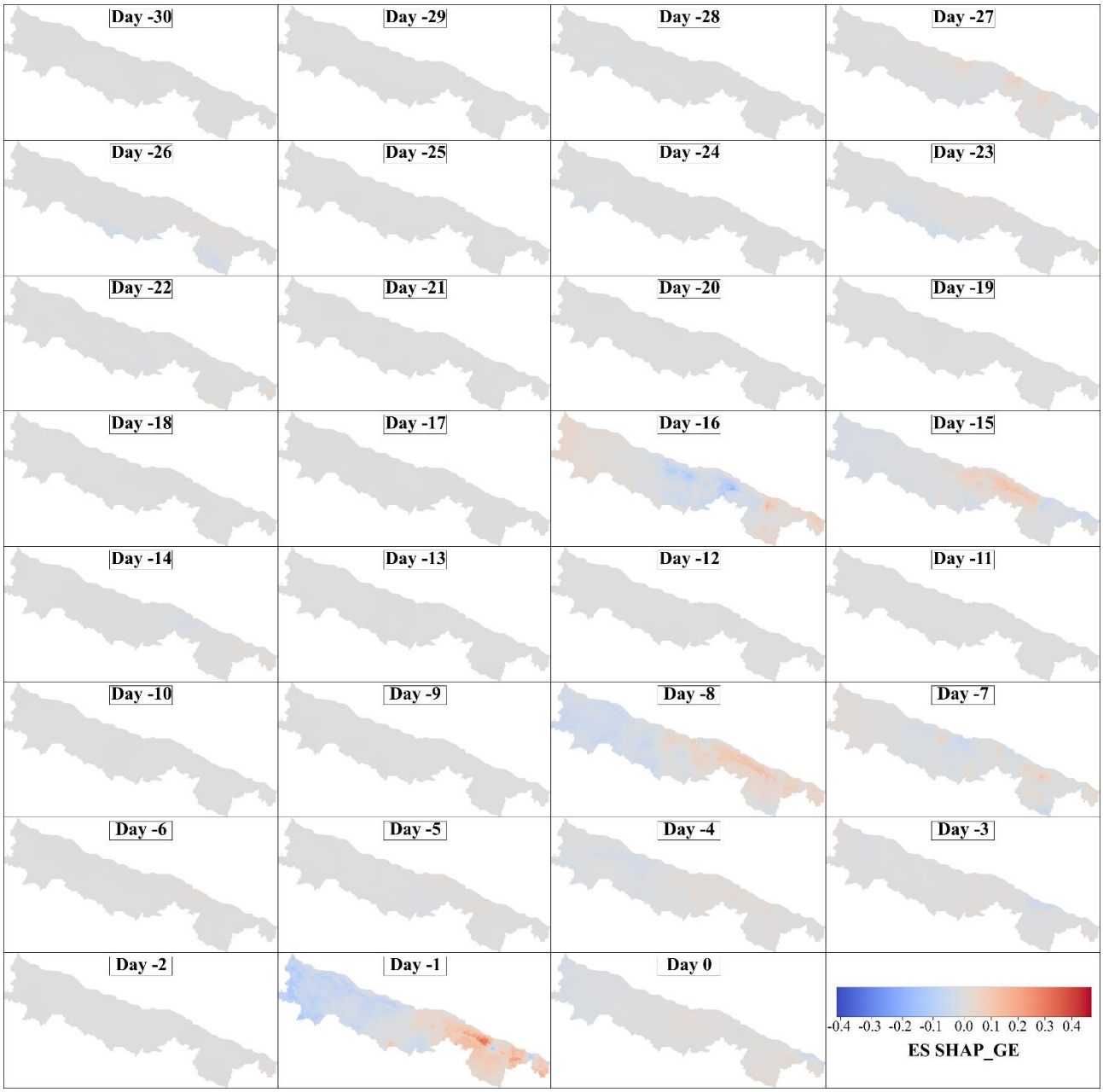
S 4.9. Spatial pseudo-probability (.1) and Confusion Matrix (*.2) Earth flow (EF) for the May 2023 Emilia Romagna events. a) Shown the "Only terrain" configuration. b) Shown the "All cumulative scalar rainfall" configurations. c) Shown the "Time series with daily intervals rainfall" configuration. d) Shown the "Time series with hourly intervals rainfall" configuration.*



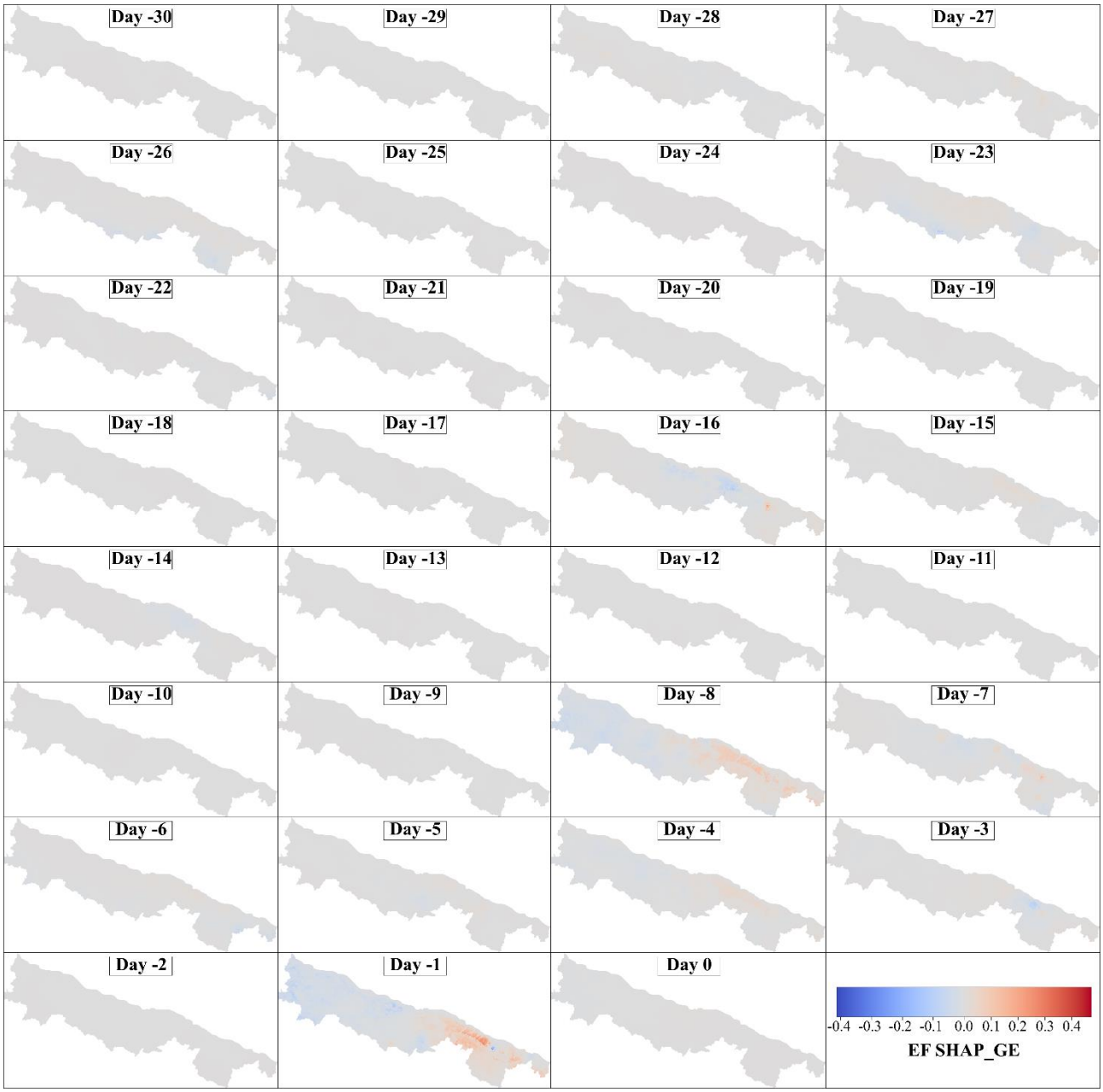
S 4.10. Spatial pseudo-probability (.1) and Confusion Matrix (*.2) Rock slide (RS) for the May 2023 Emilia Romagna events. a) Shown the "Only terrain" configuration. b) Shown the "All cumulative scalar rainfall" configurations. c) Shown the "Time series with daily intervals rainfall" configuration. d) Shown the "Time series with hourly intervals rainfall" configuration.*



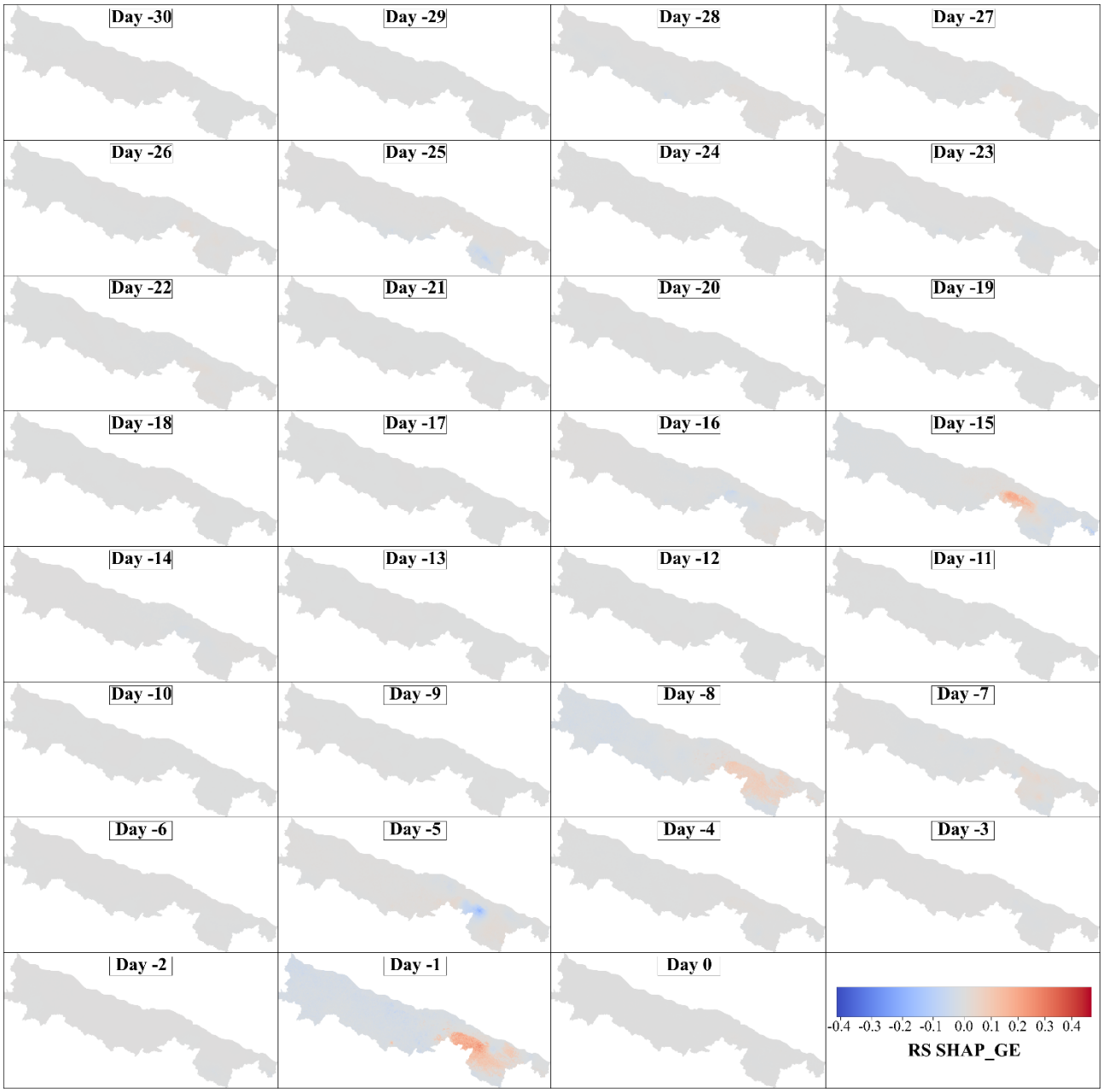
S 4.11. Spatial visualization of the SHAP_GE scores per day for the DF landslide type. Color bar shared among all landslide types.



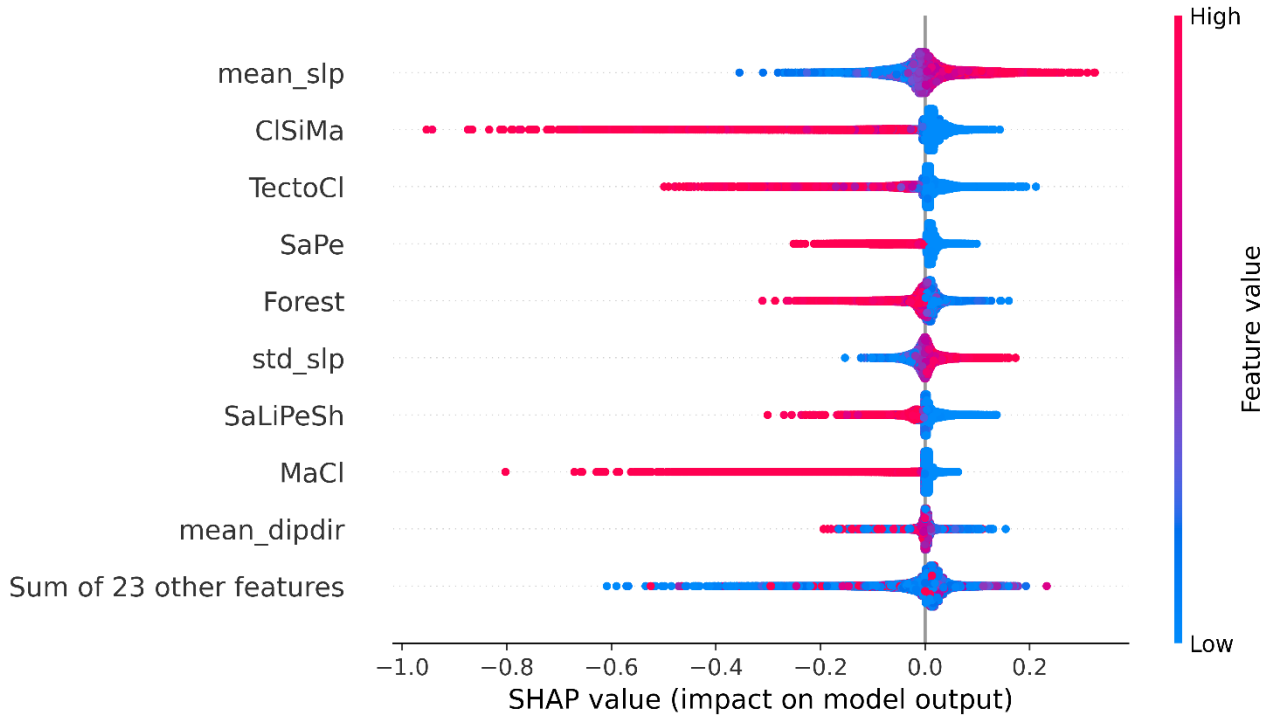
S 4.12. Spatial visualization of the SHAP_GE scores per day for the ES landslide type. Color bar shared among all landslide types.



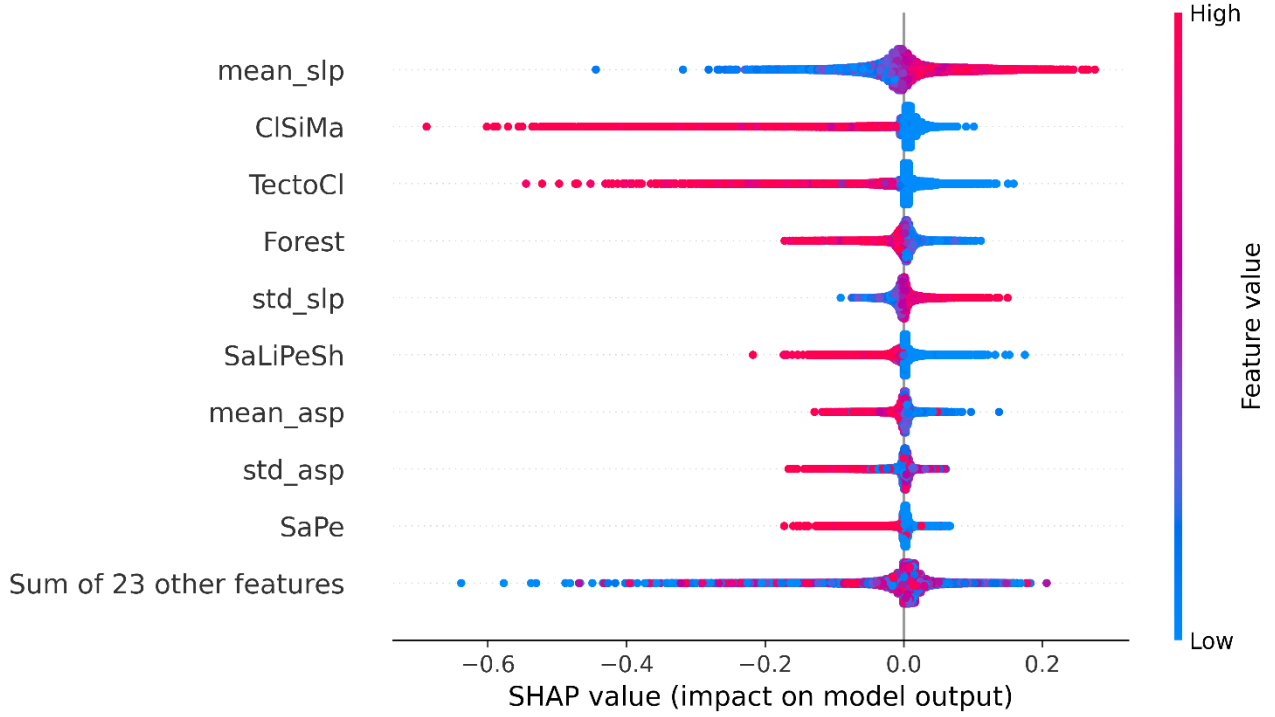
S 4.13. Spatial visualization of the SHAP_GE scores per day for the EF landslide type. Color bar shared among all landslide types.



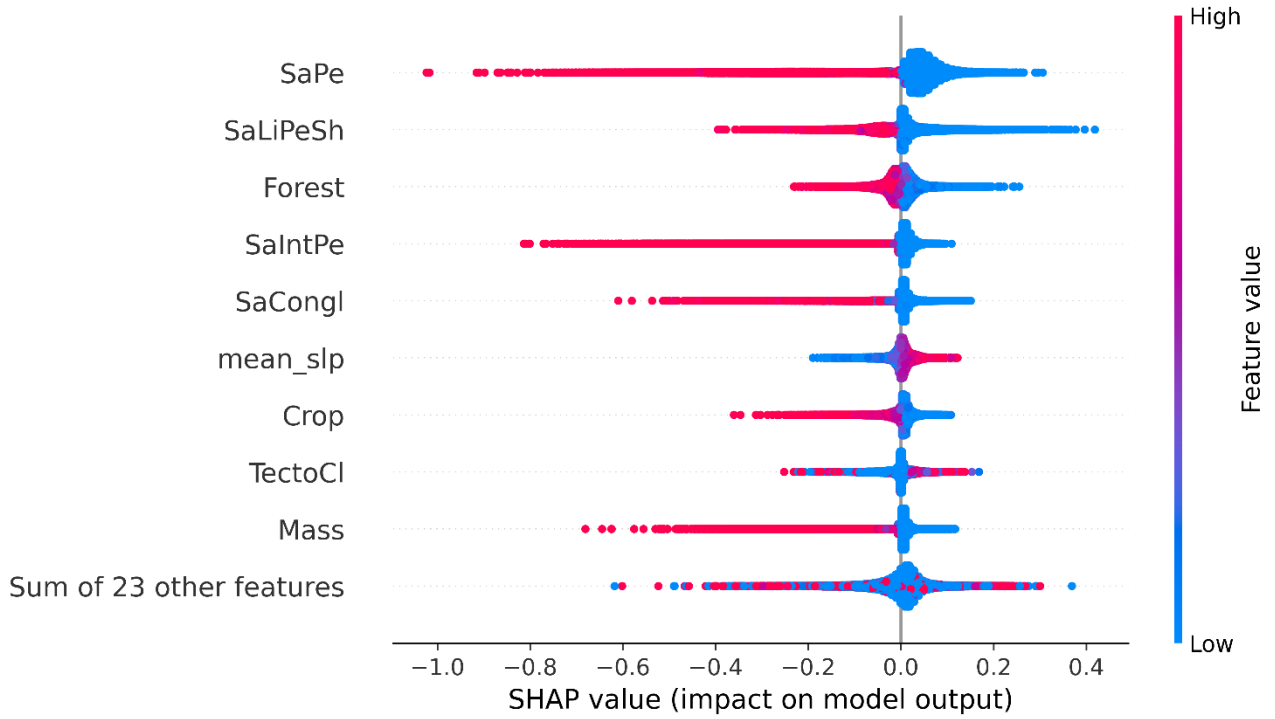
S 4.14. Spatial visualization of the SHAP_GE scores per day for the RS landslide type. Color bar shared among all landslide types.



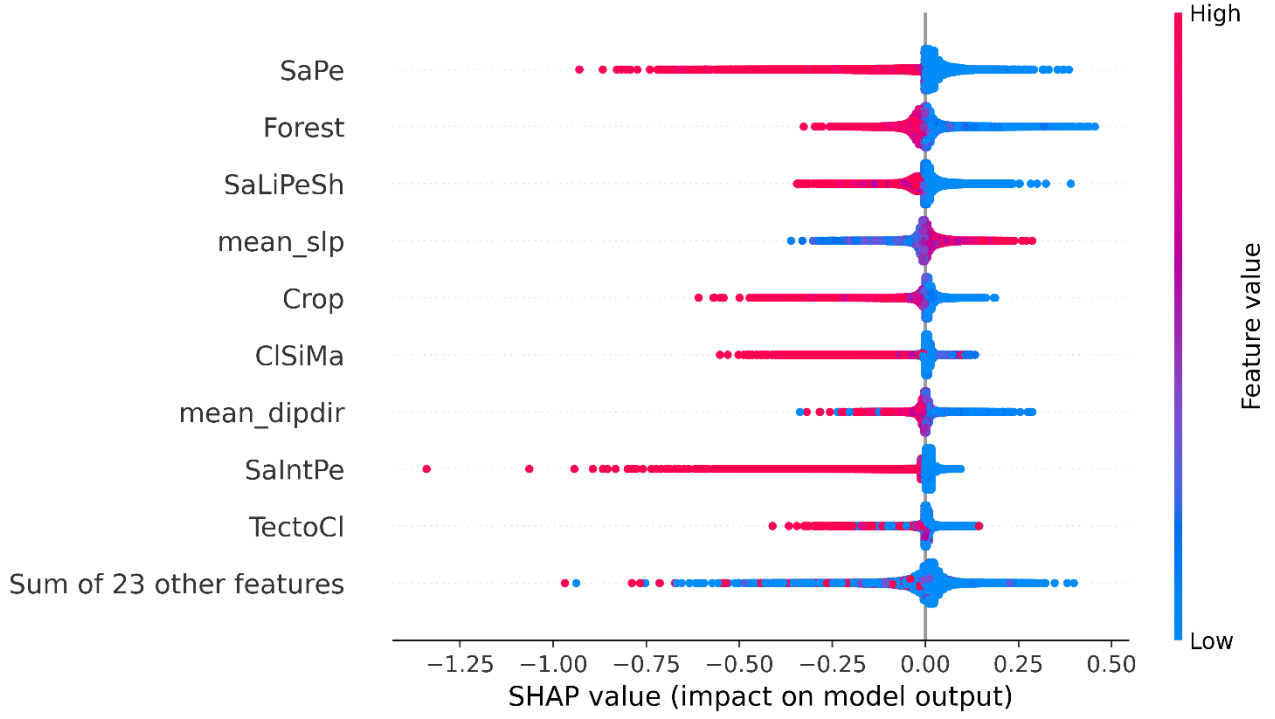
S 4.15. Beeswarm plot of SHAP_GE scores for DS landslide types of static factors (the nine most influential ones and the sum of the remaining).



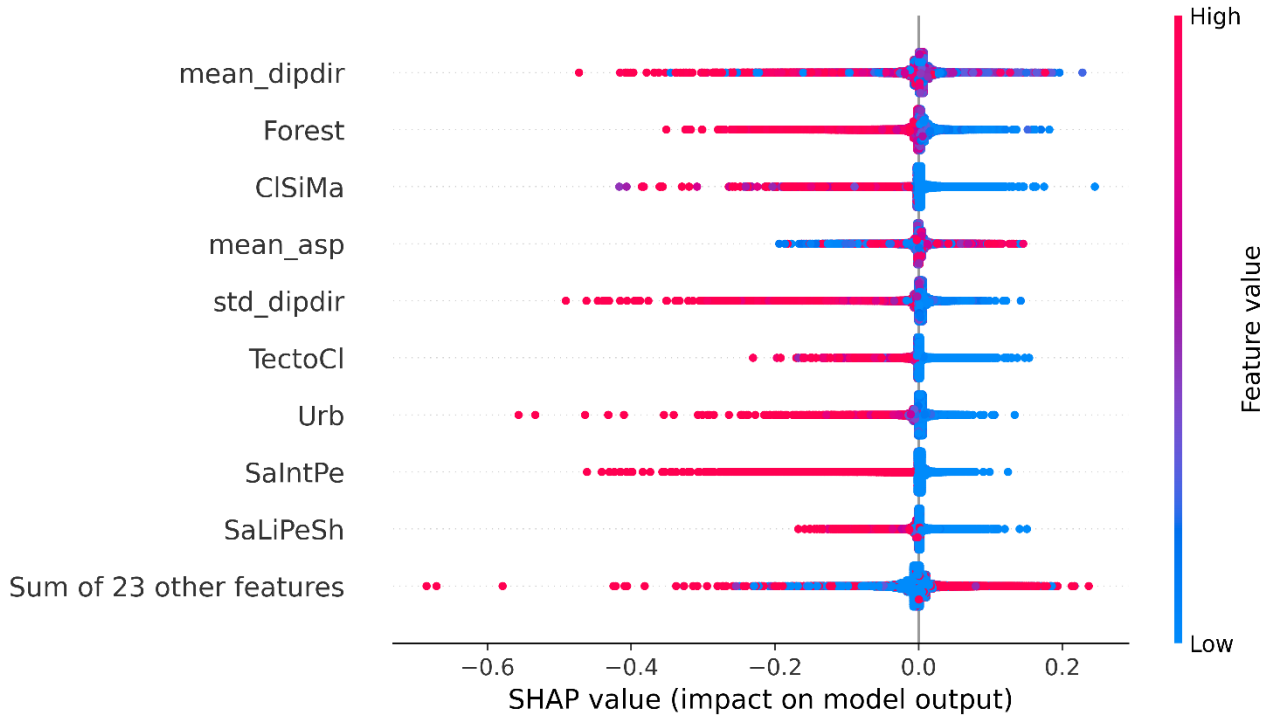
S 4.16. Beeswarm plot of SHAP_GE scores for DF landslide types of static factors (the nine most influential ones and the sum of the remaining).



S 4.17. Beeswarm plot of SHAP_GE scores for ES landslide types of static factors (the nine most influential ones and the sum of the remaining).



S 4.18. Beeswarm plot of SHAP_GE scores for EF landslide types of static factors (the nine most influential ones and the sum of the remaining).



S 4.19. Beeswarm plot of SHAP_GE scores for RS landslide types of static factors (the nine most influential ones and the sum of the remaining).

5. Chapter 5: Risk-based prioritisation for the installation of Debris Flow early warning system in drainage catchments

5.1 Preface

The chapter 5 is developed under the project "Debris Flow Risk Mapping and Early Warning" with Coalition for Disaster Resilient Infrastructure (CDRI). This project introduces a groundbreaking method to map debris flow risks and provide timely early warnings. By precisely identifying high-risk zones and refining alert systems, this approach enhances safety, bolsters infrastructure resilience, and minimizes damage.

My personal contribution in the project is presented in this chapter and consist of evaluating the susceptibility, the likely debris flow magnitude, and the impact on the built environment of 17 mountain catchments in Val Camonica (Lombardia, Italy). The assessment considers morphometric parameters, existing approaches to transported volume quantification, and publicly available data on exposure and vulnerability.

The proposed risk assessment methodology is designed to be easily replicable in other regions, relying solely on a Digital Elevation Model (DEM) to map alluvial fans and automatically delineate the corresponding catchments, combined with satellite RGB imagery to derive land cover information. This ensures that the framework can be applied even in areas with limited historical landslide data, particularly in developing countries.

The chapter also highlights the use of physical-empirical techniques for debris flow susceptibility estimation, maintaining consistency with the necessary landslide type-specific analysis emphasized throughout this thesis. This straightforward and adaptable framework demonstrates how reliable risk assessment can be achieved with minimal data, supporting proactive disaster management and early-warning strategies.

The chapter is based on a paper published on *9th International Symposium for Geotechnical Safety and Risk (ISGSR) proceedings*:

Redaelli M.[1], Rani R.[2], Ioriatti E.[2] & Berti M.[2], 2025. Risk-based prioritisation for the installation of debris flow early warning systems in drainage catchments.

[1] Studio Geotecnico Strutturale – SGS s.r.l., Italy.

[2] Department of Biological, Geological, and Environmental Sciences, University of Bologna, Bologna, Italy.

5.2 Abstract

This work shows how desk-based analysis combined with cost-effective targeted surveys can rapidly identify debris-flow hazard and risk hotspots in Central Val Camonica (Italian Alps). Hazard is assessed by combining susceptibility and magnitude where event-frequency data are limited. Susceptibility is estimated using a logistic-regression model calibrated on a global dataset of basins and driven by two GIS-derived morphometric predictors: the Melton Ratio and mean alluvial-fan slope. Magnitude is estimated with an empirical volume model coupling basin relief with an active-area index representing sediment supply, derived from land-cover classes via a weighted stability number. Results are synthesised in contingency matrices using percentile thresholds (33rd and 66th) to produce detailed and simplified hazard classes. Risk is evaluated by intersecting alluvial fans with exposed elements (buildings and transport infrastructure) from OpenStreetMap and a national topographic database, harmonised through normalisation and weighting into a single exposure score. The method highlights priority basins and critical fans along infrastructure corridors to support trial deployment of a novel monitoring and early warning system (MEWS).

5.3 Introduction

Effective geological risks mitigation relies on prioritizing limited financial resources for the highest-risk locations. This study, conducted by Regione Lombardia and Bologna University, identifies the most suitable sites for an experimental monitoring and early warning system (MEWS) in a densely populated mountainous area. Val Camonica, in the central Italian Alps, spans 1’300 km² with approximately 120’000 residents. The study focuses on its central part (423 km²), as shown in Figure 5.1.

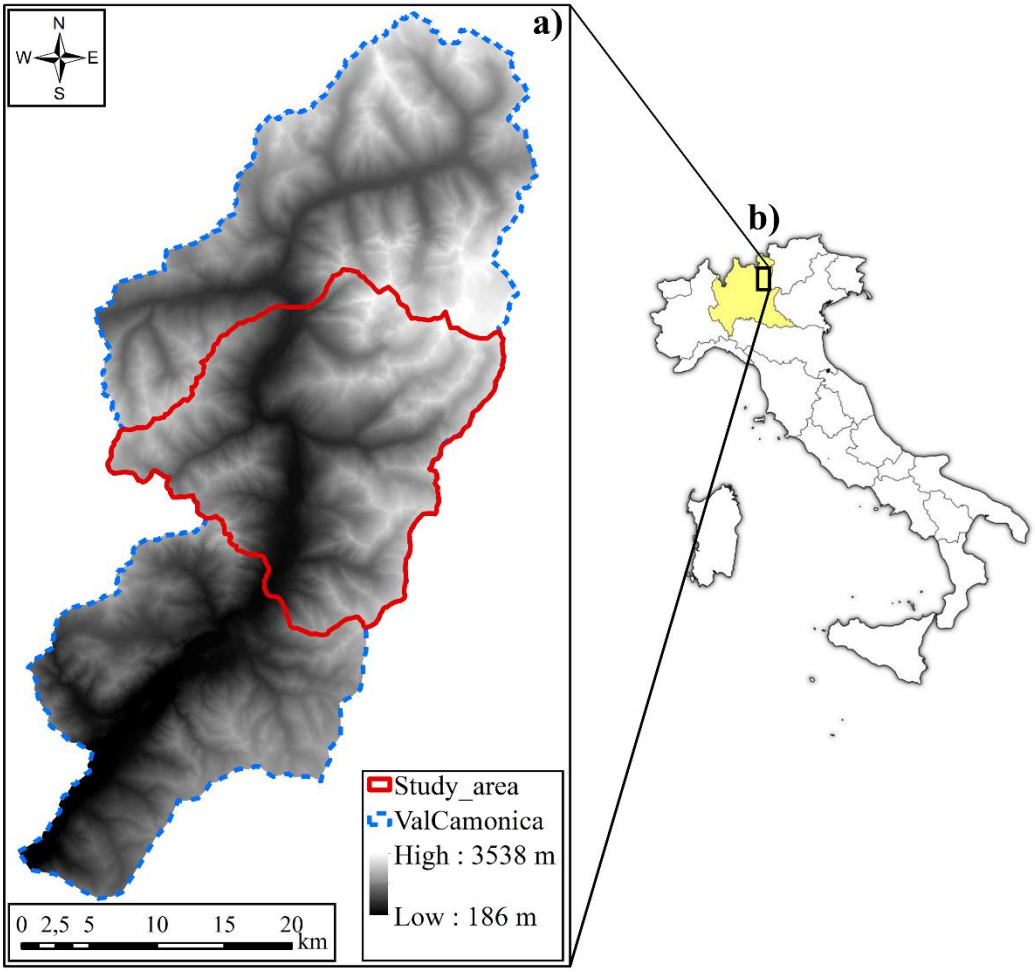


Figure 5.1. a) The selected study area: Central Valle Camonica, in the Italian Alps b) Location of the Val Camonica in Italy with the Lombardia region highlighted in light yellow.

The need to rank the drainage catchments according to their debris flow risk was prompted by a research project on a new MEWS (Ioriatti et al. 2023). The novel system combines the rainfall threshold method with the analysis of ground vibrations and infrasound from the process initiation (Schimmel & Berti 2023). The proposed risk-based prioritization can inform decisions about where new MEWS stations should be installed.

5.4 Hazard estimation

In our research, we have undertaken a comprehensive approach to debris flow hazard assessment, integrating factors of susceptibility and magnitude to delineate the potential impact of such events within targeted watersheds. This methodology is rooted in the understanding that a thorough hazard assessment encapsulates not only the physical characteristics and potential intensity of natural events but also the frequency at which these events may occur. Hazard, within the context of natural disaster science, is commonly conceptualized as a multidimensional construct, encompassing the event's magnitude, frequency, and susceptibility of the area affected (White & Haas, 1975).

However, an ideal hazard assessment, incorporating all three dimensions, often faces significant challenges. Foremost among these is the accurate estimation of event frequency, particularly

for phenomena like debris flows, which may exhibit sporadic occurrence patterns and be influenced by a complex interplay of triggers and predisposing factors (Iverson, 1997). Consequently, in scenarios where frequency data are scarce or inherently uncertain, the focus of hazard assessment may pivot towards a detailed analysis of susceptibility and magnitude. This pragmatic shift allows for the identification and prioritization of areas at risk based on the best available data, albeit with an understanding and acknowledgement of the inherent limitations posed by the incomplete triad of hazard components.

In our study area, we have information on the recent debris flow event and related infrastructure damage, but we deliberately pursued a generic approach so that it can be replicated in lesser-known areas in developing countries from a CDRI perspective. For the same reason, we want to create a simple and straightforward procedure using a pragmatic integration of what already exists in the literature.

5.4.1 Susceptibility estimation

Numerous studies in the literature have explored morphometric indicators as a means to assess the dominant flow type on alluvial fans, categorizing them into two main types: those dominated by sediment gravity flows, known as debris-flow fans, and those dominated by fluid gravity flows, referred to as fluvial fans (Jackson et al. 1987, Marchi et al. 1993; Calvache et al. 1997; Marchi and Brochot 2000; De Scally and Owens 2004; Santos Alonso 2011; Melton, et al. 2021; amongst others). These studies provide a probabilistic prediction of the dominant sediment transport process at a catchment's outlet, derived from morphometric properties of both the fans and catchments, albeit based on several assumptions. Alluvial fans are sedimentary formations created through prolonged and complex cycles of aggradation and degradation, influenced by the evolving environmental factors that govern the sediment and water dynamics within their catchments. As noted by Harvey (1984) and Blair and McPherson (1994), these fans have undergone shifts in the dominant sediment transport processes over time, often in response to changes in land use. Consequently, current surface features or the stratigraphic record may exhibit signs of processes that are no longer active. For our purposes, we consider multi-decadal to secular timescales as appropriate for identifying ongoing processes on the fans. This perspective also enables us to examine basins under various past conditions, which may mirror future scenarios in the face of climate change.

In our study, we employed the equation proposed by Rowbotham et al. (2005) to quantify the likelihood of a catchment producing debris flows, alongside the direct classification method from Wilford et al. (2004) for determining the potential of basins to generate such flows based on their morphometrics. However, when applied to our study area, both methods yielded unrealistic results, especially in light of known recent debris flow events. This discrepancy is largely attributed to the original calibration and training of these models on a limited dataset of basins under different geomorphological, geological, and climatic conditions than those found in our study area. This challenge was overcome by adopting the following equation from Bertrand et al. (2013), which is based on logistic regression analysis of 620 basins worldwide, covering a range of geo-climatic conditions. This equation (5.1), implemented in Matlab, leverages the Melton Ratio index (Mr) from Melton (1965) and the mean slope of the alluvial

fan (S), highlighting these as the most indicative and widely utilized parameters for differentiating predominant transport processes.

$$p = \frac{e^{-0.65} * Mr^{1.66} * S^{2.00}}{1 - e^{-0.65} * Mr^{1.66} * S^{2.00}} \quad [5.1]$$

5.4.2 Magnitude estimation

The magnitude of a debris flow, often quantified by the volume of material released during an event, is a critical factor in understanding the potential impact of such phenomena (D'Agostino & Marchi, 2001). Various methods, including empirical and statistical models, have been developed to estimate debris flow volume, each with its advantages and limitations (Bovis & Jacob, 1999; Franzi & Bianco, 2001; D'Agostino & Marchi, 2001; Dong et al., 2009; Gartner et al., 2014; De Haas & Densmore, 2019). Assessing magnitude is fundamental for evaluating debris-flow hazards as it provides insights into the event's severity, influences the selection of mitigation strategies (e.g., designing torrent control measures or relocating settlements), and determines the required capacity of debris retention basins. However, these methods are not without their shortcomings. Empirical formulas often suffer from low accuracy and are usually applicable only within the regions for which they were developed. Geomorphological estimations based on field observations can introduce subjectivity. Moreover, the availability of sediment supply in the catchment is a crucial parameter affecting debris flow magnitude, as highlighted by Iverson (1997), Bovis & Jacob (1999), and Dong et al. (2009).

In this study, we applied the Bovis and Jacob method (Eq. 2), calibrated in British Columbia, Canada, which incorporates land cover data and potential sediment supply, offering a more comprehensive approach despite the geological and meteorological differences from our study area.

$$M = 800 * Z_T^{2.38} * A_I^{1.35} \quad [5.2]$$

Where Z_T is the total basin relief and A_I is the active area index defined as the product of the Area actively contributing debris (A_C) and Weighted stability number ($N_S = \sum a_j * W_j$) in which a_j is the proportion of the total active area in stability class j, and W_j is stability weight in ordinal class j (w1 = 1: < 40 per cent of terrain polygon unvegetated and actively contributing debris, w2 = 2: 40 ±70 per cent of terrain polygon unvegetated and actively contributing debris, w3 = 4: > 70 per cent of terrain polygon unvegetated and actively contributing debris).

5.4.3 Hazard estimation framework implementation

Our final purpose is to identify the most critical mountain basins for the Debris Flow, particularly in scenarios where historical event data is unavailable. To achieve this, we initiated a comprehensive data collection phase aimed at mapping and analyzing watersheds within our area of interest. This phase involved gathering various information layers, which are detailed in Table 5.1. Our study area, situated in the Val Camonica Valley in the northeastern Lombardia region, benefits significantly from a robust base of data made accessible through the Geoportal of Lombardia (https://www.geoportale.regione.lombardia.it/) and National service.

Name	Description	Source
Orthophoto	Aerial shots taken by AGEA (Agricultural Disbursement Agency) in the 2021 - WGS84/UTM32N coordinates - WMS data	Lombardia Geoportal
DEM 5x5	Digital Elevation Model with 5-meter resolution, 2015 edition - WGS84/UTM32N coordinates - Raster data	Lombardia Geoportal
ADBPO Cartography	Cartography of basins and fans prepared in 2003 by the public organization: the Po River District Basin Authority (ADBPO) - ED50/UTM32N coordinates – Vector Dataset.	ADBPO
IFFI Lombardia	The Inventory of the Phenomena Franosi (IFFI) of Lombardy - WGS84/UTM32N coordinates - Vector Dataset.	Lombardia Geoportal
DUSUF 2015	DUSAF (agricultural and forest land use destination) is a detailed geographic database compiled by Lombardia region in 2015 - WGS84/UTM32N coordinates - Vector Dataset.	Lombardia Geoportal

Table 5.1. Data used for mapping and analysing the mountain basins in Val Camonica.

Focusing on the Val Camonica Valley, we leveraged the ADBPO and IFFI information layers to identify 17 alluvial fans and their corresponding catchments (Figure 5.2), located within the valley segment of interest. This initial step, conducted as part of the CDRI project, utilized pre-mapped data to expedite our analysis. It is important to acknowledge, however, that such detailed map data may not be universally available across all global regions. Nevertheless, it is crucial to highlight that similar information can be obtained through satellite imagery, Digital Elevation Models (DEMs), and field surveys, albeit with varying degrees of detail, making our methodology suitable for developing countries with poor information about past debris flow events.

Following the identification process, we derived all required geometrical parameters (as detailed in table 5.2 and 5.3) to apply the literature-based methodology previously discussed for estimating the probability and magnitude of debris flow events in each identified basin. These parameters were accurately determined using GIS software, employing straightforward and easily replicable steps that utilize basic geometric computation tools. The catchment data were instrumental in determining event magnitude, while the alluvial fans were more indicative of the probability of debris flow occurrence, owing to their preservation of historical depositional process records.

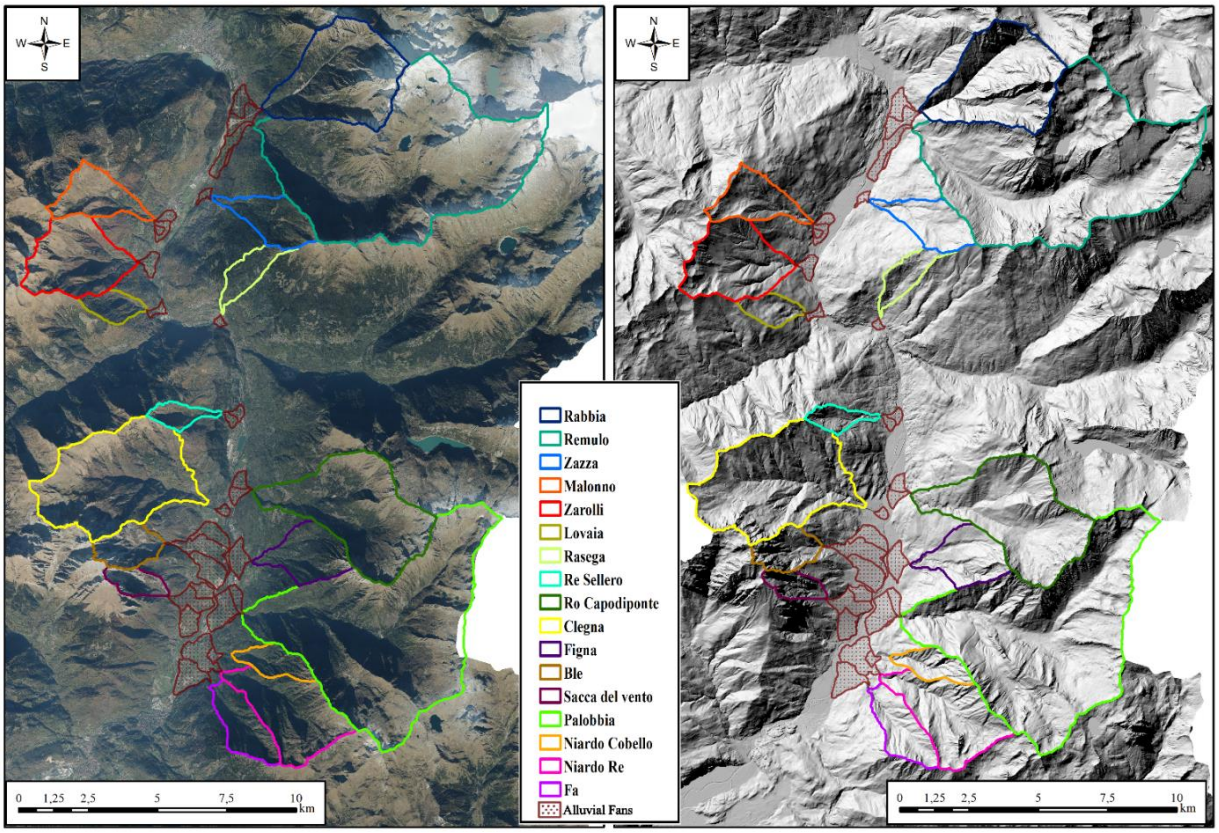


Figure 5.2. Mapped mountain catchments with related valley names and fans.

Basins name	**Area (km^2)**	**Relief (km)**	**Melton Ratio (R)**	**Mean Slope (°)**	**Loose Debris and Unvegetated Area (km^2)**	**Gravelly Riverbed Area (km^2)**	**Sparse Vegetation Area (km^2)**
Rabbia	13,2	2,5	0,68	42,0	4,5	0,1	2,1
Remulo	40,3	2,8	0,44	32,6	0,6	0,0	0,3
Zazza	4,0	1,9	0,94	28,8	0,0	0,0	0,1
Malonno	4,1	1,9	0,94	32,3	0,1	0,0	0,2
Zarolli	8,1	2,0	0,69	34,7	0,2	0,0	0,6
Lovaia	1,8	1,3	0,97	32,4	0,0	0,0	0,0
Rasega	2,2	1,7	1,16	28,7	0,0	0,0	0,0
Re Sellero	1,5	1,6	1,30	39,3	0,0	0,0	0,0
Ro Capodiponte	15,5	2,3	0,60	30,9	2,8	0,0	1,9
Clegna	16,8	2,0	0,48	32,5	1,1	0,0	1,9
Figna	4,7	2,0	0,91	32,0	0,1	0,0	0,1
Ble	2,9	1,8	1,04	40,8	1,5	0,0	0,4
Sacca del vento	1,7	1,7	1,32	46,2	0,7	0,0	0,3
Palobbia	38,1	2,4	0,40	33,8	9,6	0,1	6,5
Niardo Cobello	2,3	1,8	1,15	40,2	0,0	0,0	0,1
Niardo Re	8,5	1,9	0,67	35,3	0,4	0,0	0,3
Fa	3,9	1,7	0,88	37,4	0,0	0,0	0,0

Table 5.2. Geometrical parameters from catchment areas.

Basins name	Area (km^2)	Relief (m)	Metlon Ratio	Mean Slope (°)
Rabbia	1,2	235,3	0,22	10,7
Remulo	0,8	176,6	0,20	5,7
Zazza	0,1	78,3	0,21	12,2
Malonno	0,5	130,5	0,18	10,0
Zarolli	0,4	129,1	0,21	11,9
Lovaia	0,2	125,0	0,30	14,2
Rasega	0,1	79,0	0,25	8,9
Re Sellero	0,3	107,2	0,18	9,4
Ro Capodiponte	0,5	89,6	0,12	6,3
Clegna	0,2	64,7	0,16	6,9
Figna	0,6	126,9	0,16	7,4
Ble	2,7	430,9	0,26	8,6
Sacca del vento	1,8	375,1	0,28	11,9
Palobbia	1,3	147,5	0,13	8,4
Niardo Cobello	0,2	146,5	0,36	11,0
Niardo Re	1,1	220,6	0,21	8,4
Fa	0,9	184,6	0,19	7,7

Table 5.3. Geometrical parameters from alluvial fan areas.

For the probability assessment, we directly applied Equation 5.1 to our dataset listed in Table 5.3. Regarding magnitude estimations, we utilized Equation 5.2, referencing the dataset presented in Table 5.2. In place of the original vegetation degree categorizations suggested by Bovis and Jacob (1999) within the Weighted Stability Number (Ns), we opted for the pre-mapped classes provided in the DUSUF 2015. Drawing from our extensive understanding and expertise in debris flow dynamics, we selected the last three columns of Table 5.2 as the most relevant categories for assessing sediment instability and availability. Consequently, we assigned the geometrically progressive weights recommended by Bovis and Jacob (1999) to our newly defined classes as follows: a weight of 1 for the sparse vegetation class, 2 for the gravelly riverbed class, and 4 for the class characterized by loose debris and absence of vegetation. The entire computational procedure was automated in MATLAB for efficiency.

The culmination of our pragmatic approach entails generating a hazard ranking for basins, leveraging the framework of a contingency matrix. This methodology plots magnitude against the probability for each basin, with magnitude values being log-transformed to accommodate the substantial variance observed across basins and enhance visual interpretability. In this project, two sets of thresholds for probability and magnitude have been established to facilitate the reclassification of each basin within the contingency matrix. We adopted a straightforward and widely used method for determining these thresholds, specifically employing percentiles. By dividing the contingency matrix into nine distinct sections based on the 33rd and 66th percentiles along both the probability and magnitude axes, we achieved a nuanced stratification reflective of varying hazard levels. We employed two approaches to classify the risk levels of basins in terms of debris flow (DF) hazards: a detailed method with seven hazard classes and a more straightforward method with three classes. This dual classification strategy acknowledges the importance of DF probability over magnitude, recognizing that even low-magnitude debris

flows can result in significant economic losses and endanger lives. This prioritization is crucial since our primary objective is to mitigate any potential harm irrespective of the debris flow's scale. For the detailed classification:

We created a contingency matrix where each basin is assigned a hazard level based on its DF probability and magnitude.

- In the first column, where the DF probability is less than the 33rd percentile, the classification begins with 'Very Low' (VL) at the lowest magnitude, ascending to 'Low' (L) and 'Medium Low' (ML) as magnitude increases.
- In the middle column, for probabilities between the 33rd and 66th percentiles, the levels start at 'Medium Low' (ML) at the lowest magnitude and rise to 'Medium' (M) and 'Medium High' (MH).
- In the third column, representing the highest DF probability, the classification starts with 'Medium High' (MH) at the lowest magnitude and moves up to 'High' (H) and 'Very High' (VH) for increasing magnitudes.

For the straightforward classification:

- The 'Low' (L) class is designated for lower magnitude sections within the first column, extending to the lowest magnitude section of the second column.
- The 'Medium' (M) class spans diagonally from the upper-left to the lower-right section of the matrix.
- The 'High' (H) class is assigned to the highest magnitude sections in the second and third columns and includes the mid-magnitude section of the third column.

This dual approach enables detailed risk assessments for specific needs while providing a simplified classification for general planning and communication.

5.5 Risk estimation

5.5.1 Exposed elements

To estimate risk effectively, we collected data on elements exposed to hazards in areas potentially affected by debris flows (DF). Our methodology simplifies data collection and enhances replicability by assuming the entire alluvial fan is at risk. Although this approach is an approximation, it is a prudent strategy in regions where detailed local knowledge is lacking. We identified vulnerable man-made structures within these fans using two distinct datasets. The first dataset is derived from OpenStreetMap, which provides global coverage and aids in generalizing our approach. The second dataset comes from the Italian military's national synthesis database (database di sintesi nazionale - DBSN), which offers finer detail. We considered the following exposed elements (as separate data): the buildings (all types), the railway, the most important national road in the Val Camonica (SS42), and the secondary roads. Our analysis revealed that the roads and railways are consistently mapped in both the OpenStreetMap and DBSN datasets. However, there is a significant discrepancy in building data between the two sources. OpenStreetMap often omits smaller villages and buildings, evident from satellite imagery checks, whereas DBSN captures nearly all buildings comprehensively. By overlaying these datasets on our fan maps, we extracted detailed

information on the presence of buildings (counts) and the lengths of railways, SS42, and other roads within each fan area. This method allowed us to quantify the exposure of built environments to potential debris flow hazards effectively.

5.5.2 Risk analysis framework implementation

In our risk analysis, we employed two contingency risk matrices, detailed and straightforward, following the hazard assessment. A key challenge was integrating diverse exposed elements into a single, coherent metric for classifying basins based on hazard classes and exposure values. This challenge arises from the need to merge count data (buildings) with linear measurements (others). To address this, we normalised each exposed element relative to the data from the fans, scaling them to a range between 0 and 1. This normalisation allowed us to convert different types of data into a uniform format, facilitating aggregation. We then weighted the secondary roads by an arbitrary factor of 0.2, under the assumption that these roads are less critical in the context of debris flow events, both in terms of human presence during such events and their importance to the overall connectivity of the valley. Primary elements like railways and the SS42 road were given full weight due to their importance. Thus, we aggregated the normalised and weighted values of all exposed elements to produce a composite exposure score for each fan, which we then re-normalised to range from 0 to 1. This standardised metric integrates all relevant exposure data, providing a consistent basis for further risk classification in our contingency matrices. The risk classification incorporates both detailed and straightforward methods, which follow the contingency matrix structure established in the hazard assessment. The classification uses the 33^{rd} and 66^{th} percentiles of the computed final exposure values to assign risk categories to basins/fans based on their initial hazard classification.

In the detailed method, the risk class is adjusted from the hazard class based on the exposure value's percentile ranking:

- Basins with exposure values below the 33^{rd} percentile have their hazard class downgraded one level (e.g., M becomes ML).
- Basins with exposure values above the 66^{th} percentile have their hazard class upgraded one level (e.g., M becomes MH).
- Basins with exposure values between the 33^{rd} and 66^{th} percentiles retain their original hazard class.

The exceptions to these adjustments are the extreme categories in the matrix, where basins cannot be shifted further up or down. In these cases, basins in the lowest hazard category remain at the lowest risk, and those in the highest hazard category are placed at the highest risk.

The straightforward method also reclassifies hazard levels, from the simplified hazard method, based on exposure value percentiles, using four broad classes:

- Basins in the L hazard class are classified as L risk if their exposure is below the 66^{th} percentile and M risk if above.
- Basins in the M hazard class are categorized as M risk if their exposure is below the 66^{th} percentile and H risk if above.

- For the H hazard class, an additional risk category, VH, is introduced. Basins are classified as H risk if their exposure is below the 33rd percentile and VH risk if above.

This dual framework enables nuanced risk management based on detailed data while providing a simpler, more accessible model for broader strategic planning and communication.

These 2 methods were applied separately using the exposure values obtained from the 2 datasets, OpenStreetMap and DBSN.

5.6 Results

5.6.1 Debris flow hazard

As direct results from the application of equations 1 and 2 in our study area, we obtained the probability of being a debris flow catchment, and the expected magnitude in volume is shown in the Table 5.4. It suddenly appears the huge variation in predicted volumes, thus the need of plotting them in a log-axis.

Basins name	**Probability (Bertrand et al 2013)**	**Magnitude (Bovis & Jacob 1999) [km^3]**
Rabbia	0.826	218027.81
Remulo	0.533	30712.39
Zazza	0.852	570.88
Malonno	0.752	1705.78
Zarolli	0.851	5265.77
Lovaia	**0.934**	1.10
Rasega	0.799	2.24
Re Sellero	0.735	149.82
Re Capodiponte	0.382	114920.83
Clegna	0.551	34641.57
Figna	0.576	1419.29
Ble	0.807	26577.16
Sacca del Vento	0.900	10639.10
Palobbia	0.558	**535771.24**
Niardo Cobello	0.919	191.49
Niardo Re	0.743	8857.61
Fa	0.669	276.42

Table 5.4. Probability of each catchment produce a debris flow with the expected magnitude (volume).

This hazard contingency matrix arranges basins such that those with higher probabilities and magnitudes of debris flow events are positioned towards the upper-right corner, signifying elevated critical conditions. The implementation of percentile thresholds further accentuates the segregation of basins more hazardous. The Rabbia basin emerges as the most critical, closely followed by the Ble and Sacca del Vento basins, with the Zarolli basin also showing significant hazards. All the basins were then reclassified based on their position in the hazard contingency matrices (Figure 5.3 and Figure 5.4), as described in the previous section. Using this pragmatic

approach provides valuable quantitative information, enabling us and local authorities to prioritize areas based on the estimated hazard class of each basin.

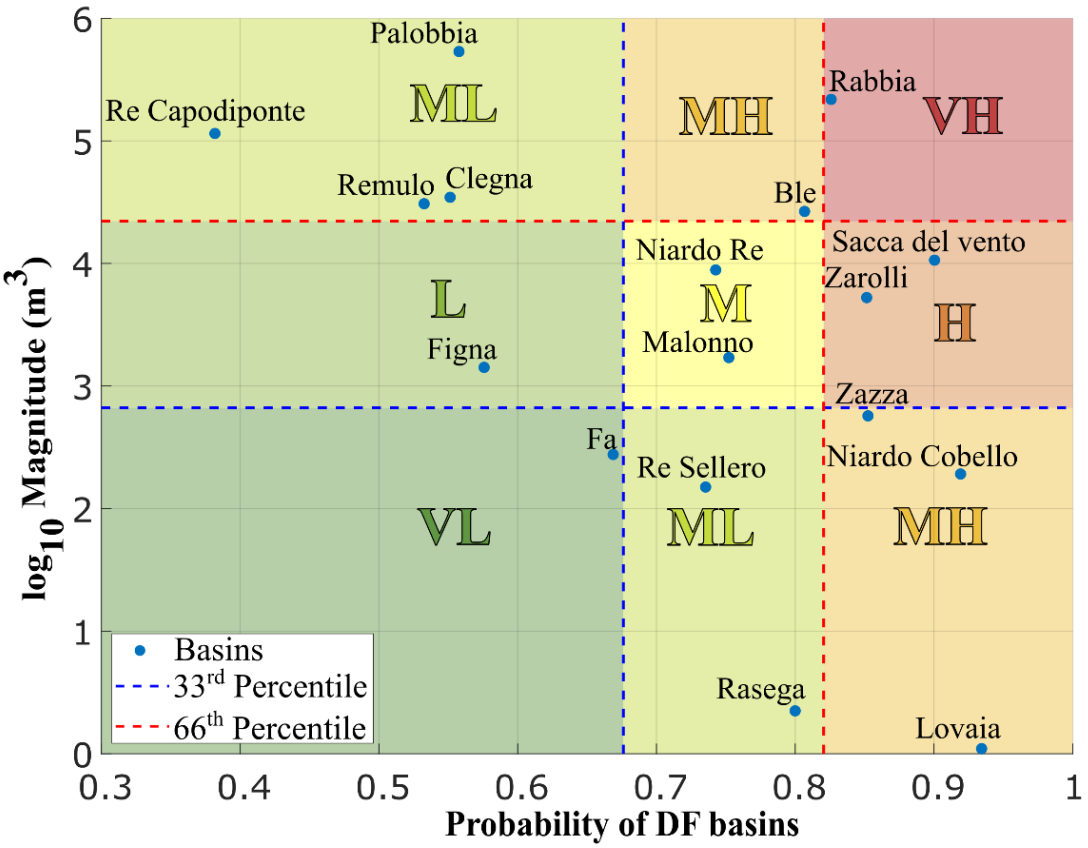


Figure 5.3 Detailed resulting hazard matrix

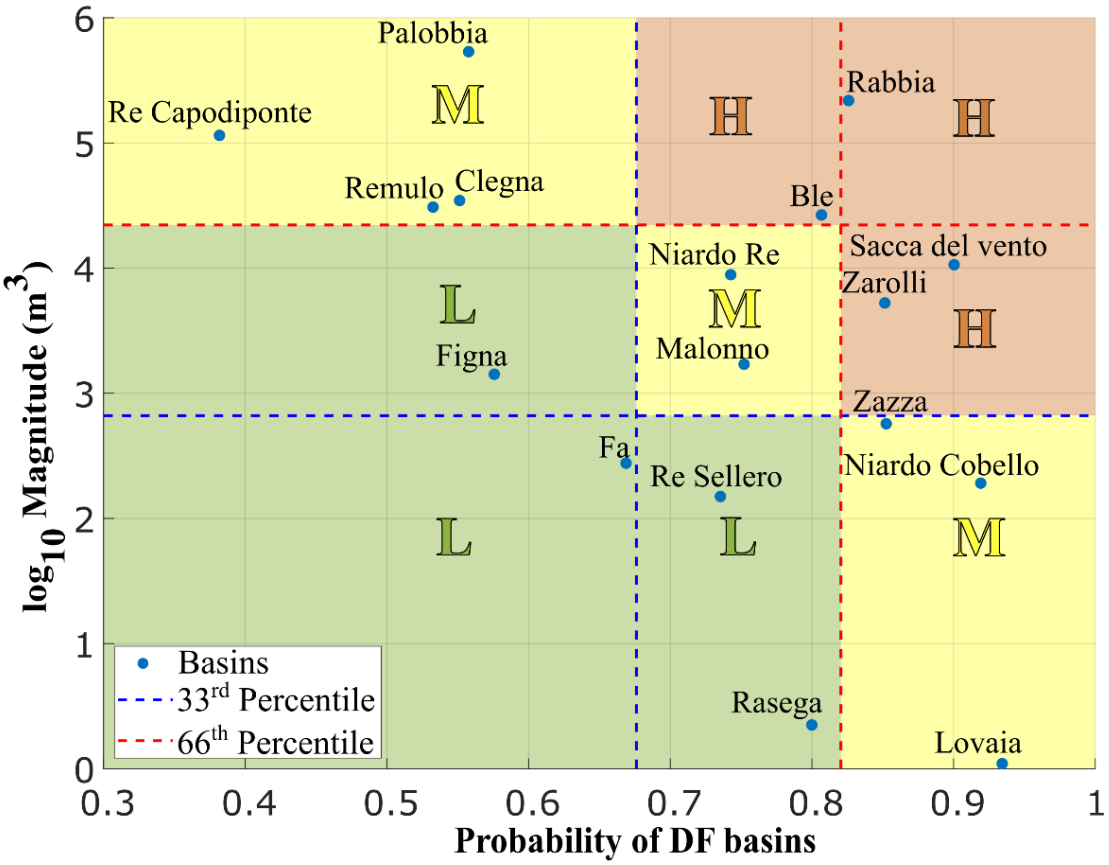


Figure 5.4. Straightforward resulting hazard matrix.

5.6.2 Debris flow risk

The information of the exposed elements, showed in Table 5.5, were summarized in a single value via the weighted normalization process described in the paragraph 3.2.

Basins name	**Exposed_elemnts_DBSN**	**Exposed_elemnts_Open**
Rabbia	0.249	0.200
Remulo	0.066	0.021
Zazza	0	0
Malonno	0.599	0.467
Zarolli	0.173	0.204
Lovaia	0.087	0.136
Rasega	0.069	0.0911
Re Sellero	0.214	0.162
Re Capodiponte	0.507	0.528
Clegna	0.029	0.003
Figna	0.686	0.694
Ble	0.391	0.179
Sacca del Vento	0.200	0.230
Palobbia	**1**	0.895
Niardo Cobello	0.073	0.337
Niardo Re	0.635	0.927
Fa	0.641	**1**

Table 5.5. Resulting aggregation values of weighted and normalized exposed elements for both database (DBSN central column, OpenStreetMap third column).

The risk reclassification matrices, both detailed and straightforward, are displayed in Figure 5.5 and Figure 5.6. These matrices are based on exposure values from the DBSN dataset, which is considered more reliable. Due to space constraints, only results from the DBSN dataset are presented. Detailed hazard and risk reclassification results for each basin are tabulated Table 5.6, for easily consultation. Notably, the Rabbia catchment emerges as the highest risk area, classified as very high (VH) risk in both reclassification methods. Following closely are the Zarolli and Sacca del Vento catchments, categorized as high (H) risk in the detailed method and very high (VH) in the simplified method. The Ble catchment is also significant, classified as very high (VH) in the simplified method and medium-high (MH) in the detailed approach. This classification system enables focused prioritization of specific catchments for targeted risk management and resource allocation.

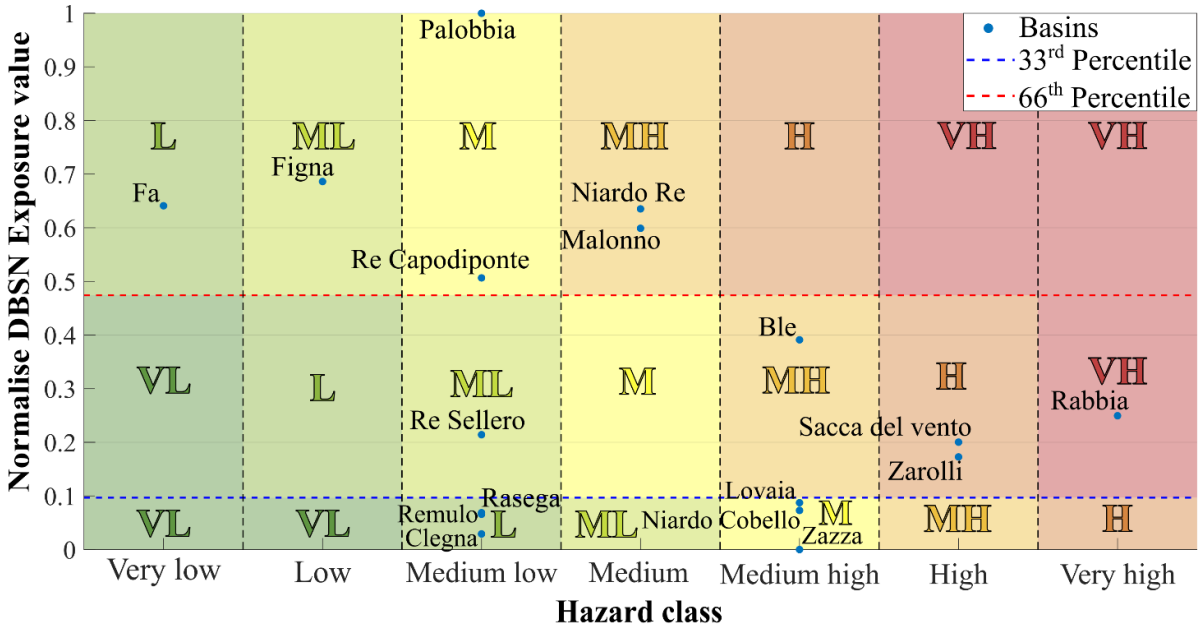


Figure 5.5. Detailed resulting risk matrix for DBSN exposed elements.

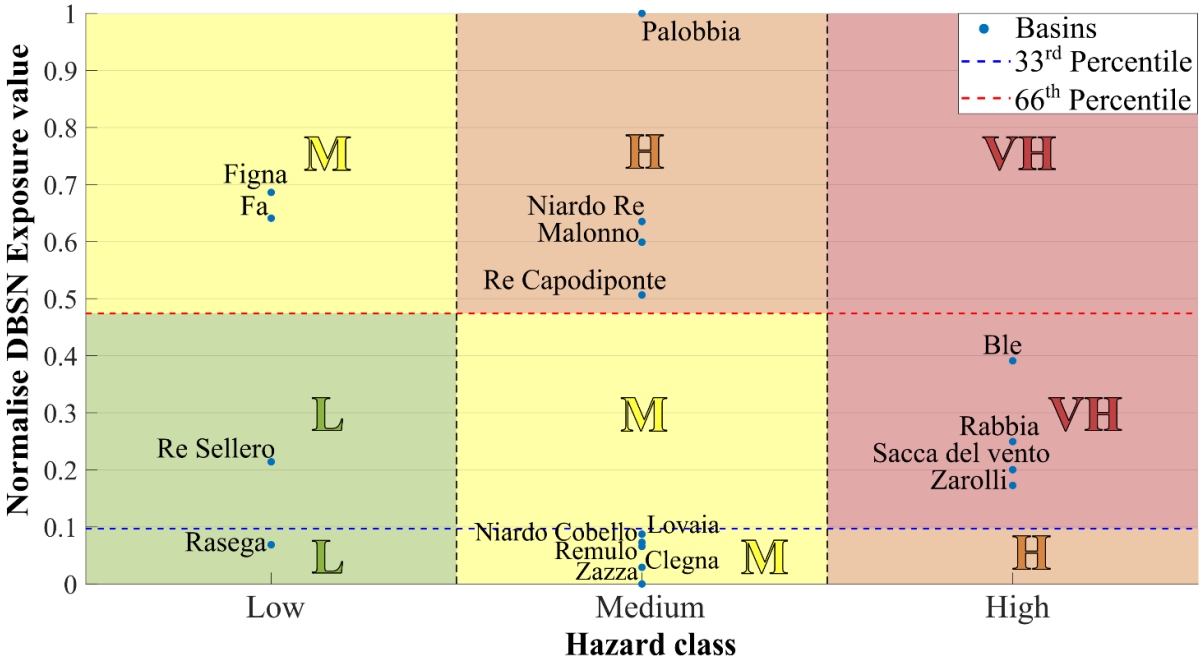


Figure 5.6. Straightforward resulting risk matrix for DBSN exposed elements.

Basins name	**Detailed Hazard Level**	**Simplified Hazard Level**	**Detailed Risk Level Open**	**Simplified Risk Level Open**	**Detailed Risk Level DBSN**	**Simplified Risk Level DBSN**
Rabbia	Very High	High	Very High	Very High	Very High	Very High
Remulo	Medium Low	Medium	Low	Medium	Low	Medium
Zazza	Medium High	Medium	Medium	Medium	Medium	Medium
Malonno	Medium	Medium	Medium High	High	Medium High	High
Zarolli	High	High	High	Very High	High	Very High
Lovaia	Medium High	Medium	Medium	Medium	Medium	Medium
Rasega	Medium Low	Low	Low	Low	Low	Low
Re Sellero	Medium Low	Low	Low	Low	Medium Low	Low
Re Capodiponte	Medium Low	Medium	Medium	High	Medium	High
Clegna	Medium Low	Medium	Low	Medium	Low	Medium
Figna	Low	Low	Medium Low	Medium	Medium Low	Medium
Ble	Medium High	High	Medium High	Very High	Medium High	Very High
Sacca del Vento	High	High	High	Very High	High	Very High

Palobbia	Medium Low	Medium	Medium	High	Medium	High
Niardo Cobello	Medium High	Medium	Medium High	Medium	Medium	Medium
Niardo Re	Medium	Medium	Medium High	High	Medium High	High
Fa	Very Low	Low	Low	Medium	Low	Medium

Table 5.6. Comprehensive hazard and risk summary categorized results.

Finally, a helpful output in the context of linear infrastructure design is the overlay of the proposed alignment on the computed debris-flow hazard across each alluvial fan (e.g. Figure 5.7 with existing infrastructure). An eventual debris flow event will affect approximately the same area of the mapped alluvial fan; it is therefore necessary to assume the whole alluvial fan is uniformly hazardous (out of detailed local analysis). Under this assumption, during the planning stage of major linear works (e.g., a railway or national road), plotting candidate alignments against the hazard map helps identify, and therefore avoid, the most critical, highly hazardous fans.

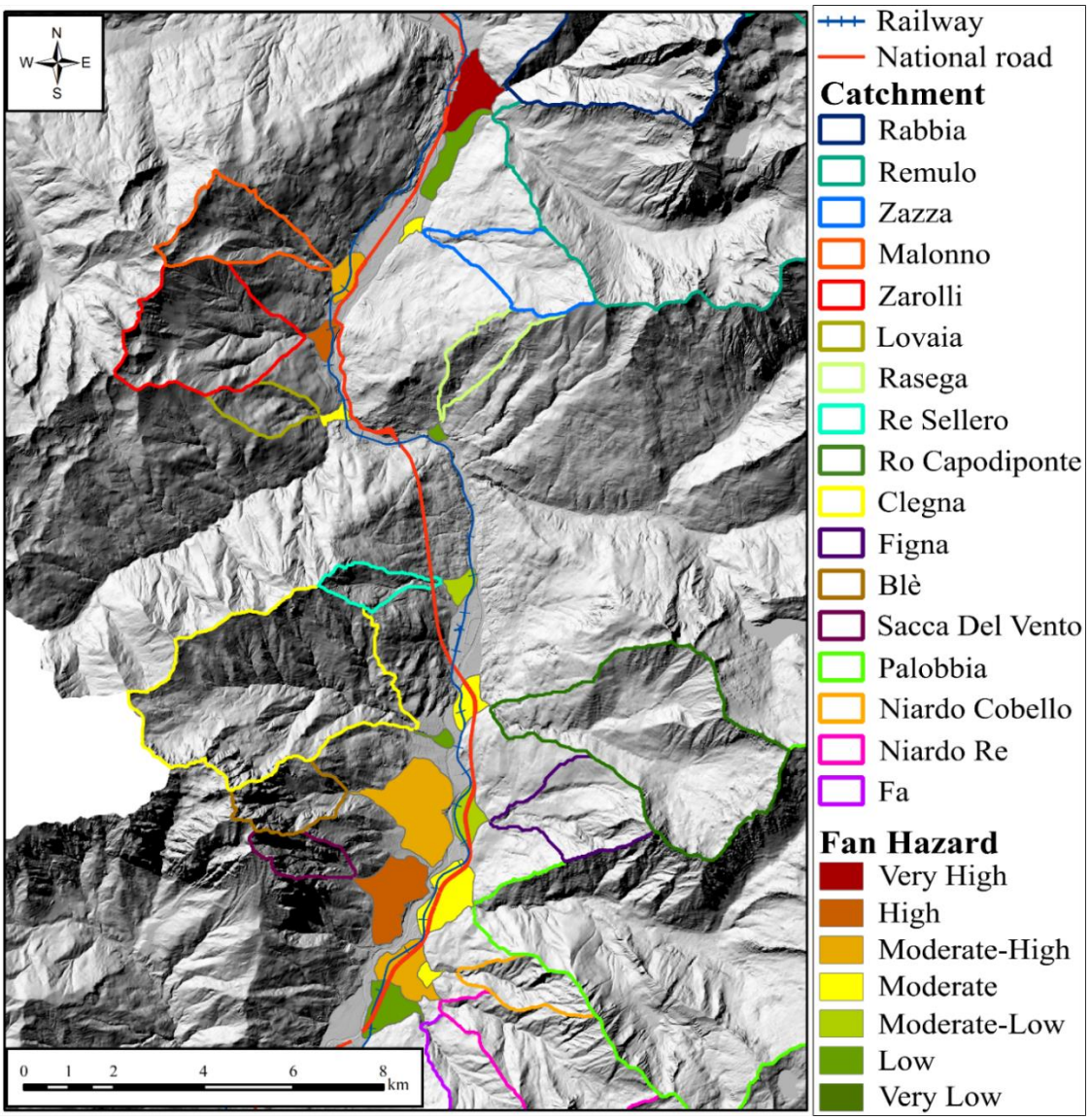


Figure 5.7. Alluvial fans' debris flow hazard according to detailed classification overlaid by the Railway line and the road SS42.

5.7 Conclusion

A methodology to rapidly perform debris-flow hazard-ranking and risk-ranking for the drainage catchments in a study area extending more than 400 km² was developed. The presented approach can be easily extended to larger areas thanks to its implementation in GIS and the use of MATLAB routines to automate the calculations. The methodology was developed in relation to the field trial of a new Monitoring and Early Warning System (MEWS). In this respect, it is desirable to install experimental systems in high-hazard catchments. In high-risk basins, it is desirable to install highly reliable, tried-and-tested systems that are compatible with civil-protection needs. The implemented methodology involves the use of alluvial-fan morphometry information to obtain the general susceptibility of a catchment to produce debris-flow events and relies on the land-cover information and topography of the above basins to estimate the expected magnitude in terms of volume. Then, the methodology uses a series of contingency matrices to rank the catchment hazard, followed by a risk estimation intersecting the alluvial

fan with the exposed elements. In the study area of Val Camonica (Italy), the results show the Val Rabbia catchment as the most critical, with a very high debris-flow risk. The presented approach can be refined by introducing a runout analysis for each drainage channel to obtain a better delineation of the elements at risk. Moreover, the hazard map of the alluvial fans could enhance the linear infrastructure planning phase, avoiding the critical ones. The simplicity of implementation and the efficiency of the data processing make the presented methodology suitable for the application to larger study areas in a wide range of geographies and developing countries.

5.8 References chapter 5

Autorità di bacino distrettuale del fiume Po (ADBPo). (2018). *Piano di Assetto Idrogeologico – Elaborato 2: Atlante dei rischi idraulici e idrogeologici*. https://pai.adbpo.it/index.php/documentazione-pai/

Bertrand, M., Liébault, F., & Piégay, H. (2013). Debris-flow susceptibility of upland catchments. *Natural Hazards, 67*, 497–511. https://doi.org/10.1007/s11069-013-0575-4

Bovis, M. J., & Jakob, M. (1999). The role of debris supply conditions in predicting debris flow activity. *Earth Surface Processes and Landforms, 24*(11), 1039–1054. https://doi.org/10.1002/(SICI)1096-9837(199910)24:11%3C1039::AID-ESP29%3E3.0.CO;2-U

D'Agostino, V., & Marchi, L. (2001). Debris flow magnitude in the Eastern Italian Alps: Data collection and analysis. *Physics and Chemistry of the Earth, Part C: Solar, Terrestrial and Planetary Science, 26*(9), 657–663. https://doi.org/10.1016/S1464-1917(01)00064-2

de Haas, T., & Densmore, A. L. (2019). Debris-flow volume quantile prediction from catchment morphometry. *Geology, 47*(8), 791–794. https://doi.org/10.1130/G45950.1

Dong, J.-J., Lee, C.-T., Tung, Y.-H., Liu, C.-N., Lin, K.-P., & Lee, J.-F. (2009). The role of sediment budget in understanding debris flow susceptibility. *Earth Surface Processes and Landforms, 34*, 1612–1624. https://doi.org/10.1002/esp.1850

Istituto Geografico Militare (IGM). (2024). *DBSN – Database di Sintesi Nazionale* [Data set]. https://www.igmi.org/it/dbsn-database-di-sintesi-nazionale

Ioriatti, I., Coviello, V., Berti, M., & Comiti, F. (2023). Study of debris-flow initiation through the analysis of seismic signals. *E3S Web of Conferences, 415*, 03014. https://doi.org/10.1051/e3sconf/202341503014

ISPRA – Istituto Superiore per la Protezione e la Ricerca Ambientale. (2024). *IFFI – Inventario dei Fenomeni Franosi in Italia* [Database]. https://www.progettoiffi.isprambiente.it/

Marchi, L., & Brochot, S. (2000). Les cônes de déjection torrentiels dans les Alpes françaises: Morphométrie et processus de transport solide torrentiel. *Revue de Géographie Alpine, 88*(3), 23–38. https://doi.org/10.3406/rga.2000.2999

Melton, M. A. (1965). The geomorphic and paleoclimatic significance of alluvial deposits in southern Arizona. *The Journal of Geology, 73*(1), 1–38. https://doi.org/10.1086/627044

Mueting, A., Bookhagen, B., & Strecker, M. R. (2021). Identification of debris-flow channels using high-resolution topographic data: A case study in the Quebrada del Toro, NW Argentina. *Journal of Geophysical Research: Earth Surface, 126*, e2021JF006330. https://doi.org/10.1029/2021JF006330

OpenStreetMap Foundation (OSMF). (2024). *OpenStreetMap* [Data set]. https://www.openstreetmap.org

Schimmel, A., & Berti, M. (2023). MAMODIS: A low-cost monitoring system for debris flows based on infrasound and seismic signals. *E3S Web of Conferences, 415*, 03025. https://doi.org/10.1051/e3sconf/202341503025

Wilford, D., Sakals, M., Innes, J., Sidle, R., & Bergerud, W. (2004). Recognition of debris flow, debris flood, and flood hazard through watershed morphometrics. *Landslides, 1*(1), 61–66. https://doi.org/10.1007/s10346-004-0008-2

6. Chapter 6: General conclusions

This PhD research aimed to investigate how advanced landslide susceptibility and probability modeling can support hazard assessment and risk-informed decision-making, with particular attention to complex geological settings and real-world infrastructure planning. The work focused on three fundamental aspects: improving the reliability and interpretability of susceptibility models, integrating dynamic triggering factors such as rainfall into probability estimates, and developing methodologies that are practical and accessible for decision-makers and stakeholders. Below, we summarize the main outcomes of the thesis, discussing the answers to the specific research questions.

6.1 Addressing the research questions

How can we conduct a comprehensive susceptibility analysis in an area prone to different types of landslides?

Simplifying the analysis by assuming a single landslide type, or aggregating multiple types into a single category, can lead to misleading results and a loss of critical information. This thesis was developed on the premise that a type-specific approach is essential to obtain comprehensive and actionable susceptibility and probability assessments.

Chapter 2 highlights the importance of differentiating landslide types. The original IFFI inventory, characterized by long temporal coverage but low spatial resolution, predominantly reflects large-scale phenomena, classified with classical categories (e.g. Cruden and Varnes 1996), which sometimes could be unclear and operator-depending. Despite these limitations, discriminating between landslide types along the planned railway line provides decision-makers with vital information: they can anticipate the predominant hazards the railway may encounter and adapt the track to avoid the most dangerous phenomena (e.g., rockfalls versus slow-moving landslides). This differentiation also allows the selection of appropriate, cost-effective mitigation strategies, avoiding unnecessary expenditure on measures suited to a different hazard type (e.g. huge complex/sliding type against rockfall ones).

In Chapter 3, a carefully pondered reclassification of mapped landslides was carried out to characterize their features and better understand the underlying processes. The majority of the landslides documented in this chapter are shallow or superficial processes, often unmapped in traditional inventories due to their small scale and transient nature. Despite their subtlety in long-term datasets, these shallow landslides can have severe impacts on land-use planning. Therefore, producing a detailed, type-classified inventory for shallow landslides is crucial. Different shallow landslide types may behave differently and require distinct mitigation strategies; for instance, shallow debris slides are rapid and highly hazardous compared to slower earth slides, while rock slides, due to their volume and velocity, present even higher risks.

The type-classified inventory developed in Chapter 3 served as the ground truth in Chapter 4. Here, landslide movement was aggregated into five primary types, forming the basis for five separate landslide probability maps. These maps reveal how different types respond to both static conditioning factors and dynamic triggering factors such as rainfall time-series. The analysis shows geotechnically consistent patterns: mean slope predominantly influences debris

slides and debris flows, while earth slides and earth flows are more affected by geology and land cover; rock slides are additionally influenced by geology bedding orientation. This approach enables the generation of type-specific probability maps even for shallow landslides, such as those triggered by the extreme rainfall events in May 2023 in Emilia-Romagna, providing the foundation for a more comprehensive susceptibility assessment.

Finally, Chapter 5 focuses on a single landslide type (alpine debris flows) but the analysis remains type-specific, adhering to the principle of separate treatment for each landslide category. Importantly, this type-specific approach could involve the use of distinct modeling techniques to better tailor the landslide types hazard: for example, empirical based methods were applied to alpine debris flows, while deep learning models (TNN) were used for multi-type rainfall-triggered landslides in Chapter 4. The selection of methodology depends not only on landslide type but also on environmental context, data availability, and computational requirements.

Do we need detailed and classified landslide inventories for susceptibility analysis? How can we produce them?

As anticipated in the introduction, the quality and completeness of the landslide inventory used to train susceptibility models play a critical role in the accuracy of the resulting maps. The contrast in inventory quality between IFFI (Chapter 2) and the manually mapped Emilia-Romagna dataset (Chapters 3 and 4) is striking. Differences in spatial resolution and the nature of mapped phenomena directly influence the outputs of susceptibility and probability analyses. Even when different partition units and sampling strategies are applied, the high-resolution, type-classified inventory used in Chapter 4 provides more reliable and informative results.

Completeness and spatial detail are therefore fundamental in susceptibility modelling. The creation of a reliable, detailed, and comprehensive inventory for the May 2023 Emilia-Romagna event was essential. Producing such inventories is challenging, despite recent technological advances. Chapter 3 demonstrates the value of manual mapping: it provides a robust, accurate baseline while allowing researchers to better understand the processes at work. Through direct GIS inspection, it becomes possible to interpret morphologies, colour variations, and the underlying physical processes, translating into insights that automated approaches tend to hide from our understanding. In this sense, manual mapping represents a meaningful investment of time and expertise.

However, manual mapping is labor-intensive and may not always be feasible in situations requiring rapid response. In such cases, ML models, particularly convolutional neural networks (CNNs), can provide effective tools for landslide mapping during emergencies, especially in large-scale disasters. One key advantage is the substantial reduction in time required to map vast numbers of landslides, providing a reliable starting point for experts to refine and validate outputs, thereby ensuring both speed and accuracy.

How can we effectively reclassify the landslide susceptibly results to communicate with non-technical final decision makers?

The need for a clear and accessible reclassification output, easily understood by non-technical decision-makers, is one of the most important concepts, yet it is frequently overlooked in the literature. Chapter 2 highlights the absence of a standardized reclassification methodology, emphasizing the confusion that often arises in published works. In this chapter, the proposed methodology employs several reclassification approaches in an ensemble framework to enhance the reliability of the results, demonstrating the differences between methods and capturing the most reasonable thresholds. The methodology also quantifies the reliability of each association with the specific classes, yielding a more comprehensive outcome.

The resulting reclassification maps are clear and less influenced by the choice of a single technique, making them both objective and reliable. Currently, the ensemble integrates six reclassification methods; however, future developments could expand this framework to include a broader set of techniques. Such an improvement would also help identify the most robust objective methods; for instance, in our case, the ROC curve discretization with AUC optimization. This approach represents a step toward a unified, rigorously tested, and objective methodology that could be standardized for territorial planning purposes.

How can preparatory factors (antecedent rainfall) be successfully integrated into susceptibility analysis with specific or multiple landslide events?

Predisposing factors are generally static and easily available over long periods (spanning seasons or even centuries), whereas preparatory factors are transient, varying over days or weeks, and are more challenging to capture. Further complication arises because triggering and preparatory factors are directly linked to the landslide occurrence, requiring knowledge of the exact event date to correctly associate these factors with the target. In many cases, however, the precise date of the landslide is ambiguous, making probability-based landslide modeling less common due to the lack of detailed, date-specific inventories.

In this context, the introduction of a Transformer Neural Network (TNN) offers a promising solution. The TNN model can automatically capture the natural spatial and temporal correlations between triggering and preparatory factors and the resulting landslide. This allows the model to detect the relevant time frames (tokens) within the sequence without explicitly differentiating preparatory from triggering factors or knowing the exact day of occurrence, which may become superfluous. Furthermore, the TNN eliminates the need to predefine a time window from which to derive an aggregate value for these temporal factors; the model inherently understands the relationship between antecedent preparatory conditions, their position in the sequence, and the target landslide.

In Chapter 4, the TNN was applied to investigate the relationship between antecedent rainfall and two major landslide events, separated by only 14 days, in Emilia-Romagna, Italy. The methodology used 30 days of antecedent rainfall before the second event, along with the triggering events, treated together as a single continuous rainfall time series. The resulting model produced highly promising metrics and visualizations, and, when coupled with explainability techniques, allowed a deeper understanding of model behavior and increased reliability. Chapter 4 thus represents a first step toward a more comprehensive model, which can be trained over multiple events with their associated preparatory and triggering rainfall,

enabling the model to learn the general role of rainfall and rainfall extremes, including sequential extreme events like those observed here.

In developing countries with limited data on landslides, how can we conduct a reliable analysis of debris flow risk?

In many regions, particularly in tropical environments, documentation of past landslides is often sparse or inconsistent. Rapidly growing vegetation, complex topography, and frequent cloud cover make mapping landslide inventories challenging, even with advanced automated machine learning techniques. Moreover, climate change is exacerbating the frequency and intensity of extreme rainfall, increasing the likelihood of debris flows in areas with little historical record. In contexts where the detailed information required to train statistical or machine learning models is lacking, empirical and physically-based equations can provide a practical and reliable alternative.

In Chapter 5, a comprehensive methodology for assessing debris flow risk was implemented in a test area in Lombardy, Italy. This approach relies only on widely available datasets: a Digital Elevation Model (DEM) to delineate catchments and alluvial fans, and satellite-derived land cover information to assess potential exposure. The methodology integrates hazard assessment with risk matrices to rank basins according to their susceptibility and potential impact, highlighting priority areas where experimental early warning systems could be installed.

This chapter demonstrates that even relatively simple empirical equations, when coupled with globally available geospatial data, can effectively categorize debris flow risk in alpine environments. Such an approach is particularly valuable in developing countries, where landslide inventories (specifically for debris flows) are often missing. By relying solely on morphometric and land cover information, the methodology provides a fast, reproducible, and scalable framework, allowing authorities to prioritize interventions and enhance community safety even in data-scarce regions.

6.2 Integration of physically-based and data-driven models: an open question

As introduced in Section 1.3, physically based models aim to reproduce slope failure mechanisms by explicitly representing the governing processes (e.g., shear strength, pore-pressure changes, and geometry). They rely on geotechnical and hydrological properties to estimate stability through the Factor of Safety (Fs). This makes them conceptually consistent with site-specific slope stability analyses usually adopted by geologists and engineers during design (e.g., limit equilibrium and FEM/FDM tools such as GeoStudio/SLOPE-W, FLAC 2D/3D). In addition, Fs-based verification is embedded in engineering practice and aligns with the safety checks required by national building regulations (e.g., the Italian *Norme Tecniche per le Costruzioni*, NTC). For these reasons, spatial Fs patterns provide an interpretable bridge between susceptibility assessments and engineering decision-making.

However, physically based approaches typically require detailed and spatially representative geotechnical/hydrogeological inputs. This requirement often limits their feasibility at regional

scales, where input heterogeneity is high and data availability is sparse, making them less suitable as stand-alone tools for large-area infrastructure screening (e.g., railway planning) (Rai & Sahu 2020). In this context, data-driven susceptibility models remain attractive because they are scalable, can exploit widely available predictors (DEM derivatives, geology, land cover), and can provide robust first-order screening.

A promising direction is therefore a hybrid framework that preserves (i) the scalability and statistical power of data-driven models and (ii) the process coherence and engineering interpretability of physically based modelling. Three complementary integration strategies can be considered.

- Scalability-driven hybridisation:
  a) Source-area prediction + physically based runout. A practical strategy is to use a data-driven model to identify potential initiation zones (source areas), and then apply a physically based runout model to simulate propagation and deposition (Lima et al. 2023, Bornaetxea et al. 2025). This produces outputs that are closer to hazard-relevant footprints (initiation and impact area), particularly where runout controls exposure.
  b) Two-tier workflow of regional screening + local mechanics. Site-specifc physically based integration analys, not just run out but fully mechanical analisys. At the feasibility stage, susceptibility is first mapped using scalable data-driven models over the full corridor. Then, only the most critical sectors (high susceptibility) are analysed using site-specific or semi-distributed physically based stability models. This preserves computational feasibility, gicen the high computation cost of physiscally based models (Rai & Sahu 2020), while improving consistency with engineering practice where decisions are ultimately made.
- Parameter-enhancement coupling:
  a) Data-driven spatialisation of measured parameters. Data-driven methods can be used to spatialise sparse field measurements (e.g., cohesion, friction angle, unit weight, hydraulic conductivity) by learning their relationship with spatial proxies such as lithology, terrain attributes, and geomorphological settings (McBratney et al. 2003, Kim & Ji 2022). The resulting parameter fields can then serve as inputs for physically based modelling over targeted areas, reducing the gap between point-scale data and distributed modelling needs. Moreover, data-driven models can support parameter calibration of physical models (e.g., optimising hydraulic or strength parameters within plausible bounds to reproduce observed failures).
  b) Physics to ML features parameter optimisation. Physical-model outputs (e.g., transient pore pressure, wetness index proxies, stability indices, displacement indicators) can be incorporated as additional predictors in ML models to inject process information into purely statistical mapping. Beyond a one-way coupling, future systems could adopt bidirectional (closed-loop) architectures (Ye et al. 2025), where ML models both use physical outputs and update physical parameters iteratively, enabling adaptive calibration under changing boundary conditions (e.g., evolving rainfall regimes).
- Meta-modelling (stacked hybrid approach):

a) A second-stage learner is trained to integrate (i) the original conditioning factors (e.g., geology, topography, land cover, hydrology proxies) and (ii) the parallel outputs of two first-stage models (e.g., susceptibility score 0-1, spatial Fs patterns, pore-pressure or wetness indicators). The meta-model learns how these inputs jointly relate to observed landslide occurrences, susceptibility and Fs, producing a final hazard-relevant score that can be calibrated against or expressed in a form of Fs, keeping the engegniring applicability. In this way, physical meaning is injected through the physically based component, while interpretability can be added to the final prediction using explainability tools such as SHAP (or other attribution methods), clarifying which predictors and which model outputs drive the result in each location. A key limitation is that this strategy can be computationally demanding, because it requires running (and often re-running for tuning/validation) both base models and then training the meta-learner, potentially over large areas and multiple scenarios.

6.3 Overall concluding aspects

Over the past decade, landslide susceptibility modeling has experienced remarkable growth, driven primarily by the integration of Machine Learning techniques and substantial advances in computational power. These developments have enabled researchers to handle larger datasets, incorporate complex spatial and temporal variables, and produce increasingly accurate susceptibility and probability maps. Despite these technical advancements, however, the practical application of landslide susceptibility and probability models in real-world planning and risk management remains limited.

This gap between research and application arises from two main factors. The first concerns technical and methodological challenges, which include the reliability of landslide inventories, the treatment of different landslide types, the integration of dynamic triggering factors such as rainfall, and the communication of complex model outputs to decision-makers. These challenges are systematically addressed throughout this thesis, which develops robust, type-specific, and interpretable methodologies for landslide susceptibility and probability modeling.

The second factor is the institutional inertia of public administrations, which often hesitates to adopt innovative methodologies despite their potential benefits. Changing established planning practices requires both trust in the models and a clear demonstration of their practical utility. The work presented in this thesis contributes to bridging this gap by producing methodologies that are transparent, replicable, and directly applicable to real-world scenarios, thereby facilitating the adoption of advanced modeling approaches by stakeholders.

A key demonstration of the practical relevance of this research is its collaboration with ITALFERR, the national railway construction company, which actively contributed to the development of the methodologies presented. In the context of railway planning, landslide susceptibility maps are not merely academic exercises, they provide essential information for preventing line disruptions and potential disasters. By identifying high-risk areas before construction, planners can implement targeted mitigation measures, avoiding costly interruptions, infrastructure damage, and, most importantly, potential casualties.

In summary, this PhD research represents a strategic integration of cutting-edge modeling techniques with practical, decision-oriented applications, emphasizing not only methodological innovation but also tangible societal and economic benefits. It demonstrates how advanced landslide susceptibility and probability models can move beyond theoretical frameworks to directly inform infrastructure planning, risk reduction, and disaster resilience strategies.

6.4 General references

Abbate, A., Papini, M., & Longoni, L. (2021). Analysis of meteorological parameters triggering rainfall-induced landslide: A review of 70 years in Valtellina. *Natural Hazards and Earth System Sciences*, 21(7), 2041–2058. https://doi.org/10.5194/nhess-21-2041-2021

Ado, M., Amitab, K., Maji, A. K., Jasińska, E., Gono, R., Leonowicz, Z., & Jasiński, M. (2022). Landslide susceptibility mapping using machine learning: A literature survey. *Remote Sensing*, 14(13), 3029. https://doi.org/10.3390/rs14133029

AGSS-RER (Area Geologia, Suoli e Sismica, Regione Emilia-Romagna). (1986). *Carta Geologica dell'Appennino emiliano-romagnolo, scala 1:10.000*. Regione Emilia-Romagna. Retrieved March 3, 2025, from https://geoportale.regione.emilia-romagna.it/

Ahmad, S. M., Sadhasivam, N., Lisa, M., Lombardo, L., Emil, M. K., Fadel, I., van Westen, C. J., Ahmed, A., & Tanyas, H. (2022). Standing on the shoulder of a giant landslide: An InSAR look at a slow-moving hillslope under melting glaciers in the western Karakoram. *EarthArXiv*. https://doi.org/10.31223/X5P63T

Ahmed, M., Tanyaș, H., Huser, R., Dahal, A., Titti, G., Borgatti, L., Francioni, M., & Lombardo, L. (2023). Dynamic rainfall-induced landslide susceptibility: A step towards a unified forecasting system. *International Journal of Applied Earth Observation and Geoinformation*, 125, 103593. https://doi.org/10.1016/j.jag.2023.103593

Agterberg, F. P. (1992). Combining indicator patterns in weights of evidence modeling for resource evaluation. *Nonrenewable Resources*, 1(1), 39–50. https://doi.org/10.1007/BF01782111

Aleotti, P., & Chowdhury, R. (1999). Landslide hazard assessment: Summary review and new perspectives. *Bulletin of Engineering Geology and the Environment*, 58(1), 21–44. https://doi.org/10.1007/s100640050066

Altman, D. G., & Bland, J. M. (1994). Diagnostic tests. 1: Sensitivity and specificity. *BMJ*, 308(6943), 1552. https://doi.org/10.1136/bmj.308.6943.1552

Amato, G., Eisank, C., Castro-Camilo, D., & Lombardo, L. (2019). Accounting for covariate distributions in slope-unit-based landslide susceptibility models: A case study in the alpine environment. *Engineering Geology*, 260, 105237. https://doi.org/10.1016/j.enggeo.2019.105237

Amatya, P., Kirschbaum, D., Stanley, T., & Tanyas, H. (2021). Landslide mapping using object-based image analysis and open source tools. *Engineering Geology*, 282, 106000. https://doi.org/10.1016/j.enggeo.2021.106000

Angelov, P. P., Soares, E. A., Jiang, R., Arnold, N. I., & Atkinson, P. M. (2021). Explainable artificial intelligence: An analytical review. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 11(5), e1424. https://doi.org/10.1002/widm.1424

Ardizzone, F., Cardinali, M., Carrara, A., Guzzetti, F., & Reichenbach, P. (2002). Impact of mapping errors on the reliability of landslide hazard maps. *Natural Hazards and Earth System Sciences*, 2(1/2), 3–14. https://doi.org/10.5194/nhess-2-3-2002

Ardizzone, F., Bucci, F., Cardinali, M., Fiorucci, F., Pisano, L., Santangelo, M., & Zumpano, V. (2023). Geomorphological landslide inventory map of the Daunia Apennines, southern Italy. *Earth System Science Data*, 15, 753–767. https://doi.org/10.5194/essd-15-753-2023

Arslan, A., Tari, E., Ziatdinov, R., & Nabiyev, R. I. (2014). Transition curve modeling with kinematical properties: Research on log-aesthetic curves. *Computer-Aided Design and Applications*, 11(5), 509–517. https://doi.org/10.1080/16864360.2014.902680

Atkinson, P. M., & Massari, R. (1998). Generalised linear modelling of susceptibility to landsliding in the central Apennines, Italy. *Computers & Geosciences*, 24(4), 373–385. https://doi.org/10.1016/S0098-3004(97)00117-9

Ayalew, L., & Yamagishi, H. (2005). The application of GIS-based logistic regression for landslide susceptibility mapping in the Kakuda-Yahiko Mountains, Central Japan. *Geomorphology*, 65(1–2), 15–31. https://doi.org/10.1016/j.geomorph.2004.06.010

Auflič, M. J., Bezak, N., Šegina, E., Frantar, P., Gariano, S. L., Medved, A., & Peternel, T. (2023). Climate change increases the number of landslides at the juncture of the Alpine, Pannonian and Mediterranean regions. *Scientific Reports*, 13, 23085. https://doi.org/10.1038/s41598-023-50314-x

Autorità di bacino distrettuale del fiume Po (ADBPo). (2018). *Piano di Assetto Idrogeologico – Elaborato 2: Atlante dei rischi idraulici e idrogeologici*. https://pai.adbpo.it/index.php/documentazione-pai/

Azarafza, M., Azarafza, M., Akgün, H., Atkinson, P. M., & Derakhshani, R. (2021). Deep learning-based landslide susceptibility mapping. *Scientific Reports*, 11(1), 24112. https://doi.org/10.1038/s41598-021-03585-1

Ba, Q., Chen, Y., Deng, S., Yang, J., & Li, H. (2018). A comparison of slope units and grid cells as mapping units for landslide susceptibility assessment. *Earth Science Informatics*, 11, 373–388. https://doi.org/10.1007/s12145-018-0335-9

Batar, A. K., & Watanabe, T. (2021). Landslide susceptibility mapping and assessment using geospatial platforms and weights of evidence (WoE) method in the Indian Himalayan Region: Recent developments, gaps, and future directions. *ISPRS International Journal of Geo-Information*, 10(3), 114. https://doi.org/10.3390/ijgi10030114

Bayer, B., Simoni, A., Mulas, M., Corsini, A., & Schmidt, D. (2018). Deformation responses of slow moving landslides to seasonal rainfall in the Northern Apennines, measured by InSAR. *Geomorphology*, 308, 293–306. https://doi.org/10.1016/j.geomorph.2018.02.020

Baum, R. L., Savage, W. Z., & Godt, J. W. (2008). *TRIGRS: A Fortran program for transient rainfall infiltration and grid-based regional slope-stability analysis, version 2.0* (USGS Open-File Report 2008-1159, 75 p.). Reston, VA: U.S. Geological Survey. https://doi.org/10.3133/ofr20081159

Bednarik, M., Magulová, B., Matys, M., & Marschalko, M. (2010). Landslide susceptibility assessment of the Kraľovany–Liptovský Mikuláš railway case study. *Physics and Chemistry of the Earth, Parts A/B/C*, 35(3–5), 162–171. https://doi.org/10.1016/j.pce.2009.12.002

Berrar, D. (2019). Performance measures for binary classification. https://doi.org/10.1016/B978-0-12-809633-8.20351-8

Berti, M., & Simoni, A. (2010). Field evidence of pore pressure diffusion in clayey soils prone to landsliding. *Journal of Geophysical Research*, 115, F03031. https://doi.org/10.1029/2009JF001463

Berti, M., & Simoni, A. (2012). Observation and analysis of near-surface pore pressure measurements in clay shale slopes. *Hydrological Processes*, 26, 2187–2205. https://doi.org/10.1002/hyp.9316

Berti, M., Martina, M. L. V., Franceschini, S., Pignone, S., Simoni, A., & Pizziolo, M. (2012). Probabilistic rainfall thresholds for landslide occurrence using a Bayesian approach. *Journal of Geophysical Research: Earth Surface*, 117(F4). https://doi.org/10.1029/2012JF002367

Berti, M., Ciccarese, G., Fanti, R., Lapietra, I., Parise, M., & Liso, I. S. (2025b). *Deliverable title: Rationale for selecting and scale-dependent weighing of predisposing factors*.

Berti, M., Pizziolo, M., Scaroni, M., Generali, M., Critelli, V., Mulas, M., Tondo, M., Lelli, F., Fabbiani, C., Ronchetti, F., Ciccarese, G., Dal Seno, N., Ioriatti, E., Rani, R., Zuccarini, A., Simonelli, T., & Corsini, A. (2025). RER2023: The landslide inventory dataset of the May 2023 Emilia-Romagna event. *Earth System Science Data*, 17, 1055–1078. https://doi.org/10.5194/essd-17-1055-2025 *(published version; replaces preprint citation)*

Bertrand, M., Liébault, F., & Piégay, H. (2013). Debris-flow susceptibility of upland catchments. *Natural Hazards*, 67, 497–511. https://doi.org/10.1007/s11069-013-0575-4

Bertolini, G., Corsini, A., & Tellini, C. (2017). Fingerprints of large-scale landslides in the landscape of the Emilia Apennines. In M. Soldati & M. Marchetti (Eds.), *Landscapes and landforms of Italy. World geomorphological landscapes* (pp. 215–224). Springer. https://doi.org/10.1007/978-3-319-26194-2_18

Bertolini, G., Guida, M., & Pizziolo, M. (2005). Landslides in Emilia-Romagna region (Italy): Strategies for hazard assessment and risk management. *Landslides*, 2, 302–312. https://doi.org/10.1007/s10346-005-0020-1

Beven, K. J., & Kirkby, M. J. (1979). A physically based, variable contributing area model of basin hydrology. *Hydrological Sciences Journal*, 24(1), 43–69. https://doi.org/10.1080/02626667909491834

Bhuyan, K., Rana, K., Ferrer, J. V., Cotton, F., Ozturk, U., Catani, F., & Malik, N. (2024). Landslide topology uncovers failure movements. *Nature Communications*, 15(1), 2633. https://doi.org/10.1038/s41467-024-46741-7

Bhuyan, K., Tanyaş, H., Nava, L., Puliero, S., Meena, S. R., Floris, M., van Westen, C., & Catani, F. (2023). Generating multi-temporal landslide inventories through a general deep transfer learning strategy using HR EO data. *Scientific Reports*, 13, 162. https://doi.org/10.1038/s41598-022-27352-y

Bonham-Carter, G. (1994). *Geographic information systems for geoscientists: Modelling with GIS* (Vol. 13). Elsevier.

Bordoni, M., Persichillo, M. G., Meisina, C., Crema, S., Cavalli, M., Bartelletti, C., Galanti, Y., Barsanti, M., Giannecchini, R., & D'Amato Avanzi, G. (2018). Estimation of the susceptibility of a road network to shallow landslides with the integration of the sediment connectivity. *Natural Hazards and Earth System Sciences*, 18(6), 1735–1758. https://doi.org/10.5194/nhess-18-1735-2018

Bordoni, M., Vivaldi, V., Lucchelli, L., Ciabatta, L., Brocca, L., Galve, J. P., & Meisina, C. (2021). Development of a data-driven model for spatial and temporal shallow landslide probability of occurrence at catchment scale. *Landslides*, 18, 1209–1229. https://doi.org/10.1007/s10346-020-01592-3

Bornaetxea, T., Blais-Stevens, A., Miller, B., & Marchesini, I. (2025). Combination of statistical and conceptual approaches for debris-flow susceptibility modelling at a regional scale, British Columbia, Canada. *Catena*, 256, 109044. https://doi.org/10.1016/j.catena.2025.109044

Bovis, M. J., & Jakob, M. (1999). The role of debris supply conditions in predicting debris flow activity. *Earth Surface Processes and Landforms*, 24(11), 1039–1054. https://doi.org/10.1002/(SICI)1096-9837(199910)24:11%3C1039::AID-ESP29%3E3.0.CO;2-U

Bozzolan, E., Holcombe, E. A., Pianosi, F., Marchesini, I., Alvioli, M., & Wagener, T. (2023). A mechanistic approach to include climate change and unplanned urban sprawl in landslide susceptibility maps. *Science of The Total Environment*, 858, 159412. https://doi.org/10.1016/j.scitotenv.2022.159412

Brabb, E. E., Pampeyan, E. H., & Bonilla, M. G. (1972). *Landslide susceptibility in San Mateo County, California* (Map I-360). U.S. Geological Survey. https://doi.org/10.3133/mf360

Brath, A., Casagli, N., Marani, M., Mercogliano, P., & Motta, R. (2023). *Rapporto della Commissione tecnico-scientifica istituita con deliberazione della Giunta Regionale n. 984/2023 e determinazione dirigenziale 14641/2023, al fine di analizzare gli eventi meteorologici estremi del mese di maggio 2023* (147 pp.). Regione Emilia-Romagna. Retrieved March 3, 2025, from https://www.regione.emilia-romagna.it/alluvione/rapporto-della-commissione-tecnico-scientifica

Brenning, A. (2005). Spatial prediction models for landslide hazards: Review, comparison and evaluation. *Natural Hazards and Earth System Sciences*, 5(6), 853–862. https://doi.org/10.5194/nhess-5-853-2005

Brenning, A. (2008). Statistical geocomputing combining R and SAGA: The example of landslide susceptibility analysis with generalized additive models. *Hamburger Beiträge zur Physischen Geographie und Landschaftsökologie*, 19(23–32), 410.

Broothaerts, N., Kissi, E., Poesen, J., Van Rompaey, A., Getahun, K., Van Ranst, E., & Diels, J. (2012). Spatial patterns, causes and consequences of landslides in the Gilgel Gibe catchment, SW Ethiopia. *Catena*, 97, 127–136. https://doi.org/10.1016/j.catena.2012.05.011

Budimir, M. E. A., Atkinson, P. M., & Lewis, H. G. (2015). A systematic review of landslide probability mapping using logistic regression. *Landslides*, 12, 419–436. https://doi.org/10.1007/s10346-014-0550-5

Burns, W. J., Mickelson, K. A., & Madin, I. (2016). *Landslide susceptibility overview map of Oregon*. Oregon Department of Geology and Mineral Industries.

Caleca, F., Lombardo, L., Steger, S., Tanyas, H., Raspini, F., Dahal, A., Nefros, C., Mărgărint, M. C., Drouin, V., Jemec-Auflič, M., Novellino, A., Tonini, M., Loche, M., Casagli, N., & Tofani, V. (2025). Pan-European landslide risk assessment: From theory to practice. *Reviews of Geophysics*, 63(1), e2023RG000825. https://doi.org/10.1029/2023RG000825

Campforts, B., Shobe, C. M., Overeem, I., & Tucker, G. E. (2022). The art of landslides: How stochastic mass wasting shapes topography and influences landscape dynamics. *Journal of Geophysical Research: Earth Surface*, 127(8), e2022JF006745. https://doi.org/10.1029/2022JF006745

Cantarino, I., Carrion, M. A., Goerlich, F., & Martinez Ibañez, V. (2019). A ROC analysis-based classification method for landslide susceptibility maps. *Landslides*, 16(2), 265–282. https://doi.org/10.1007/s10346-018-1071-6

Cao, C., Chicco, D., & Hoffman, M. M. (2020). The MCC-F1 curve: A performance evaluation technique for binary classification. *arXiv preprint*, arXiv:2006.11278. https://doi.org/10.48550/arXiv.2006.11278

Cardellini, S., Galdenzi, S., Guerrera, F., Pennacchioni, E., Piergiovanni, A., Pieruccini, P., Sandroni, P., Tosti, S., & Tramontana, M. (2009). *Carta Geologica d'Italia alla scala 1:50.000, Foglio 292 Jesi*. https://www.isprambiente.gov.it/Media/carg/note_illustrative/292_Jesi.pdf

Carrara, A. (1983). Multivariate models for landslide hazard evaluation. *Journal of the International Association for Mathematical Geology*, 15(3), 403–426. https://doi.org/10.1007/BF01031290

Carrara, A., Crosta, G., & Frattini, P. (2003). Geomorphological and historical data in assessing landslide hazard. *Earth Surface Processes and Landforms*, 28(10), 1125–1142. https://doi.org/10.1002/esp.545

Carrara, A., Cardinali, M., Detti, R., Guzzetti, F., Pasqui, V., & Reichenbach, P. (1991). GIS techniques and statistical models in evaluating landslide hazard. *Earth Surface Processes and Landforms*, 16(5), 427–445. https://doi.org/10.1002/esp.3290160505

Carrara, A., Crosta, G., & Frattini, P. (2008). Comparing models of debris-flow susceptibility in the alpine environment. *Geomorphology*, 94(3–4), 353–378. https://doi.org/10.1016/j.geomorph.2006.10.033

Castro-Camilo, D., Lombardo, L., Mai, P. M., Dou, J., & Huser, R. (2017). Handling high predictor dimensionality in slope-unit-based landslide susceptibility models through LASSO-penalized generalized linear model. *Environmental Modelling & Software*, 97, 145–156. https://doi.org/10.1016/j.envsoft.2017.08.003

Catani, F. (2021). Landslide detection by deep learning of non-nadiral and crowdsourced optical images. *Landslides*, 18, 1025–1044. https://doi.org/10.1007/s10346-020-01513-4

Catani, F., Lagomarsino, D., Segoni, S., & Tofani, V. (2013). Landslide susceptibility estimation by random forests technique: Sensitivity and scaling issues. *Natural Hazards and Earth System Sciences*, 13(11), 2815–2831. https://doi.org/10.5194/nhess-13-2815-2013

Centamore, E., Chiocchini, M., Chiocchini, U., Dramis, F., Giardini, G., Jacobacci, A., Martelli, G., Micarelli, A., & Potetti, M. (1979). *Note illustrative del Foglio 301 "Fabriano" alla scala 1:50.000*. Servizio Geologico d'Italia, Roma. https://www.isprambiente.gov.it/Media/carg/note_illustrative/301_Fabriano.pdf

Che, Z., Purushotham, S., Cho, K., Sontag, D., & Liu, Y. (2018). Recurrent neural networks for multivariate time series with missing values. *Scientific Reports*, 8(1), 6085. https://doi.org/10.1038/s41598-018-24271-9

Chen, C., & Fan, L. (2024). Interpretability of statistical, machine learning, and deep learning models for landslide susceptibility mapping in Three Gorges Reservoir Area. *arXiv preprint*, arXiv:2405.11762. https://doi.org/10.48550/arXiv.2405.11762

Chen, T., Wang, Q., Zhao, Z., Liu, G., Dou, J., & Plaza, A. (2024). LCFSTE: Landslide conditioning factors and Swin Transformer ensemble for landslide susceptibility assessment.

IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. https://doi.org/10.1109/JSTARS.2024.3373029

Chen, W., Xie, X., Wang, J., Pradhan, B., Hong, H., Bui, D. T., Duan, Z., & Ma, J. (2017). A comparative study of logistic model tree, random forest, and classification and regression tree models for spatial prediction of landslide susceptibility. *Catena*, 151, 147–160. https://doi.org/10.1016/j.catena.2016.11.032

Chen, W., Peng, J., Hong, H., Shahabi, H., Pradhan, B., Liu, J., Zhu, A., Pei, X., & Duan, Z. (2018). Landslide susceptibility modelling using GIS-based machine learning techniques for Chongren County, Jiangxi Province, China. *Science of the Total Environment*, 626, 1121–1135. https://doi.org/10.1016/j.scitotenv.2018.01.160

Chester, M., & Horvath, A. (2010). Life-cycle assessment of high-speed rail: The case of California. *Environmental Research Letters*, 5(1), 014003. https://doi.org/10.1088/1748-9326/5/1/014003

Chikalamo, E. E., Mavrouli, O. C., Ettema, J., van Westen, C. J., Muntohar, A. S., & Mustofa, A. (2020). Satellite-derived rainfall thresholds for landslide early warning in Bogowonto Catchment, Central Java, Indonesia. *International Journal of Applied Earth Observation and Geoinformation*, 89, 102093. https://doi.org/10.1016/j.jag.2020.102093

Chicco, D., & Jurman, G. (2020). The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation. *BMC Genomics*, 21, 6. https://doi.org/10.1186/s12864-019-6413-7

Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: Synthetic minority over-sampling technique. *Journal of Artificial Intelligence Research*, 16, 321–357. https://doi.org/10.1613/jair.953

Childs, C. (2004). Interpolating surfaces in ArcGIS Spatial Analyst. *ArcUser*, July–September, 32–35.

Chung, C. J. F., & Fabbri, A. G. (1999). Probabilistic prediction models for landslide hazard mapping. *Photogrammetric Engineering & Remote Sensing*, 65(12), 1389–1399.

Chung, C. J. F., & Fabbri, A. G. (2003). Validation of spatial prediction models for landslide hazard mapping. *Natural Hazards*, 30(3), 451–472. https://doi.org/10.1023/B:NHAS.0000007172.62651.2b

Clerici, A., Perego, S., Tellini, C., & Vescovi, P. (2006). A GIS-based automated procedure for landslide susceptibility mapping by the conditional analysis method: The Baganza Valley case study (Italian Northern Apennines). *Environmental Geology*, 50(7), 941–961. https://doi.org/10.1007/s00254-006-0264-7

Coltorti, M., Farabollini, P., Gentili, B., & Pambianchi, G. (1996). Geomorphological evidence for anti-Apennine faults in the Umbro-Marchean Apennines and in the peri-Adriatic basin, Italy. *Geomorphology*, 15(1), 33–45. https://doi.org/10.1016/0169-555X(95)00117-N

Conti, P., Cornamusini, G., & Carmignani, L. (2020). An outline of the geology of the Northern Apennines (Italy), with geological map at 1:250,000 scale. *Italian Journal of Geosciences*, 139(2), 149–194. https://doi.org/10.3301/IJG.2019.25

Costantini, M., Minati, F., Trillo, F., Ferretti, A., Novali, F., Passera, E., Dehls, J., Larsen, Y., Marinkovic, P., Eineder, M., Brčić, R., Siegmund, R., Kotzerke, P., Probeck, M., Kenyeres, A., Proietti, S., Solari, L., & Andersen, H. S. (2021, July). European Ground Motion Service

(EGMS). In *2021 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)* (pp. 3293–3296). IEEE. https://doi.org/10.1109/IGARSS47720.2021.9553562

Crawford, M. M., Dortch, J. M., Koch, H. J., Killen, A. A., Zhu, J., Zhu, Y., … Haneberg, W. C. (2021). Using landslide-inventory mapping for a combined bagged-trees and logistic-regression approach to determining landslide susceptibility in eastern Kentucky, USA. *Quarterly Journal of Engineering Geology and Hydrogeology*, 54(4), qjegh2020-177. https://doi.org/10.1144/qjegh2020-177

Crosetto, M., Solari, L., Balasis-Levinsen, J., Bateson, L., Casagli, N., Frei, M., Oyen, A., Moldestad, D. A., & Mróz, M. (2021). Deformation monitoring at European scale: The Copernicus Ground Motion Service. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLIII-B3-2021, 141–146. https://doi.org/10.5194/isprs-archives-XLIII-B3-2021-141-2021

Crosta, G. B., & Frattini, P. (2008). Rainfall-induced landslides and debris flows. *Hydrological Processes*, 22(4), 473–477. https://doi.org/10.1002/hyp.6885

Crozier, M. (2005). Multiple-occurrence regional landslide events in New Zealand: Hazard management issues. *Landslides*, 2, 247–256. https://doi.org/10.1007/s10346-005-0019-7

Cruden, D. M., & Varnes, D. J. (1996). Landslide types and processes. In *Transportation Research Board Special Report* (Vol. 247, pp. 36–75). https://doi.org/10.17226/11057

D'Agostino, V., & Marchi, L. (2001). Debris flow magnitude in the Eastern Italian Alps: Data collection and analysis. *Physics and Chemistry of the Earth, Part C*, 26(9), 657–663. https://doi.org/10.1016/S1464-1917(01)00064-2

Dahal, A., & Lombardo, L. (2023). Explainable artificial intelligence in geoscience: A glimpse into the future of landslide susceptibility modeling. *Computers & Geosciences*, 176, 105364. https://doi.org/10.1016/j.cageo.2023.105364

Dahal, A., Tanyaş, H., & Lombardo, L. (2024). Full seismic waveform analysis combined with transformer neural networks improves coseismic landslide prediction. *Communications Earth & Environment*, 5(1), 75. https://doi.org/10.1038/s43247-024-01243-8

Dahal, A., Huser, R., & Lombardo, L. (2024). At the junction between deep learning and statistics of extremes: Formalizing the landslide hazard definition. *Journal of Geophysical Research: Machine Learning and Computation*, 1(3), e2024JH000164. https://doi.org/10.1029/2024JH000164

Dahal, A., & Lombardo, L. (2025). Towards physics-informed neural networks for landslide prediction. *Engineering Geology*, *344*, 107852. https://doi.org/10.1016/j.enggeo.2024.107852

Dai, F. C., Lee, C. F., & Ngai, Y. Y. (2002). Landslide risk assessment and management: An overview. *Engineering Geology*, 64(1), 65–87. https://doi.org/10.1016/S0013-7952(01)00093-X

Dal Seno, N., Evangelista, D., Piccolomini, E., & Berti, M. (2024). Comparative analysis of conventional and machine learning techniques for rainfall threshold evaluation under complex geological conditions. *Landslides*, 21(12), 2893–2911. https://doi.org/10.1007/s10346-024-02336-3

Das, I., Sahoo, S., van Westen, C., Stein, A., & Hack, R. (2010). Landslide susceptibility assessment using logistic regression and its comparison with a rock mass classification system,

along a road section in the northern Himalayas (India). *Geomorphology*, 114(4), 627–637. https://doi.org/10.1016/j.geomorph.2009.09.023

Das, A., & Rad, P. (2020). Opportunities and challenges in explainable artificial intelligence (XAI): A survey. *arXiv preprint*, arXiv:2006.11371.

de Haas, T., & Densmore, A. L. (2019). Debris-flow volume quantile prediction from catchment morphometry. *Geology*, 47(8), 791–794. https://doi.org/10.1130/G45950.1

De Kemp, E. A. (1998). Three-dimensional projection of curvilinear geological features through direction cosine interpolation of structural field observations. *Computers & Geosciences*, 24(3), 269–284. https://doi.org/10.1016/S0098-3004(97)00066-6

Di Napoli, M., Eroglu, C., van den Bout, B., Di Martire, D., Tanyas, H., & Lombardo, L. (2024). Space-time modeling of cascading hazards: Chaining wildfires, rainfall and landslide events through machine learning. *Catena*, 246, 108452. https://doi.org/10.1016/j.catena.2024.108452

Donat, M. G., Lowry, A. L., Alexander, L. V., O'Gorman, P. A., & Maher, N. (2016). More extreme precipitation in the world's dry and wet regions. *Nature Climate Change*, 6(5), 508–513. https://doi.org/10.1038/nclimate2941

Dong, J.-J., Lee, C.-T., Tung, Y.-H., Liu, C.-N., Lin, K.-P., & Lee, J.-F. (2009). The role of sediment budget in understanding debris flow susceptibility. *Earth Surface Processes and Landforms*, 34, 1612–1624. https://doi.org/10.1002/esp.1850

Dou, H. Q., Huang, S. Y., Jian, W. B., & Wang, H. (2023). Landslide susceptibility mapping of mountain roads based on machine learning combined model. *Journal of Mountain Science*, 20(5), 1232–1248. https://doi.org/10.1007/s11629-022-7657-2

Dou, J., Yamagishi, H., Pourghasemi, H. R., Yunus, A. P., Song, X., Xu, Y., & Zhu, Z. (2015). An integrated artificial neural network model for the landslide susceptibility assessment of Osado Island, Japan. *Natural Hazards*, 78, 1749–1776. https://doi.org/10.1007/s11069-015-1799-2

Drummond, C., & Holte, R. C. (2000, August). Explicitly representing expected cost: An alternative to ROC representation. In *Proceedings of the Sixth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 198–207).

Eker, R., & Aydin, A. (2016). Landslide susceptibility assessment of forest roads. *European Journal of Forest Engineering*, 2(2), 54–60.

ESA (European Space Agency). (2024). Copernicus Global Digital Elevation Model. *OpenTopography*. https://doi.org/10.5069/G9028PQB

Fan, X., Scaringi, G., Korup, O., West, A. J., van Westen, C. J., Tanyas, H., Hovius, N., Hales, T. C., Jibson, R. W., Allstadt, K. E., & Zhang, L. (2019). Earthquake-induced chains of geologic hazards: Patterns, mechanisms, and impacts. *Reviews of Geophysics*, 57(2), 421–503. https://doi.org/10.1029/2018RG000626

Fang, Z., Tanyas, H., Gorum, T., Dahal, A., Wang, Y., & Lombardo, L. (2023). Speech-recognition in landslide predictive modelling: A case for a next generation early warning system. *Environmental Modelling & Software*, 170, 105833. https://doi.org/10.1016/j.envsoft.2023.105833

Fang, Z., Wang, Y., van Westen, C., & Lombardo, L. (2024). Space–time landslide susceptibility modeling based on data-driven methods. *Mathematical Geosciences*, 56(6), 1335–1354. https://doi.org/10.1007/s11004-023-10105-6

Fang, Z., Morales, A. B., Wang, Y., & Lombardo, L. (2025). Climate change has increased rainfall-induced landslide damages in central China. *International Journal of Disaster Risk Reduction*, 105320. https://doi.org/10.1016/j.ijdrr.2025.105320

Fatourechi, M., Ward, R. K., Mason, S. G., Huggins, J., Schlögl, A., & Birch, G. E. (2008, December). Comparison of evaluation metrics in classification applications with imbalanced datasets. In *2008 Seventh International Conference on Machine Learning and Applications* (pp. 777–782). IEEE. https://doi.org/10.1109/ICMLA.2008.131

Fawcett, T. (2006). An introduction to ROC analysis. *Pattern Recognition Letters*, 27(8), 861–874. https://doi.org/10.1016/j.patrec.2005.10.010

Fell, R., Corominas, J., Bonnard, C., Cascini, L., Leroi, E., & Savage, W. Z. (2008). Guidelines for landslide susceptibility, hazard and risk zoning for land use planning. *Engineering Geology*, 102(3–4), 85–98. https://doi.org/10.1016/j.enggeo.2008.03.022

Ferrario, M. F., & Livio, F. (2024). Rapid mapping of landslides induced by heavy rainfall in the Emilia-Romagna (Italy) region in May 2023. *Remote Sensing*, 16, 122. https://doi.org/10.3390/rs16010122

Fidan, S., Tanyaş, H., Akbaş, A., Lombardo, L., Petley, D. N., & Görüm, T. (2024). Understanding fatal landslides at global scales: A summary of topographic, climatic, and anthropogenic perspectives. *Natural Hazards*, 120(7), 6437–6455. https://doi.org/10.1007/s11069-024-06487-3

Fleuchaus, P., Blum, P., Wilde, M., Terhorst, B., & Butscher, C. (2021). Retrospective evaluation of landslide susceptibility maps and review of validation practice. *Environmental Earth Sciences*, 80, 1–15. https://doi.org/10.1007/s12665-021-09770-9

Foraci, R., Tesini, M. S., Nanni, S., Antolini, G., & Pavan, V. (2023). L'inquadramento meteo e idrologico degli eventi. *Ecoscienza* (ARPAE Emilia-Romagna), XIV(5), 20–24.

Franke, R. (1982). Scattered data interpolation: Tests of some methods. *Mathematics of Computation*, 38(157), 181–200. https://doi.org/10.1090/S0025-5718-1982-0637296-4

Frattini, P., Crosta, G., & Carrara, A. (2010). Techniques for evaluating the performance of landslide susceptibility models. *Engineering Geology*, 111(1–4), 62–72. https://doi.org/10.1016/j.enggeo.2009.12.004

Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. *Annals of Statistics*, 29(5), 1189–1232. https://doi.org/10.1214/aos/1013203451

Froude, M. J., & Petley, D. N. (2018). Global fatal landslide occurrence from 2004 to 2016. *Natural Hazards and Earth System Sciences*, 18(8), 2161–2181. https://doi.org/10.5194/nhess-18-2161-2018

Fuad, N., Meandad, J., Haque, A., Sultana, R., Anwar, S. B., & Sultana, S. (2024). Landslide vulnerability analysis using frequency ratio (FR) model: A study on Bandarban district, Bangladesh. *arXiv preprint*, arXiv:2407.20239. https://doi.org/10.48550/arXiv.2407.20239

Gaidzik, K., & Ramírez-Herrera, M. T. (2021). The importance of input data on landslide susceptibility mapping. *Scientific Reports*, 11(1), 19334. https://doi.org/10.1038/s41598-021-98830-y

Galli, M., Ardizzone, F., Cardinali, M., Guzzetti, F., & Reichenbach, P. (2008). Comparing landslide inventory maps. *Geomorphology*, 94(3–4), 268–289. https://doi.org/10.1016/j.geomorph.2006.09.023

Gariano, S. L., & Guzzetti, F. (2016). Landslides in a changing climate. *Earth-Science Reviews*, 162, 227–252. https://doi.org/10.1016/j.earscirev.2016.08.011

Gariano, S. L., & Rianna, G. (2025). How will the projected climate change influence rainfall-induced landslides in Europe? A review of modelling approaches. *Landslides*, 1–17. https://doi.org/10.1007/s10346-025-02550-7

Ghojogh, B., & Crowley, M. (2019). The theory behind overfitting, cross validation, regularization, bagging, and boosting: Tutorial. *arXiv preprint*, arXiv:1905.12787. https://doi.org/10.48550/arXiv.1905.12787

Goetz, J. N., Guthrie, R. H., & Brenning, A. (2011). Integrating physical and empirical landslide susceptibility models using generalized additive models. *Geomorphology*, 129(3–4), 376–386. https://doi.org/10.1016/j.geomorph.2011.03.001

Goodchild, M. (2010). Twenty years of progress: GIScience in 2010. *Journal of Spatial Information Science*, 1, 3–20. https://doi.org/10.5311/JOSIS.2010.1.2

Grelle, G., Revellino, P., Donnarumma, A., & Guadagno, F. M. (2011). Bedding control on landslides: A methodological approach for computer-aided mapping analysis. *Natural Hazards and Earth System Sciences*, 11(5), 1395–1409. https://doi.org/10.5194/nhess-11-1395-2011

Gu, T., Duan, P., Wang, M., Li, J., & Zhang, Y. (2024). Effects of non-landslide sampling strategies on machine learning models in landslide susceptibility mapping. *Scientific Reports*, 14(1), 7201. https://doi.org/10.1038/s41598-024-57964-5

Guinau, M., Vilajosana, I., & Vilaplana, J. M. (2007). GIS-based debris flow source and runout susceptibility assessment from DEM data: A case study in NW Nicaragua. *Natural Hazards and Earth System Sciences*, 7(6), 703–716. https://doi.org/10.5194/nhess-7-703-2007

Guo, Z., Tian, B., Zhu, Y., He, J., & Zhang, T. (2024). How do the landslide and non-landslide sampling strategies impact landslide susceptibility assessment?—A catchment-scale case study from China. *Journal of Rock Mechanics and Geotechnical Engineering*, 16(3), 877–894. https://doi.org/10.1016/j.jrmge.2023.07.026

Guo, Z., Zeng, T., Zhang, Y., Yu, W., Wang, L., Guo, Z., & Glade, T. (2025). A novel hybrid model integrating high-resolution remote sensing and stacking ensemble techniques for landslide susceptibility mapping: Application to event-based landslide inventory. *Geomorphology*, 109886. https://doi.org/10.1016/j.geomorph.2025.109886

Gunning, D., & Aha, D. (2019). DARPA's explainable artificial intelligence (XAI) program. *AI Magazine*, 40(2), 44–58. https://doi.org/10.1609/aimag.v40i2.2850

Gunning, D., Stefik, M., Choi, J., Miller, T., Stumpf, S., & Yang, G. Z. (2019). XAI—Explainable artificial intelligence. *Science Robotics*, 4(37), eaay7120. https://doi.org/10.1126/scirobotics.aay7120

Günther, A., Reichenbach, P., Malet, J. P., Van Den Eeckhaut, M., Hervás, J., Dashwood, C., & Guzzetti, F. (2013). Tier-based approaches for landslide susceptibility assessment in Europe. *Landslides*, 10(5), 529–546. https://doi.org/10.1007/s10346-013-0435-z

Gupta, K., & Satyam, N. (2024). Optimizing seismic hazard inputs for co-seismic landslide susceptibility mapping: A probabilistic analysis. *Natural Hazards*, 120(9), 8459–8481. https://doi.org/10.1007/s11069-024-06517-0

Guzzetti, F., Cardinali, M., & Reichenbach, P. (1996). The influence of structural setting and lithology on landslide type and pattern. *Environmental & Engineering Geoscience*, 2(4), 531–555. https://doi.org/10.2113/gseegeosci.2.4.531

Guzzetti, F., Carrara, A., Cardinali, M., & Reichenbach, P. (1999). Landslide hazard evaluation: A review of current techniques and their application in a multi-scale study, Central Italy. *Geomorphology*, 31(1–4), 181–216. https://doi.org/10.1016/S0169-555X(99)00078-1

Guzzetti, F., Galli, M., Reichenbach, P., Ardizzone, F., & Cardinali, M. (2006). Landslide hazard assessment in the Collazzone area, Umbria, Central Italy. *Natural Hazards and Earth System Sciences*, 6(1), 115–131. https://doi.org/10.5194/nhess-6-115-2006

Guzzetti, F., Mondini, A. C., Cardinali, M., Fiorucci, F., Santangelo, M., & Chang, K. T. (2012). Landslide inventory maps: New tools for an old problem. *Earth-Science Reviews*, 112(1–2), 42–66. https://doi.org/10.1016/j.earscirev.2012.02.001

Guzzetti, F., Reichenbach, P., Cardinali, M., Galli, M., & Ardizzone, F. (2005). Probabilistic landslide hazard assessment at the basin scale. *Geomorphology*, 72(1–4), 272–299. https://doi.org/10.1016/j.geomorph.2005.06.002

Hanley, J. A., & McNeil, B. J. (1982). The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology*, 143(1), 29–36. https://doi.org/10.1148/radiology.143.1.7063747

Handwerger, A. L., Fielding, E. J., Sangha, S. S., & Bekaert, D. P. S. (2022). Landslide sensitivity and response to precipitation changes in wet and dry climates. *Geophysical Research Letters*, 49, e2022GL099499. https://doi.org/10.1029/2022GL099499

Hao, L., Rajaneesh, A., van Westen, C., Sajinkumar, K. S., Martha, T. R., Jaiswal, P., & McAdoo, B. G. (2020). Constructing a complete landslide inventory dataset for the 2018 monsoon disaster in Kerala, India, for land-use change analysis. *Earth System Science Data*, 12, 2899–2918. https://doi.org/10.5194/essd-12-2899-2020

Hastie, T., & Tibshirani, R. (1986). Generalized additive models. Statistical science, 1(3), 297-310.

Hastie, T., & Tibshirani, R. (1990). Exploring the nature of covariate effects in the proportional hazards model. *Biometrics*, 46(4), 1005–1016. https://doi.org/10.2307/2532444

He, J., Tanyas, H., Dahal, A., Huang, D., & Lombardo, L. (2025). Spatial prediction of InSAR-derived hillslope velocities via deep learning. *Bulletin of Engineering Geology and the Environment*, 84(3), 1–14. https://doi.org/10.1007/s10064-025-04161-x

Hölbling, D., Eisank, C., Albrecht, F., Vecchiotti, F., Friedl, B., Weinke, E., & Kociu, A. (2017). Comparing manual and semi-automated landslide mapping based on optical satellite images from different sensors. Geosciences, 7(2), 37. https://doi.org/10.3390/geosciences7020037

Huabin, W., Gangjun, L., Weiya, X., & Gonghui, W. (2005). GIS-based landslide hazard assessment: an overview. Progress in Physical geography, 29(4), 548-567. https://doi.org/10.1191/0309133305pp462

Huang, Y., & Zhao, L. (2018). Review on landslide susceptibility mapping using support vector machines. Catena, 165, 520-529. https://doi.org/10.1016/j.catena.2018.03.003

Hughes, K. S., & Schulz, W. (2020). Map depicting susceptibility to landslides triggered by intense rainfall, Puerto Rico (No. 2020-1022). US Geological Survey. https://doi.org/10.3133/ofr20201022

Hungr, O., Leroueil, S., & Picarelli, L. (2014). The Varnes classification of landslide types, an update. *Landslides*, 11(2), 167–194. https://doi.org/10.1007/s10346-013-0436-y

Hussin, H. Y., Zumpano, V., Reichenbach, P., Sterlacchini, S., Micu, M., van Westen, C., & Bălteanu, D. (2016). Different landslide sampling strategies in a grid-based bi-variate statistical susceptibility model. *Geomorphology*, 253, 508–523. https://doi.org/10.1016/j.geomorph.2015.10.030

Istituto Geografico Militare (IGM). (2024). *DBSN – Database di Sintesi Nazionale* [Data set]. https://www.igmi.org/it/dbsn-database-di-sintesi-nazionale

Ioriatti, I., Coviello, V., Berti, M., & Comiti, F. (2023). Study of debris-flow initiation through the analysis of seismic signals. *E3S Web of Conferences*, 415, 03014. https://doi.org/10.1051/e3sconf/202341503014

ISPRA – Istituto Superiore per la Protezione e la Ricerca Ambientale. (2024). *IFFI – Inventario dei Fenomeni Franosi in Italia* [Database]. https://www.progettoiffi.isprambiente.it/

Jaboyedoff, M., Oppikofer, T., Abellan, A., & Pedrazzini, A. (2012). Use of LiDAR in landslide investigations: A review. *Natural Hazards*, 61, 5–28. https://doi.org/10.1007/s11069-010-9634-2

Jacobs, L., Kervyn, M., Reichenbach, P., Rossi, M., Marchesini, I., Alvioli, M., & Dewitte, O. (2020). Regional susceptibility assessments with heterogeneous landslide information: Slope unit- vs. pixel-based approach. *Geomorphology*, 356, 107084. https://doi.org/10.1016/j.geomorph.2020.107084

Jacovi, A. (2023). Trends in explainable AI (XAI) literature. *arXiv preprint*, arXiv:2301.05433.

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An introduction to statistical learning*. New York: Springer. https://doi.org/10.1007/978-1-4614-7138-7

Jebur, M. N., Pradhan, B., & Tehrany, M. S. (2014). Optimization of landslide conditioning factors using very high-resolution airborne laser scanning (LiDAR) data at catchment scale. *Remote Sensing of Environment*, 152, 150–165. https://doi.org/10.1016/j.rse.2014.05.013

Jiang, N., Li, Y., Han, Z., Yang, J., Fu, B., Li, J., & Li, C. (2024). A side-sampling based Linformer model for landslide susceptibility assessment: a case study of the railways in China. Geomatics, Natural Hazards and Risk, 15(1), 2354507. https://doi.org/10.1080/19475705.2024.2354507

Kawagoe, S., Kazama, S., & Sarukkalige, P. R. (2009). Assessment of snowmelt-triggered landslide hazard and risk in Japan. *Cold Regions Science and Technology*, 58(3), 120–129. https://doi.org/10.1016/j.coldregions.2009.05.004

Kavzoglu, T., Colkesen, I., & Sahin, E. K. (2014). Landslide susceptibility mapping using GIS-based multi-criteria decision analysis, support vector machines, and logistic regression. *Landslides*, 11, 425–439. https://doi.org/10.1007/s10346-013-0391-7

Kavzoglu, T., Sahin, E. K., & Colkesen, I. (2015). Selecting optimal conditioning factors in shallow translational landslide susceptibility mapping using genetic algorithm. *Engineering Geology*, 192, 101–112. https://doi.org/10.1016/j.enggeo.2015.04.004

Kavzoglu, T., Colkesen, I., & Sahin, E. K. (2019). Machine learning techniques in landslide susceptibility mapping: A survey and a case study. In *Landslides: Theory, practice and modelling* (pp. 283–301). https://doi.org/10.1007/978-3-319-77377-3_13

Kikstra, J. S., Nicholls, Z. R., Smith, C. J., Lewis, J., Lamboll, R. D., Byers, E., … Riahi, K. (2022). The IPCC Sixth Assessment Report WGIII climate assessment of mitigation pathways: From emissions to global temperatures. *Geoscientific Model Development*, 15(24), 9075–9109. https://doi.org/10.5194/gmd-15-9075-2022

Kim, H., Lee, J. H., Park, H. J., & Heo, J. H. (2021). Assessment of temporal probability for rainfall-induced landslides based on nonstationary extreme value analysis. *Engineering Geology*, 294, 106372. https://doi.org/10.1016/j.enggeo.2021.106372

Kim, H. S., & Ji, Y. (2022). Three-dimensional geotechnical-layer mapping in Seoul using borehole database and deep neural network-based model. *Engineering Geology*, *297*, 106489. https://doi.org/10.1016/j.enggeo.2021.106489

Kingma, D. P., & Ba, J. (2014). Adam: A method for stochastic optimization. *arXiv preprint*, arXiv:1412.6980.

Koukouvelas, I., Litoseliti, A., Nikolakopoulos, K., & Zygouri, V. (2015). Earthquake-triggered rock falls and their role in the development of a rock slope: The case of Skolis Mountain, Greece. *Engineering Geology*, 191, 71–85. https://doi.org/10.1016/j.enggeo.2015.03.010

Kritikos, T., & Davies, T. (2015). Assessment of rainfall-generated shallow landslide/debris-flow susceptibility and runout using a GIS-based approach: Application to western Southern Alps of New Zealand. *Landslides*, 12, 1051–1075. https://doi.org/10.1007/s10346-014-0533-6

Lee, S., & Choi, J. (2004). Landslide susceptibility mapping using GIS and the weight-of-evidence model. *International Journal of Geographical Information Science*, 18(8), 789–814. https://doi.org/10.1080/13658810410001702003

Lee, S., & Sambath, T. (2006). Landslide susceptibility mapping in the Damrei Romel area, Cambodia using frequency ratio and logistic regression models. *Environmental Geology*, 50, 847–855. https://doi.org/10.1007/s00254-006-0256-7

Lei, X., Liu, H., Chen, Z., Li, S., Chen, H., Zeng, S., … Picco, L. (2025). Investigating the landslide susceptibility assessment methods for multi-scale slope units based on SDGSAT-1 and graph neural networks. *International Journal of Digital Earth*, 18(1), 2468913. https://doi.org/10.1080/17538947.2025.2468913

Leshchinsky, B., Olsen, M. J., Mohney, C., O'Banion, M., Bunn, M., Allan, J., & McClung, R. (2019). Quantifying the sensitivity of progressive landslide movements to failure geometry, undercutting processes and hydrological changes. *Journal of Geophysical Research: Earth Surface*, 124(2), 616–638. https://doi.org/10.1029/2018JF004833

Li, J., Zhou, Z., & Ma, W. (2025). Assessment of landslide susceptibility along the Lanzhou-Xinjiang high-speed railway: A case study of Menyuan–Shandanmachang. *Transportation Geotechnics*, 50, 101473. https://doi.org/10.1016/j.trgeo.2024.101473

Li, L., Lan, H., Guo, C., Zhang, Y., Li, Q., & Wu, Y. (2017). A modified frequency ratio method for landslide susceptibility assessment. *Landslides*, 14, 727–741. https://doi.org/10.1007/s10346-016-0771-x

Liao, M., Wen, H., & Yang, L. (2022). Identifying the essential conditioning factors of landslide susceptibility models under different grid resolutions using hybrid machine learning: A case of Wushan and Wuxi counties, China. *Catena*, 217, 106428. https://doi.org/10.1016/j.catena.2022.106428

Lim, J., Santinelli, G., Dahal, A., Vrieling, A., & Lombardo, L. (2024). An ensemble neural network approach for space–time landslide predictive modelling. *International Journal of Applied Earth Observation and Geoinformation*, 132, 104037. https://doi.org/10.1016/j.jag.2024.104037

Lima, P., Steger, S., Glade, T., & Mergili, M. (2023). Conventional data-driven landslide susceptibility models may only tell us half of the story: Potential underestimation of landslide impact areas depending on the modeling design. *Geomorphology*, 430, 108638. https://doi.org/10.1016/j.geomorph.2023.108638

Liu, K., Wang, M., Cao, Y., Zhu, W., & Yang, G. (2018). Susceptibility of existing and planned Chinese railway systems subjected to rainfall-induced multi-hazards. *Transportation Research Part A: Policy and Practice*, 117, 214–226. https://doi.org/10.1016/j.tra.2018.08.030

Liu, S., Wang, L., Zhang, W., He, Y., & Pijush, S. (2023). A comprehensive review of machine-learning-based methods in landslide susceptibility mapping. *Geological Journal*, 58(6), 2283–2301. https://doi.org/10.1002/gj.4666

Liu, X., Shao, S., & Shao, S. (2024). Landslide susceptibility zonation using the analytical hierarchy process (AHP) in the Great Xi'an Region, China. *Scientific Reports*, 14(1), 2941. https://doi.org/10.1038/s41598-024-53630-y

Liu, X., Shao, S., Zhang, C., & Shao, S. (2025). Effect of different mapping units, spatial resolutions, and machine learning algorithms on landslide susceptibility mapping at the township scale. *Environmental Earth Sciences*, 84(5), 138. https://doi.org/10.1007/s12665-025-12148-w

Loche, M., Alvioli, M., Marchesini, I., Bakka, H., & Lombardo, L. (2022). Landslide susceptibility maps of Italy: Lesson learnt from dealing with multiple landslide types and the uneven spatial distribution of the national inventory. *Earth-Science Reviews*, 232, 104125. https://doi.org/10.1016/j.earscirev.2022.104125

Lombardo, L., & Mai, P. M. (2018). Presenting logistic regression-based landslide susceptibility results. *Engineering Geology*, 244, 14–24. https://doi.org/10.1016/j.enggeo.2018.07.019

Lombardo, L., Cama, M., Conoscenti, C., Märker, M., & Rotigliano, E. (2015). Binary logistic regression versus stochastic gradient boosted decision trees in assessing landslide susceptibility for multiple-occurring landslide events: Application to the 2009 storm event in Messina (Sicily, Southern Italy). *Natural Hazards*, 79, 1621–1648. https://doi.org/10.1007/s11069-015-1915-8

Lombardo, L., & Tanyas, H. (2020). Chrono-validation of near-real-time landslide susceptibility models via plug-in statistical simulations. *Engineering Geology*, 278, 105818. https://doi.org/10.1016/j.enggeo.2020.105818

Lombardo, L., & Tanyas, H. (2022). From scenario-based seismic hazard to scenario-based landslide hazard: Fast-forwarding to the future via statistical simulations. *Stochastic Environmental Research and Risk Assessment*, 36(8), 2229–2242. https://doi.org/10.1007/s00477-021-02020-1

Lombardo, L., Opitz, T., Ardizzone, F., Guzzetti, F., & Huser, R. (2020). Space-time landslide predictive modeling. *Earth-Science Reviews*, 209, 103318. https://doi.org/10.1016/j.earscirev.2020.103318

Lou, Y., Caruana, R., Gehrke, J., & Hooker, G. (2013). Accurate intelligible models with pairwise interactions. In *Proceedings of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 623–631). https://doi.org/10.1145/2487575.2487579

Lu, J., He, Y., Zhang, L., Zhang, Q., Gao, B., Chen, H., & Fang, Y. (2024). Ensemble learning landslide susceptibility assessment with optimized non-landslide sample selection. *Geomatics, Natural Hazards and Risk*, 15(1), 2378176. https://doi.org/10.1080/19475705.2024.2378176

Lu, Z. Y., Liu, G. Y., Zhao, X. D., Sun, K., Chen, Y. S., Song, Z. H., … Yang, M. S. (2025). Landslide susceptibility assessment based on an interpretable coupled FR–RF model: A case study of Longyan City, Fujian Province, Southeast China. *China Geology*, 8(2), 281–294. https://doi.org/10.31035/cg2024123

Luetzenburg, G., Svennevig, K., Bjørk, A. A., Keiding, M., & Kroon, A. (2022). A national landslide inventory for Denmark. *Earth System Science Data*, 14, 3157–3165. https://doi.org/10.5194/essd-14-3157-2022

Lundberg, S. (2022). SHAP (SHapley Additive exPlanations). Retrieved October 2022, from https://github.com/slundberg/shap

Lundberg, S. M., & Lee, S. I. (2017). A unified approach to interpreting model predictions. *arXiv preprint*, arXiv:1705.07874.

Ma, P., Chen, L., Yu, C., Zhu, Q., Ding, Y., Wu, Z., Li, H., Tian, C., & Fan, X. (2025). Dynamic landslide susceptibility mapping over last three decades to uncover variations in landslide causation in subtropical urban mountainous areas. *Remote Sensing of Environment*, 326, 114800. https://doi.org/10.1016/j.rse.2025.114800

Ma, S., Shao, X., & Xu, C. (2023). Landslide susceptibility mapping in terms of the slope-unit or raster-unit, which is better? *Journal of Earth Science*, 34(2), 386–397. https://doi.org/10.1007/s12583-021-1407-1

Malamud, B. D., Turcotte, D. L., Guzzetti, F., & Reichenbach, P. (2004). Landslide inventories and their statistical properties. *Earth Surface Processes and Landforms*, 29, 687–711. https://doi.org/10.1002/esp.1064

Marchi, L., & Brochot, S. (2000). Les cônes de déjection torrentiels dans les Alpes françaises: Morphométrie et processus de transport solide torrentiel. *Revue de Géographie Alpine*, 88(3), 23–38. https://doi.org/10.3406/rga.2000.2999

Marchesini, I., Althuwaynee, O., Santangelo, M., Alvioli, M., Cardinali, M., Mergili, M., Reichenbach, P., Peruccacci, S., Balducci, V., Agostino, I., Esposito, R., & Rossi, M. (2024).

National-scale assessment of railways exposure to rapid flow-like landslides. *Engineering Geology*, 332, 107474. https://doi.org/10.1016/j.enggeo.2024.107474

Marín-Rodríguez, N. J., Vega, J., Zanabria, O. B., González-Ruiz, J. D., & Botero, S. (2024). Towards an understanding of landslide risk assessment and its economic losses: A scientometric analysis. *Landslides*, 21(8), 1865–1881. https://doi.org/10.1007/s10346-024-02272-2

Martinello, C., Cappadonia, C., Conoscenti, C., Agnesi, V., & Rotigliano, E. (2021). Optimal slope units partitioning in landslide susceptibility mapping. *Journal of Maps*, 17(3), 152–162. https://doi.org/10.1080/17445647.2021.1916790

Martinello, C., Cappadonia, C., Conoscenti, C., & Rotigliano, E. (2022). Landform classification: A high-performing mapping unit partitioning tool for landslide susceptibility assessment—A test in the Imera River basin (Northern Sicily, Italy). *Landslides*. https://doi.org/10.1007/s10346-022-01870-6

Martinović, K., Gavin, K., & Reale, C. (2016). Development of a landslide susceptibility assessment for a rail network. *Engineering Geology*, 215, 1–9. https://doi.org/10.1016/j.enggeo.2016.10.011

Matthews, B. W. (1975). Comparison of the predicted and observed secondary structure of T4 phage lysozyme. *Biochimica et Biophysica Acta (BBA) – Protein Structure*, 405(2), 442–451. https://doi.org/10.1016/0005-2795(75)90109-9

McBratney, A. B., Santos, M. M., & Minasny, B. (2003). On digital soil mapping. *Geoderma*, *117*(1-2), 3-52. https://doi.org/10.1016/S0016-7061(03)00223-4

McColl, S. T., & Cook, S. J. (2024). A universal size classification system for landslides. *Landslides*, 21, 111–120. https://doi.org/10.1007/s10346-023-02131-6

Meena, S. R., Puliero, S., Bhuyan, K., Floris, M., & Catani, F. (2022). Assessing the importance of conditioning factor selection in landslide susceptibility for the province of Belluno (Veneto, NE Italy). *Natural Hazards and Earth System Sciences*, 22(4), 1395–1417. https://doi.org/10.5194/nhess-22-1395-2022

Meentemeyer, R. K., & Moody, A. (2000). Automated mapping of conformity between topographic and geological surfaces. *Computers & Geosciences*, 26(7), 815–829. https://doi.org/10.1016/S0098-3004(99)00126-2

Menditto, A., Patriarca, M., & Magnusson, B. (2007). Understanding the meaning of accuracy, trueness and precision. *Accreditation and Quality Assurance*, 12(1), 45–47. https://doi.org/10.1007/s00769-006-0191-z

Merghadi, A., Yunus, A. P., Dou, J., Whiteley, J., ThaiPham, B., Bui, D. T., Avtar, R., & Abderrahmane, B. (2020). Machine learning methods for landslide susceptibility studies: A comparative overview of algorithm performance. *Earth-Science Reviews*, 207, 103225. https://doi.org/10.1016/j.earscirev.2020.103225

Mergili, M., Schwarz, L., & Kociu, A. (2019). Combining release and runout in statistical landslide susceptibility modeling. *Landslides*, 16, 2151–2165. https://doi.org/10.1007/s10346-019-01222-7

Meyer, N. K., Schwanghart, W., Korup, O., Romstad, B., & Etzelmüller, B. (2014). Estimating the topographic predictability of debris flows. *Geomorphology*, 207, 114–125. https://doi.org/10.1016/j.geomorph.2013.10.030

Minder, J. R., Roe, G. H., & Montgomery, D. R. (2009). Spatial patterns of rainfall and shallow landslide susceptibility. *Water Resources Research*, 45(4). https://doi.org/10.1029/2008WR007027

Molnar, C. (2019). *Interpretable Machine Learning: A Guide for Making Black Box Models Explainable*. https://christophm.github.io/interpretable-ml-book/

Montgomery, D. R., & Dietrich, W. E. (1994). A physically based model for the topographic control on shallow landsliding. *Water Resources Research*, 30(4), 1153–1171. https://doi.org/10.1029/93WR02979

Monsieurs, E., Dewitte, O., & Demoulin, A. (2019). A susceptibility-based rainfall threshold approach for landslide occurrence. *Natural Hazards and Earth System Sciences*, 19(4), 775–789. https://doi.org/10.5194/nhess-19-775-2019

Moreno, M., Steger, S., Tanyas, H., & Lombardo, L. (2023). Modeling the area of co-seismic landslides via data-driven models: The Kaikōura example. *Engineering Geology*, 320, 107121. https://doi.org/10.1016/j.enggeo.2023.107121

Moreno, M., Lombardo, L., Crespi, A., Zellner, P. J., Mair, V., Pittore, M., van Westen, C. J., & Steger, S. (2024). Space-time data-driven modeling of precipitation-induced shallow landslides in South Tyrol, Italy. *Science of the Total Environment*, 912, 169166. https://doi.org/10.1016/j.scitotenv.2023.169166

Mueting, A., Bookhagen, B., & Strecker, M. R. (2021). Identification of debris-flow channels using high-resolution topographic data: A case study in the Quebrada del Toro, NW Argentina. *Journal of Geophysical Research: Earth Surface*, 126, e2021JF006330. https://doi.org/10.1029/2021JF006330

Myhre, G., Alterskjær, K., Stjern, C. W., Hodnebrog, Ø., Marelle, L., Samset, B. H., Sillman, J., Schaller, N., Fisher, E., Schulz, M., & Stohl, A. (2019). Frequency of extreme precipitation increases extensively with event rareness under global warming. *Scientific Reports*, 9(1), 16063. https://doi.org/10.1038/s41598-019-52277-4

Nair, V., & Hinton, G. E. (2010). Rectified linear units improve restricted Boltzmann machines. In *Proceedings of the 27th International Conference on Machine Learning (ICML-10)* (pp. 807–814).

Nam, K., Kim, J., & Chae, B. G. (2024). Exploring class imbalance with under-sampling, over-sampling, and hybrid sampling based on Mahalanobis distance for landslide susceptibility assessment: A case study of the 2018 Iburi earthquake-induced landslides in Hokkaido, Japan. *Geosciences Journal*, 28(1), 71–94. https://doi.org/10.1007/s12303-023-0033-6

Nava, L., Carraro, E., Reyes-Carmona, C., Puliero, S., Bhuyan, K., Rosi, A., Monserrat, O., Floris, M., Meena, S. R., Galve, J. P., & Catani, F. (2023). Landslide displacement forecasting using deep learning and monitoring data across selected sites. *Landslides*, 20(10), 2111–2129. https://doi.org/10.1007/s10346-023-02104-9

Neuland, H. (1976). A prediction model of landslips. *Catena*, 3(2), 215–230. https://doi.org/10.1016/0341-8162(76)90011-4

Nocentini, N., Rosi, A., Piciullo, L., Liu, Z., Segoni, S., & Fanti, R. (2024). Regional-scale spatiotemporal landslide probability assessment through machine learning and potential applications for operational warning systems: A case study in Kvam (Norway). *Landslides*, 21(10), 2369–2387. https://doi.org/10.1007/s10346-024-02287-9

Notti, D., Cignetti, M., Godone, D., Cardone, D., & Giordan, D. (2024). The unsupervised shallow landslide rapid mapping: PANDA method applied to severe rainfalls in northeastern Apennine (Italy). *International Journal of Applied Earth Observation and Geoinformation*, 129, 103806. https://doi.org/10.1016/j.jag.2024.103806

Oliveira, S. C., Zêzere, J. L., Guillard-Gonçalves, C., Garcia, R. A., & Pereira, S. (2017). Integration of landslide susceptibility maps for land use planning and civil protection emergency management. In *Advancing Culture of Living with Landslides: Volume 1 ISDR-ICL Sendai Partnerships 2015–2025* (pp. 543–553). Springer. https://doi.org/10.1007/978-3-319-59469-9_49

OpenStreetMap Foundation (OSMF). (2024). *OpenStreetMap* [Data set]. https://www.openstreetmap.org

Opitz, T., Bakka, H., Huser, R., & Lombardo, L. (2022). High-resolution Bayesian mapping of landslide hazard with unobserved trigger event. *The Annals of Applied Statistics*, 16(3), 1653–1675. https://doi.org/10.1214/21-AOAS1561

Ori, G. G., Serafini, G., Visentin, C., Ricci Lucchi, F., Casnedi, R., Colalongo, M. L., & Mosna, S. (1991). *The Pliocene–Pleistocene Adriatic Foredeep (Marche and Abruzzo, Italy): An integrated approach to surface and subsurface geology*.

Ozturk, U., Saito, H., Matsushi, Y., Crisologo, I., & Schwanghart, W. (2021). Can global rainfall estimates (satellite and reanalysis) aid landslide hindcasting? *Landslides*, 18(9), 3119–3133. https://doi.org/10.1007/s10346-021-01689-3

Ozdemir, A., & Altural, T. (2013). A comparative study of frequency ratio, weights of evidence and logistic regression methods for landslide susceptibility mapping: Sultan Mountains, SW Turkey. *Journal of Asian Earth Sciences*, 64, 180–197. https://doi.org/10.1016/j.jseaes.2012.12.014

Pacheco Quevedo, R., Velastegui-Montoya, A., Montalván-Burbano, N., Morante-Carballo, F., Korup, O., & Daleles Rennó, C. (2023). Land use and land cover as a conditioning factor in landslide susceptibility: A literature review. *Landslides*, 20(5), 967–982. https://doi.org/10.1007/s10346-022-02020-4

Parker, R. N., Hales, T. C., Mudd, S. M., Grieve, S. W., & Constantine, J. A. (2016). Colluvium supply in humid regions limits the frequency of storm-triggered landslides. *Scientific Reports*, 6(1), 34438. https://doi.org/10.1038/srep34438

Pasang, S., & Kubíček, P. (2020). Landslide susceptibility mapping using statistical methods along the Asian Highway, Bhutan. *Geosciences*, 10(11), 430. https://doi.org/10.3390/geosciences10110430

Patané, G., Bortolotti, T., Yordanov, V., Biagi, L. G. A., Brovelli, M. A., Truong, X. Q., & Vantini, S. (2025). An interpretable and transferable model for shallow landslides detachment combining spatial Poisson point processes and generalized additive models. *Stochastic Environmental Research and Risk Assessment*, 39(4), 1723–1740. https://doi.org/10.1007/s00477-025-02943-z

Peiró, Y., Volpe, E., Ciabatta, L., & Cattoni, E. (2024). High-resolution precipitation and soil moisture data integration for landslide susceptibility mapping. *Geosciences*, 14(12), 330. https://doi.org/10.3390/geosciences14120330

Pellicani, R., Argentiero, I., & Spilotro, G. (2017). GIS-based predictive models for regional-scale landslide susceptibility assessment and risk mapping along road corridors. *Geomatics, Natural Hazards and Risk*, 8(2), 1012–1033. https://doi.org/10.1080/19475705.2017.1292411

Peng, L., Niu, R., Huang, B., Wu, X., Zhao, Y., & Ye, R. (2014). Landslide susceptibility mapping based on rough set theory and support vector machines: A case of the Three Gorges area, China. *Geomorphology*, 204, 287–301. https://doi.org/10.1016/j.geomorph.2013.08.013

Petschko, H., Brenning, A., Bell, R., Goetz, J., & Glade, T. (2014). Assessing the quality of landslide susceptibility maps—Case study Lower Austria. *Natural Hazards and Earth System Sciences*, 14(1), 95–118. https://doi.org/10.5194/nhess-14-95-2014

Piacentini, D., Troiani, F., Daniele, G., & Pizziolo, M. (2018). Historical geospatial database for landslide analysis: The Catalogue of Landslide Occurrences in the Emilia-Romagna Region (CLOCkER). *Landslides*, 15, 811–822. https://doi.org/10.1007/s103

Quinn, P. E., Hutchinson, D. J., Diederichs, M. S., & Rowe, R. K. (2010). Regional-scale landslide susceptibility mapping using the weights of evidence method: An example applied to linear infrastructure. Canadian Geotechnical Journal, 47(8), 905–927. https://doi.org/10.1139/T10-019

Radbruch-Hall, D. H. (1977). The systematic evaluation of landslide incidence and susceptibility in the United States. *Bulletin of the International Association of Engineering Geology*, 16, 82–86. https://doi.org/10.1007/BF02591456

Rai, R., & Sahu, C. K. (2020). Driven by data or derived through physics? a review of hybrid physics guided machine learning techniques with cyber-physical system (cps) focus. *IEEe Access*, *8*, 71050-71073. https://doi.org/10.1109/ACCESS.2020.2987324

Regmi, N. R., Giardino, J. R., McDonald, E. V., & Vitek, J. D. (2014). A comparison of logistic regression-based models of susceptibility to landslides in western Colorado, USA. *Landslides*, 11, 247–262. https://doi.org/10.1007/s10346-012-0380-2

Reichenbach, P., Rossi, M., Malamud, B. D., Mihir, M., & Guzzetti, F. (2018). A review of statistically-based landslide susceptibility models. *Earth-Science Reviews*, 180, 60–91. https://doi.org/10.1016/j.earscirev.2018.03.001

Rényi, A. (2007). *Probability Theory*. Courier Corporation. https://doi.org/10.1007/978-94-010-1795-4

Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). "Why should I trust you?" Explaining the predictions of any classifier. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 1135–1144). https://doi.org/10.1145/2939672.2939778

Ricci-Lucchi, F. (1975). Depositional cycles in two turbidite formations of northern Apennines. *Journal of Sedimentary Research*, 45(1), 3–43. https://doi.org/10.1306/212F6CB7-2B24-11D7-8648000102C1865D

Roccati, A., Paliaga, G., Luino, F., Faccini, F., & Turconi, L. (2021). GIS-based landslide susceptibility mapping for land use planning and risk assessment. *Land*, 10(2), 162. https://doi.org/10.3390/land10020162

Rosi, A., Frodella, W., Nocentini, N., Caleca, F., Havenith, H. B., Strom, A., Saidov, M., Bimurzaev, G. A., & Tofani, V. (2018). The new landslide inventory of Tuscany (Italy) updated

with PS-InSAR: Geomorphological features and landslide distribution. *Landslides*, 15, 5–19. https://doi.org/10.1007/s10346-017-0861-4

Rosi, A., Frodella, W., Nocentini, N., … & Tofani, V. (2023). Comprehensive landslide susceptibility map of Central Asia. *Natural Hazards and Earth System Sciences*, 23(6), 2229–2250. https://doi.org/10.5194/nhess-23-2229-2023

Rossi, F., Witt, A., Guzzetti, F., Malamud, B. D., & Peruccacci, S. (2012). Analysis of historical landslide time series in the Emilia-Romagna region, northern Italy. *Earth Surface Processes and Landforms*, 35, 1123–1137. https://doi.org/10.1002/esp.1858

Rossi, M., Guzzetti, F., Reichenbach, P., Mondini, A. C., & Peruccacci, S. (2010). Optimal landslide susceptibility zonation based on multiple forecasts. *Geomorphology*, 114(3), 129–142. https://doi.org/10.1016/j.geomorph.2009.06.020

Sabokbar, H. F., Roodposhti, M. S., & Tazik, E. (2014). Landslide susceptibility mapping using geographically-weighted principal component analysis. *Geomorphology*, 226, 15–24. https://doi.org/10.1016/j.geomorph.2014.07.028

Saha, S., Roy, J., Pradhan, B., & Hembram, T. K. (2021). Hybrid ensemble machine learning approaches for landslide susceptibility mapping using different sampling ratios at East Sikkim Himalayan, India. *Advances in Space Research*, *68*(7), 2819-2840. https://doi.org/10.1016/j.asr.2021.05.018

Sahin, E. K. (2022). Comparative analysis of gradient boosting algorithms for landslide susceptibility mapping. *Geocarto International*, 37(9), 2441–2465. https://doi.org/10.1080/10106049.2020.1831623

Sahin, E. K. (2023). Implementation of a free and open-source semi-automatic feature engineering tool in landslide susceptibility mapping using RF, SVM, and XGBoost. *Stochastic Environmental Research and Risk Assessment*, 37(3), 1067–1092. https://doi.org/10.1007/s00477-022-02330-y

Santangelo, M., Marchesini, I., Cardinali, M., Fiorucci, F., Rossi, M., Bucci, F., & Guzzetti, F. (2015). An approach to reduce mapping errors in the production of landslide inventory maps. *Natural Hazards and Earth System Sciences*, 15(9), 2111–2126. https://doi.org/10.5194/nhess-15-2111-2015

Santangelo, M., Marchesini, I., Bucci, F., Cardinali, M., Fiorucci, F., & Guzzetti, F. (2015). A method for the assessment of the influence of bedding on landslide abundance and types. *Landslides*, 12, 295–309. https://doi.org/10.1007/s10346-014-0475-1

Santangelo, M., Althuwaynee, O., Alvioli, M., Bucci, F., Cardinali, M., Fiorucci, F., & Guzzetti, F. (2023). Inventory of landslides triggered by an extreme rainfall event in Marche-Umbria, Italy, on 15 September 2022. *Scientific Data*, 10, 427. https://doi.org/10.1038/s41597-023-02336-3

Saraswat, D., Bhattacharya, P., Verma, A., Prasad, V. K., Tanwar, S., Sharma, G., Bokoro, P. N., & Sharma, R. (2022). Explainable AI for Healthcare 5.0: Opportunities and challenges. *IEEE Access*, 10, 84486–84517. https://doi.org/10.1109/ACCESS.2022.3197671

Schafer, D. H., & Barkan, C. P. L. (2008). A prediction model for broken rails and an analysis of their economic impact. In *Proceedings of the AREMA Annual Conference*.

Schlögl, M., & Matulla, C. (2018). Potential future exposure of European land transport infrastructure to rainfall-induced landslides throughout the 21st century. *Natural Hazards and Earth System Sciences*, 18(4), 1121–1132. https://doi.org/10.5194/nhess-18-1121-2018

Schlögl, M., Spiekermann, R., & Steger, S. (2025). Towards a holistic assessment of landslide susceptibility models: Insights from the Central Eastern Alps. *Environmental Earth Sciences*, 84(4), 113. https://doi.org/10.1007/s12665-024-12041-y

Schimmel, A., & Berti, M. (2023). MAMODIS: A low-cost monitoring system for debris flows based on infrasound and seismic signals. *E3S Web of Conferences*, 415, 03025. https://doi.org/10.1051/e3sconf/202341503025

Schuster, G. (2024). *Machine Learning Methods in Geoscience*. Society of Exploration Geophysicists.

Sciarra, M., Coco, L., & Urbano, T. (2017). Assessment and validation of GIS-based landslide susceptibility maps: A case study from Feltrino stream basin (Central Italy). *Bulletin of Engineering Geology and the Environment*, 76(2), 437–456. https://doi.org/10.1007/s10064-016-0954-7

Segoni, S., Rosi, A., Lagomarsino, D., Fanti, R., & Casagli, N. (2016). The impact of rainfall time series with different length in a landslide warning system, in the framework of changing precipitation trends. *Geoenvironmental Disasters*, 3, Article 1. https://doi.org/10.1186/s40677-016-0057-6

Segoni, S., Piciullo, L., & Gariano, S. L. (2018). A review of the recent literature on rainfall thresholds for landslide occurrence. *Landslides*, 15(8), 1483–1501. https://doi.org/10.1007/s10346-018-0966-4

Segoni, S., Nocentini, N., Barbadori, F., Medici, C., Gatto, A., Rosi, A., & Casagli, N. (2025). A prototype national-scale landslide nowcasting system for Italy combining rainfall thresholds and risk indicators. *Landslides*, 1–26. https://doi.org/10.1007/s10346-024-02452-0

Shao, X., & Xu, C. (2022). Earthquake-induced landslides susceptibility assessment: A review of the state-of-the-art. *Natural Hazards Research*, 2(3), 172–182. https://doi.org/10.1016/j.nhres.2022.03.002

Shapley, L. S. (1953). A value for n-person games. In H. W. Kuhn & A. W. Tucker (Eds.), *Contributions to the Theory of Games* (Vol. 2, pp. 307–317). Princeton University Press. https://doi.org/10.1515/9781400881970-018

Shi, Q., Xu, X., Li, J., Deng, H., Zhang, Q., Tan, D., & He, Y. (2025). Spatiotemporal effect-driven landslide susceptibility mapping at fine scales via deep learning with multidimensional feature fusion and source data adaptation. *Engineering Applications of Artificial Intelligence*, 156, 110924. https://doi.org/10.1016/j.engappai.2025.110924

Simoni, S., Zanotti, F., Bertoldi, G., & Rigon, R. (2008). Modelling the probability of occurrence of shallow landslides and channelized debris flows using GEOtop-FS. *Hydrological Processes*, 22(4), 532–545. https://doi.org/10.1002/hyp.6886

Singh, A., Ashuli, A., Dhiman, N., Dubey, C. S., & Shukla, D. P. (2024). Evaluating causative factors for landslide susceptibility along the Imphal–Jiribam railway corridor (NE India) using a GIS-based statistical approach. *Environmental Science and Pollution Research*, 31(41), 53767–53784. https://doi.org/10.1007/s11356-023-28966-z

Smilkov, D., Thorat, N., Kim, B., Viégas, F., & Wattenberg, M. (2017). SmoothGrad: Removing noise by adding noise. *arXiv preprint*, arXiv:1706.03825.

Soeters, R., & Van Westen, C. J. (1996). Slope instability recognition, analysis and zonation. In A. K. Turner & R. L. Schuster (Eds.), *Landslides: Investigation and Mitigation* (TRB Special Report 247). National Academy Press. https://doi.org/10.17226/11057

Steger, S., Brenning, A., Bell, R., Petschko, H., & Glade, T. (2016). Exploring discrepancies between quantitative validation results and the geomorphic plausibility of statistical landslide susceptibility maps. *Geomorphology*, 262, 8–23. https://doi.org/10.1016/j.geomorph.2016.03.015

Steger, S., Brenning, A., Bell, R., & Glade, T. (2017). The influence of systematically incomplete shallow landslide inventories on statistical susceptibility models and suggestions for improvements. *Landslides*, 14, 1767–1781. https://doi.org/10.1007/s10346-017-0820-0

Steger, S., Scorpio, V., Comiti, F., & Cavalli, M. (2022a). Data-driven modelling of joint debris-flow release susceptibility and connectivity. *Earth Surface Processes and Landforms*, 47(11), 2740–2764. https://doi.org/10.1002/esp.5421

Steger, S., Moreno, M., Crespi, A., Gariano, S. L., Brunetti, M. T., Melillo, M., Peruccacci, S., Marra, F., Kohrs, R., Goetz, J., Mair, V., & Pittore, M. (2023). Deciphering seasonal effects of triggering and preparatory precipitation for improved shallow landslide prediction using generalized additive mixed models. *Natural Hazards and Earth System Sciences*, 23, 1483–1506. https://doi.org/10.5194/nhess-23-1483-2023

Steger, S., Moreno, M., Crespi, A., … Pittore, M. (2024). Adopting the margin of stability for space–time landslide prediction: A data-driven approach for generating spatial dynamic thresholds. *Geoscience Frontiers*, 15(5), 101822. https://doi.org/10.1016/j.gsf.2024.101822

Sujatha, E. R. (2021). An integrated landslide susceptibility model to assess landslides along linear infrastructure for environmental management. *Environmental Earth Sciences*, 80(12), 447. https://doi.org/10.1007/s12665-021-09747-8

Sundararajan, M., Taly, A., & Yan, Q. (2017). Axiomatic attribution for deep networks. In D. Precup & Y. W. Teh (Eds.), *Proceedings of the 34th International Conference on Machine Learning* (Vol. 70, pp. 3319–3328). PMLR. https://proceedings.mlr.press/v70/sundararajan17a.html

Sur, U., Singh, P., & Meena, S. R. (2020). Landslide susceptibility assessment in a Lesser Himalayan road corridor (India) applying fuzzy AHP and Earth-observation data. *Geomatics, Natural Hazards and Risk*, 11(1), 2176–2209. https://doi.org/10.1080/19475705.2020.1836038

Taalab, K., Cheng, T., & Zhang, Y. (2018). Mapping landslide susceptibility and types using Random Forest. *Big Earth Data*, 2(2), 159–178. https://doi.org/10.1080/20964471.2018.1472392

Taha, A. A., & Hanbury, A. (2015). Metrics for evaluating 3D medical image segmentation: Analysis, selection, and tool. *BMC Medical Imaging*, 15, 29. https://doi.org/10.1186/s12880-015-0068-x

Tao, J., & Barros, A. P. (2014). Coupled prediction of flood response and debris-flow initiation during warm- and cold-season events in the Southern Appalachians, USA. *Hydrology and Earth System Sciences*, 18(1), 367–388. https://doi.org/10.5194/hess-18-367-2014

Tanyaş, H., & Lombardo, L. (2019). Variation in landslide-affected area under the control of ground motion and topography. *Engineering Geology*, 260, 105229. https://doi.org/10.1016/j.enggeo.2019.105229

Tarquini, S., Isola, I., Favalli, M., Mazzarini, F., Bisson, M., Pareschi, M. T., & Boschi, E. (2007). TINITALY/01: A new Triangular Irregular Network of Italy. *Annals of Geophysics*, 50(3), 407–425. https://doi.org/10.4401/ag-4424

Tarquini, S., Vinci, S., Favalli, M., Doumaz, F., Fornaciai, A., & Nannipieri, L. (2012). Release of a 10-m resolution DEM for the Italian territory: Comparison with global-coverage DEMs and anaglyph-mode exploration via the web. *Computers & Geosciences*, 38(1), 168–170. https://doi.org/10.1016/j.cageo.2011.04.018

Titti, G., van Westen, C., Borgatti, L., Pasuto, A., & Lombardo, L. (2021). When enough is really enough? On the minimum number of landslides to build reliable susceptibility models. *Geosciences*, 11(11), 469. https://doi.org/10.3390/geosciences11110469

Titti, G., Sarretta, A., Lombardo, L., Crema, S., Pasuto, A., & Borgatti, L. (2022). Mapping susceptibility with open-source tools: A new plugin for QGIS. *Frontiers in Earth Science*, 10, 842425. https://doi.org/10.3389/feart.2022.842425

Trigila, A., Iadanza, C., & Spizzichino, D. (2010). Quality assessment of the Italian Landslide Inventory using GIS processing. *Landslides*, 7(4), 455–470. https://doi.org/10.1007/s10346-010-0213-0

Trigila, A., Iadanza, C., Esposito, C., & Scarascia-Mugnozza, G. (2015). Comparison of Logistic Regression and Random Forests techniques for shallow landslide susceptibility assessment in Giampilieri (NE Sicily, Italy). *Geomorphology*, 249, 119–136. https://doi.org/10.1016/j.geomorph.2015.06.001

Tsangaratos, P., Ilia, I., Hong, H., Chen, W., & Xu, C. (2017). Applying Information Theory and GIS-based quantitative methods to produce landslide susceptibility maps in Nancheng County, China. *Landslides*, 14, 1091–1111. https://doi.org/10.1007/s10346-016-0769-4

Van Den Eeckhaut, M., & Hervás, J. (2012). State of the art of national landslide databases in Europe and their potential for assessing susceptibility, hazard and risk. *Geomorphology*, 139–140, 545–558. https://doi.org/10.1016/j.geomorph.2011.12.006

Van Rijsbergen, C. J. (1979). *Information Retrieval* (2nd ed.). Butterworths.

Van Westen, C. J., Rengers, N., & Soeters, R. (2003). Use of geomorphological information in indirect landslide susceptibility assessment. *Natural Hazards*, 30(3), 399–419. https://doi.org/10.1023/B:NHAZ.0000007097.42735.9e

Van Westen, C. J., Van Asch, T. W., & Soeters, R. (2006). Landslide hazard and risk zonation—Why is it still so difficult? *Bulletin of Engineering Geology and the Environment*, 65, 167–184. https://doi.org/10.1007/s10064-005-0023-0

Van Westen, C. J., Castellanos, E., & Kuriakose, S. L. (2008). Spatial data for landslide susceptibility, hazard, and vulnerability assessment: An overview. *Engineering Geology*, 102(3–4), 112–131. https://doi.org/10.1016/j.enggeo.2008.03.010

Varnes, D. J. (1984). *Landslide Hazard Zonation: A Review of Principles and Practice* (No. 3). UNESCO.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention is all you need. In *Advances in Neural Information Processing Systems* (Vol. 30).

Wang, N., Cheng, W., Marconcini, M., Bachofer, F., Liu, C., Xiong, J., & Lombardo, L. (2022). Space-time susceptibility modeling of hydro-morphological processes at the Chinese national scale. *Engineering Geology*, 301, 106586. https://doi.org/10.1016/j.enggeo.2022.106586

Wang, S., Zhuang, J., Mu, J., Zheng, J., Zhan, J., Wang, J., & Fu, Y. (2022). Evaluation of landslide susceptibility of the Ya'an–Linzhi section of the Sichuan–Tibet Railway based on deep learning. *Environmental Earth Sciences*, 81(9), 250. https://doi.org/10.1007/s12665-022-10375-z

Wang, Y., Wu, X., Zhou, K., Lin, G., Peng, B., & Fang, Z. (2025). Integrating a multi-dimensional deep convolutional neural network with optimized sample selection for landslide susceptibility assessment. *Geo-spatial Information Science*, 1–21. https://doi.org/10.1080/10095020.2024.2443483

Wasowski, J., Keefer, D. K., & Lee, C.-T. (2011). Toward the next generation of research on earthquake-induced landslides: Current issues and future challenges. *Engineering Geology*, 122(1–2), 1–8. https://doi.org/10.1016/j.enggeo.2011.06.001

Wilford, D., Sakals, M., Innes, J., Sidle, R., & Bergerud, W. (2004). Recognition of debris flow, debris flood, and flood hazard through watershed morphometrics. *Landslides*, 1(1), 61–66. https://doi.org/10.1007/s10346-004-0008-2

Wood, S. N. (2017). *Generalized Additive Models: An Introduction with R* (2nd ed.). Chapman & Hall/CRC.

Woodard, J. B., Mirus, B. B., Crawford, M. M., Or, D., Leshchinsky, B. A., Allstadt, K. E., & Wood, N. J. (2023). Mapping landslide susceptibility over large regions with limited data. *Journal of Geophysical Research: Earth Surface*, 128(5), e2022JF006810. https://doi.org/10.1029/2022JF006810

Wu, C.-H., & Chen, S.-C. (2009). Determining landslide susceptibility in Central Taiwan from rainfall and six site factors using the analytical hierarchy process method. *Geomorphology*, 112(3–4), 190–204. https://doi.org/10.1016/j.geomorph.2009.06.002

Xu, K., Zhao, Z., Chen, W., Ma, J., Liu, F., Zhang, Y., & Ren, Z. (2024). Comparative study on landslide susceptibility mapping based on different ratios of training samples and testing samples by using RF and FR-RF models. *Natural Hazards Research*, *4*(1), 62-74. https://doi.org/10.1016/j.nhres.2023.07.004

Yano, A., Shinohara, Y., Tsunetaka, H., Mizuno, H., & Kubota, T. (2019). Distribution of landslides caused by heavy rainfall events and an earthquake in northern Aso Volcano, Japan (1955–2016). *Geomorphology*, 327, 533–541. https://doi.org/10.1016/j.geomorph.2018.11.024

Ye, C., Wu, H., Oguchi, T., Tang, Y., Pei, X., & Wu, Y. (2025). Physically Based and Data-Driven Models for Landslide Susceptibility Assessment: Principles, Applications, and Challenges. *Remote Sensing*, *17*(13), 2280. https://doi.org/10.3390/rs17132280

Yildiz, A., Zhao, H., & Kowalski, J. (2023). Computationally-feasible uncertainty quantification in model-based landslide risk assessment. *Frontiers in Earth Science*, 10, 1032438. https://doi.org/10.3389/feart.2022.1032438

Youssef, K., Shao, K., Moon, S., & Bouchard, L. S. (2022). Landslide susceptibility modeling by interpretable neural network. *arXiv preprint*, arXiv:2201.06837.

Yu, L., Wang, Y., & Pradhan, B. (2024). Enhancing landslide susceptibility mapping by incorporating landslide typology via stacking ensemble machine learning in Three Gorges Reservoir, China. *Geoscience Frontiers*, 15(4), 101802. https://doi.org/10.1016/j.gsf.2024.101802

Yu, X., Wang, Y., Niu, R., & Hu, Y. (2016). A combination of geographically weighted regression, particle swarm optimization, and support vector machine for landslide susceptibility mapping: A case study at Wanzhou in the Three Gorges Area, China. *International Journal of Environmental Research and Public Health*, 13(5), 487. https://doi.org/10.3390/ijerph13050487

Zêzere, J. L. (2002). Landslide susceptibility assessment considering landslide typology: A case study north of Lisbon (Portugal). *Natural Hazards and Earth System Sciences*, 2(1/2), 73–82. https://doi.org/10.5194/nhess-2-73-2002

Zêzere, J. L., Neves, M., Henriques, C. S., Garcia, R., Oliveira, S., & Piedade, A. (2009). Effects of landslide inventories uncertainty on landslide susceptibility modelling. In *Landslide Processes: From Geomorphologic Mapping to Dynamic Modelling* (pp. 81–86).

Zêzere, J. L., Pereira, S., Melo, R., Oliveira, S. C., & Garcia, R. A. (2017). Mapping landslide susceptibility using data-driven methods. *Science of the Total Environment*, 589, 250–267. https://doi.org/10.1016/j.scitotenv.2017.02.188

Zhang, B., Tang, J., Huan, Y., Song, L., Shah, S. Y. A., & Wang, L. (2024). Multi-scale convolutional neural networks for landslide inventory mapping from remote sensing imagery and landslide susceptibility mapping. *Geomatics, Natural Hazards and Risk*, 15(1), 2383309. https://doi.org/10.1080/19475705.2024.2383309

Zhao, Z., Chen, T., Dou, J., Liu, G., & Plaza, A. (2024). Landslide susceptibility mapping considering local–global features based on CNN and transformer. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*. https://doi.org/10.1109/JSTARS.2024.3379350

Zhou, C., Cao, Y., Gan, L., Wang, Y., Motagh, M., Roessner, S., … Yin, K. (2024). A novel framework for landslide displacement prediction using MT-InSAR and machine learning techniques. *Engineering Geology*, 334, 107497. https://doi.org/10.1016/j.enggeo.2024.107497

Zhou, H., Zhang, S., Peng, J., Zhang, S., Li, J., Xiong, H., & Zhang, W. (2021). Informer: Beyond efficient transformer for long sequence time-series forecasting. In *Proceedings of the AAAI conference on artificial intelligence* (Vol. 35, No. 12, pp. 11106-11115). https://doi.org/10.1609/aaai.v35i12.17325

Zhou, M., Yuan, M., Yang, G., & Mei, G. (2024). Risk analysis of road networks under the influence of landslides by considering landslide susceptibility and road vulnerability: A case study. *Natural Hazards Research*, 4(3), 387–400. https://doi.org/10.1016/j.nhres.2023.09.013

Zieher, T., Perzl, F., Rössel, M., Rutzinger, M., Meißl, G., Markart, G., & Geitner, C. (2016). A multi-annual landslide inventory for the assessment of shallow landslide susceptibility—Two test cases in Vorarlberg, Austria. *Geomorphology*, 259, 40–54. https://doi.org/10.1016/j.geomorph.2016.02.008

Zou, K. H., Wells III, W. M., Kikinis, R., & Warfield, S. K. (2004). Three validation metrics for automated probabilistic image segmentation of brain tumours. *Statistics in Medicine*, 23(8), 1259–1282. https://doi.org/10.1002/sim.1723