https://doi.org/10.5194/egusphere-2026-579
Preprint. Discussion started: 2 March 2026

Spatial machine learning modelling reveals that soil indicators and tree type best explain shallow landslide release

Denise Christina Rüther[1], Kristine Flacké Haualand[1], Iris Louisa Johanna Peeters[1,2], Mark Andrew Kusk Gillespie[1,3]

[1]Department of Environmental Sciences, Western Norway University of Applied Sciences, Sogndal, Norway
[2]Vestland County Council, Bergen, Norway
[3]Department of Ecoscience, Aarhus University, Aarhus, Denmark

Correspondence to: Denise C. Rüther (denise.christina.ruther@hvl.no)

Abstract. The exploration of shallow landslide susceptibility is often impaired by incomplete and imperfect landslide inventories, and by over-optimistic performance metrics linked to inadequate models. In this article, we make use of a recently published, systematically mapped event inventory containing 571 shallow landslides triggered in southern Norway and apply a total of 32 gradient boosted decision tree models to rigorously test the effects of (1) a nested vs. simple cross-validation strategy, (2) spatial vs. non-spatial models, (3) four different cross-validation sampling strategies which were applied on (4) full vs. forest-only datasets. Model evaluation shows that the spatial model with small block nested cross-validation is a suitable compromise between modelling the spatial structure of the landslide data adequately while retaining realistic predictive power. The compromise models suggest that soil thickness explains landslide probability partly, while other important explanatory factors like elevation, aspect and bedrock weatherability serve as soil indicators, illustrating a need for improved datasets for soil thickness and heterogeneity. In the forest-only models, tree type explains landslide probability best, driven by greater landslide susceptibility in deciduous forest. Presented results suggest that susceptibility mapping may be improved significantly by considering forest variables and forest-specific threshold values.

1 Introduction

Rainfall-induced shallow landslides pose a significant and escalating threat to societies worldwide. Estimates from the United Nations Office for Disaster Risk Reduction suggest over 500 000 fatalities due to climate-related disasters in the first two decades of the 21st century, almost doubling compared to the last two decades in the 20th century (Mizutori, 2018). Likewise, landslide-related economic losses increased from an estimated US$ 50 billion per year in the 1980s to approximately US$ 200 billion per year in the 2000s (Albano and Sole, 2018). This increase is in part driven by growing exposure of humans and assets to landslides and in part by a climate change driven increase in the landslide hazard itself (Marín-Rodríguez et al., 2024). Shallow landslides can be triggered seismically during earthquakes, anthropogenically due to deforestation and constructions, by additional loading such as rockfalls as well as due to rise in porewater pressure by rainfall and/or snow melt. The majority of shallow landslides are, however, rainfall-triggered (Sim et al., 2022), e.g. close to

https://doi.org/10.5194/egusphere-2026-579
Preprint. Discussion started: 2 March 2026

90% of fatal, recent shallow landslides in Europe and Bangladesh were rainfall-triggered (Haque et al., 2016; Sultana, 2020). Therefore, a continued acceleration of landslide risk is expected in the future as increased frequency and magnitude of heavy precipitation over mid-latitude and tropical mountainous areas will likely lead to more rainfall-induced shallow landslide events (Crozier, 2010; Jakob and Owen, 2021; Stoffel et al., 2014; Gariano and Guzzetti, 2022). This calls for an urgent need to improve shallow landslide susceptibility assessments.

Studies of shallow landslide susceptibility are commonly differentiated into knowledge-driven, physically based and data-driven methods (Corominas et al., 2014). The current Norwegian susceptibility map can be classified as a knowledge-driven approach because landslide inventories are only used for validation, while shallow landslide starting points are determined based on expert considerations of suitable threshold values for slope, planar curvature and upstream catchment area, and subsequent runout modelling (Fischer et al., 2012; Pullarello et al., 2025). Physically based methods simulate the processes on single slopes to small regions, incorporating hydrological processes, soil mechanics and topography (Sorbino et al., 2010). Physically based models can be placed on a continuum between steady-state models, guilty of oversimplification and sophisticated numerical models, which are computationally expensive, and depend on detailed, often unavailable hydrometeorological and soil data (Sachinthaka et al., 2025; Ye et al., 2025). Data-driven approaches use landslide inventories both as input and for validation of statistical and machine learning models. They build on the assumption that combinations of topographic, geological and meteorological factors that have triggered shallow landslides in the past can be used to make quantitative predictions of future susceptibility to shallow landslides (Corominas et al. 2014). While data-driven approaches can be regarded as the standard method for regional or even national susceptibility mapping, they do not account for the complex physical mechanisms behind shallow landslide release, and tend to suffer from limited and biased datasets (Lima et al., 2022). More recently, attempts have also been made to combine data-driven and physically based approaches, potentially overcoming weaknesses of both methods (Han and Semnani, 2025; Ye et al. 2025).

Among the data-driven models, machine learning methods have received exponentially growing attention over the last 10 years and often outperform more classical statistical approaches as they are better suited to tackle non-linear relationships, high dimensionality and collinearity among predictors (Ye et al. 2025). In particular, gradient boosting decision tree methods have received a lot of attention due to their predictive performance (Cao et al., 2020; Wei et al., 2022), and many recent studies report almost perfect levels of landslide prediction accuracy (Choubin et al., 2025; Dai and Huang, 2025; Liu et al., 2021). However, in line with earlier work (Steger et al., 2021; Lima et al., 2021), we suggest that these are often overly optimistic reports which are almost certainly the result of model overfitting, and that the popularity and perceived ease of use of these methods require a robust set of standardised best practice protocols (Reichenbach et al., 2018; Lombardo and Mai, 2018). Machine learning methods for classification of binary outcomes, such as the occurrence of shallow landslides, are powerful algorithms but predictive accuracy and the importance of predictors are susceptible to a number of choices made by the researcher (Hastie, 2009; Probst et al., 2019). So-called hyperparameters control aspects such as model complexity, regularisation strength (the degree to which a model overfitting is constrained) and decision tree structure (the depth and complexity of decision trees used in the model). Hyperparameters such as learning rate, number of trees and maximum tree

https://doi.org/10.5194/egusphere-2026-579
Preprint. Discussion started: 2 March 2026

depth can be either chosen by default, set by the researcher, or selected following a wide range of available data-dependent optimisation procedures (Probst et al. 2019). Typically, the optimisation process involves a resampling technique such as cross-validation (CV; Bischl et al., 2012), but the CV process is itself sensitive to choices in construction of the cross-validation folds. For example, in random cross-validation, the data are randomly assigned to k-folds and the optimisation process repeatedly selects a different fold as a test set, leaving the remainder as training set and evaluates model performance for each hyperparameter combination across multiple train-test splits.

However, for some applications, an alternative method of fold assignment may be more appropriate. In some ecological applications, where the data are spatially dependent, spatial cross-validation approaches have been found to be more appropriate and less likely to lead to overly optimistic results (Linnenbrink et al., 2023; Meyer and Pebesma, 2022). In spatial CV, the folds are assigned to spatial clusters or blocks, the size and clustering of which can also be set, and cross-validation proceeds in the same way while accounting for some of the spatial autocorrelation in input data (Meyer et al., 2024). As shallow landslides typically occur in spatial clusters and inventories often exhibit a fractal structure (Ghosh et al., 2012; Liu et al., 2019; Rouai and Jaaidi, 2003), it is highly likely that landslide data exhibit spatial autocorrelation, suggesting that spatial CV is an appropriate option for both hyperparameter optimisation and model evaluation. The importance of spatial CV in landslide susceptibility studies has been advocated elsewhere (Brenning, 2005; Goetz et al., 2015; Knevels et al., 2023; Petschko et al., 2014), and it has also been shown that failure to account for spatial autocorrelation during both hyperparameter optimisation and model evaluation can lead to overfitting and inflated performance estimates that do not generalise to new spatial locations (Steger et al., 2016). Nonetheless, the majority of recent shallow landslide susceptibility studies appear to apply random or no CV methods (e.g. Choubin et al., 2025; Li et al., 2025; Utthasini et al., 2026; Wang and Deng, 2025; Qin et al., 2025) and potentially fail to consider the inherent spatial relationship of shallow landslide inventories.

Another source of potential over-optimism in landslide susceptibility models is a lack of model fitting procedures that incorporate spatial structure in response data. Apart from a recent study which incorporates spatial structure in a Generalised Additive Modelling approach (Knevels et al., 2023), the inherent spatial structure in landslide datasets is rarely modelled directly. In some cases, spatial structure may be sufficiently addressed by spatial CV, or directly in models by the inclusion of spatially explicit variables such as initiation point coordinates or distances to land features such as rivers, roads or settlements (e.g. Choubin et al., 2025; Qin et al., 2025; Steger et al., 2016; Steger et al., 2021). However, landslide susceptibility studies rarely test for spatial autocorrelation in model residuals to ensure that the spatial structure has been accounted for. If spatial structure remains in model predictions, the results may be over-optimistic, suffer from low generalisability, or both (Mets et al., 2017; Roberts et al. 2017). In this paper, we address the above issues by utilising the Gaussian Process Gradient Boosting algorithm, which is an extreme gradient boosting machine learning method that allows for the addition of random effects in the model (Sigrist, 2022). Utilising this approach, together with nested spatial cross-validation and testing for residual spatial autocorrelation, *our first objective (O1)* was to demonstrate the potential to address overfitting in landslide modelling with machine learning methods.

https://doi.org/10.5194/egusphere-2026-579
Preprint. Discussion started: 2 March 2026

The modelling procedure is applied to a landslide dataset related to an extreme rainstorm event in Norway in 2023. As the shallow landslides all occurred within a period of two to three days, the dataset is not subject to another form of uncertainty, temporal heterogeneity, which tends to affect landslide inventory data (Reichenbach et al. 2018; Ye et al. 2025). Being based on an event-inventory and including event-specific hydrometeorological variables, results from this study offer limited generalisability and a shallow landslide susceptibility map has therefore not been produced. Nevertheless, the dataset provides a unique opportunity to explore shallow landslide explanatory factors based on an inventory with a consistent and systematic mapping routine (Rüther et al., 2024), thus minimising geographic bias. Hence, *our second objective (O2)* was to analyse morphological, geological and hydrometeorological variables relevant to shallow landslide release based on gradient boosted models on the full dataset.

Moreover, the impact of forest cover and characteristics were studied based on additional gradient boosted models applied to a sub-dataset consisting of shallow landslide and control points in the forest. While the effects of vegetation and trees on slope stability are relatively well studied on the scale of single trees and tree stands (Kim et al., 2017; Schwarz et al., 2010a; Schwarz et al., 2010b), and on the scale of forested slopes (Jiang et al., 2023; Kim et al., 2020; Medina et al., 2021), examples of data-driven shallow landslide susceptibility studies which consider vegetational or forest variables beyond landuse are rare (Moos et al., 2016). Therefore, by repeating our model routine based on a sub-dataset from forested sites, *our third objective (O3)* was to demonstrate the important role of forest characteristics in influencing landslide susceptibility.

2 'Hans' event and shallow landslide inventory

On 7 to 9 August 2023 southern Norway was hit by the extreme rainstorm event 'Hans', a powerful storm which originated as two low-pressure systems that merged over the European continent and subsequently approached Norway from the southeast. This is unusual as most low-pressure systems affecting southern Norway have a westerly origin. Consequently, large areas in southeastern Norway (20 600 km^2) which are normally situated in the rain shadow, received 70-100 mm of rain in 24 hours, corresponding to rainfall with a 100-year return period (Granerød et al., 2023). While the total accumulated precipitation during the 'Hans' storm exceeded 110 mm in southeastern as well as some western areas (Fig. 1a), 'Hans' precipitation during these three days surpassed 80% of normal precipitation for the entire month of August in large parts of eastern Norway (Fig. 1b). The storm caused widespread flooding and numerous shallow landslides with severe impacts on infrastructure, damage to private property and disruption to major transport routes. No lives were lost, in part due to well-functioning early warning and evacuation of around 4600 individuals (DSB, 2024), and in part due to luck. The 'Hans' storm is amongst the most expensive natural disasters in Norway with direct cost estimated to 0.7 billion US$ (Meld. St. 27, 2023–2024).

Shallow landslides triggered during this storm were systematically mapped based on the dNDVI (differential Normalised Difference Vegetation Index) calculated for composite pre- and post-event satellite images with 10 m spatial resolution (Rüther et al. 2024). The mapping was validated with all available data sources including higher resolution satellite images,

https://doi.org/10.5194/egusphere-2026-579
Preprint. Discussion started: 2 March 2026

post-event orthophotos, drone footage and field observations. Compared to the Norwegian crowd-sourced landslide database (NSDB), the systematic mapping approach resulted in a 246% increase in the number of detected shallow landslides (from 263 to 648) and a 35% reduction of the spatial bias towards the road network for the occurrence of shallow landslides within 50 m of a road (from 61.6 to 25.9%; Rüther et al. 2024). The 'Hans' shallow landslide inventory allows for the investigation of explanatory factors of a rare event compared to results based on general or non-event-specific inventories.

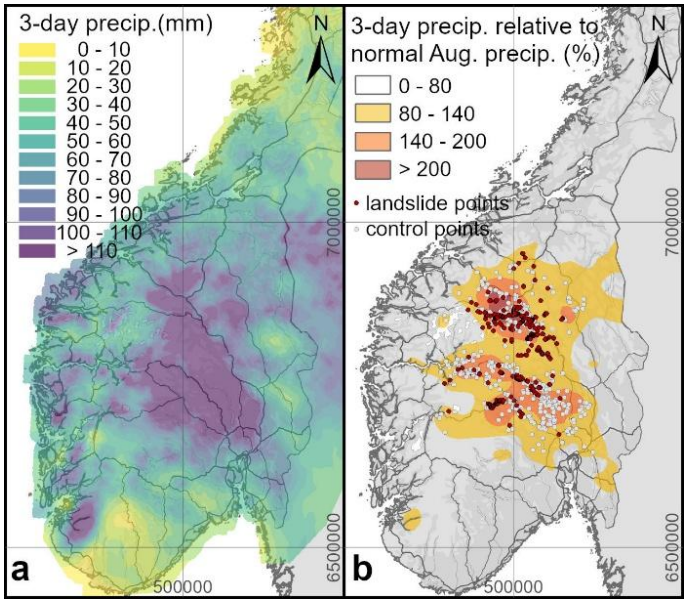


Figure 1 a) 3-day cumulative precipitation from 7 to 10 august, 2023 (6 UTC or 8 am local time) b) 3-day cumulative precipitation relative to normal precipitation for the entire month of August (classified according to its influence on landslide probability), and position of landslide and control points. Topographic map: © Kartverket, SeNorge.

3 Data preparation

The data preparation was performed in two separate Python codes: the main code was run on a supercomputer which was beneficial for the computational expensive raster operations and a second code which was run on a local computer where some shapefile-related Python packages worked better (Haualand, 2025). Data preparation was further supported by additional analyses and visualisation in ArcGIS Pro on a local computer.

3.1 Landslide points

The statistical and machine learning supported analyses presented in this study are based on a refined version of the 'Hans'-Landslide-dataset presented in Rüther et al. (2024). The original polygon dataset is mapped based on seamless, cloud-free greenest-pixel Sentinel-2 composites (Lindsay et al., 2022) together with additional data sources locally, including post-event orthophoto images, the NSDB database and field work by the main author. The published version of the dataset

https://doi.org/10.5194/egusphere-2026-579
Preprint. Discussion started: 2 March 2026

consists of 648 mapped polygons which were differentiated into 56 debris floods, 388 debris flows and 204 debris slides. The refinement of the dataset included a re-evaluation of the placement of starting points, and a revisited differentiation into debris flows and debris floods assisted by all available data including slope angle. Inferred debris floods were then excluded from the point dataset as they initiate by continued riverbed and riverbank erosion. Many studies choose either the centroid points of mapped shallow landslide polygons (e.g. Ageenko et al., 2022; Arabameri et al., 2022) or sample distributed points throughout the landslide scars (e.g. Quevedo et al., 2022). Predictive capability of different susceptibility models has been shown to vary with different sampling strategies with distributed landslide scar points scoring highest (Dou et al., 2020). However, with the aim to establish causal relationships, it is more obvious to investigate the influencing factors at the initiation points rather than at landslide centroids or distributed landslide scar points, which happen to lie downstream of a given release area. The shallow landslide point dataset (debris slides and debris flows, including debris avalanches) comprises 571 starting points released during the 'Hans' storm.

3.2 Control points

In line with other studies in the field, the machine learning models are based on a similar amount of landslide and non-landslide or control points (Huang and Zhao, 2018; Wu et al., 2020), and 571 control points have been sampled based on the following workflow:

1. Selection of a large random sample of points (571 000) which are within the landslide point attribute ranges for elevation, slope and relative precipitation.
2. Normalisation of the three attributes relative to the maximum observed value for the landslide points. This means that all landslide and control points can be seen in a three-dimensional (3D) attribute space for elevation, slope and relative precipitation with values ranging from 0 to 1.
3. Find the 5 710 closest non-landslide points in the normalised 3D attribute space and randomly select 571. Control points are hence relatively close to the elevation, slope and relative precipitation of a corresponding landslide point, but not necessarily the closest. Choosing only the closest point in the 3D attribute space will create control points that will tend to be geographically close to the landslide point.

3.3 Attributes considered as explanatory factors

The current national DEM (Digital Elevation Model) is predominantly based on lidar-data, and available in 1-meter and 10-meter resolution. However, 54.5% of Norway's landmasses were laser scanned with relatively low resolution (2 pt/m^2) and 11.4% are based on image-matching (kartverket.no). Therefore, the elevation accuracy in dense forest can be relatively poor and above the treeline, inaccuracies may include severe elevation artefacts due to image-matching of glaciated and snow-covered areas. In addition, the 1-meter DEM consists of one-pixel-wide holes at most project boundaries. This study is therefore based on a 5-meter DEM produced by the Norwegian Geological Survey (NGU) where holes in the national DEM have been filled while the other weaknesses mentioned above have not been solved (Pullarello, 2024).

A series of *morphological attributes* are considered:

a. Elevation (m a.s.l.)

https://doi.org/10.5194/egusphere-2026-579
Preprint. Discussion started: 2 March 2026

b. Slope (degrees)
c. Planar Curvature
d. Profile Curvature
e. Northerness (cosine of aspect)
f. Easterness (sine of aspect)
g. Landforms (based on the Topographic Position Index)

Elevation is the 5-meter DEM provided by NGU (Pullarello, 2024) while all other morphological attributes are derived from the same DEM. To account for the circular nature of aspect values, northerness and easterness are considered in our machine learning models. The profile curvature describes the radius of curvature parallel to the direction of the steepest slope. Positive profile curvature values indicate that the surface is concave downslope leading to flow acceleration while negative values correspond to convex downslope surfaces and decelerating flow. The planar curvature describes the radius of curvature perpendicular to the direction of the maximum slope. Positive planar curvature values indicate that the surface is laterally convex causing divergent flow while negative values correspond to laterally concave surfaces and convergent flow. The Terrain Ruggedness Index (TRI) measures the topographic variation by calculating the elevation difference between each cell and its eight neighbours (Riley et al., 1999). While Gradient Boosting is more robust to multicollinearity than linear models, highly correlated predictors can still lead to instability and since TRI correlated strongly with slope, it was omitted from the final models. Topographic position has been analysed and landforms classified following the procedure suggested by Weiss (2001). The Topographic Position Index (TPI) compares the elevation of each cell in a DEM to the mean elevation of an annulus-shaped neighbourhood around that cell. TPI was calculated for small-scale neighbourhoods (TPI300 considering cells within 150- and 300-meters radius) and for large-scale neighbourhoods (TPI2000 considering cells within 1850- and 2000-meters radius). Building on TPI300 and TPI2000, landscapes were classified into discrete slope position classes and further into the following landforms (De Reu et al., 2013; Weiss, 2001): 1: canyons, 2: shallow valleys, 3: headwaters, 4: u-shaped valleys, 5: plains, 6: open slopes, 7: upper slopes, 8: hills in valleys, 9: midslope ridges and 10: high ridges.

We further considered the following *geological attributes* with data provided by the Norwegian Geological Survey (NGU):

h. Bedrock weatherability classes 1:50 000-1:250 000
i. Superficial deposits 1:20 000-1:500 000
j. Deposit thickness 1:20 000-1:500 000

The national bedrock map (NGU, 2025a) has 121 categories for the main bedrock type. These are reduced to five weatherability classes (see Table A1 in Appendix A for details) differentiating bedrock which weathers: 1: very easily (carbonous rocks), 2: easily (shale, phyllite), 3: relatively easily (mafic plutonic rocks), 4: slowly (sandstones, conglomerates, felsic volcanic rocks, i.e. granites, and metamorphic rocks, i.e. gneiss) and 5: very slowly (quartzite and quartz sandstones). The number of categories of the national Superficial deposits (NGU, 2025b) were also reduced (see Table A2 in Appendix A for details). The soil condition ("grunnforhold") dataset provided by the Norwegian Institute of Bioeconomy Research (NIBIO) was also considered but excluded since it is not seamlessly available for the study area. In

https://doi.org/10.5194/egusphere-2026-579
Preprint. Discussion started: 2 March 2026

addition, the models include a deposit thickness map (NGU, 2025c) which considers three categories: 1: bare rock, 2: thin deposits (< 0.5m) and 3: thick deposits (> 0.5m). The bedrock, superficial deposits and deposit thickness maps are vector data compiled from datasets with highly variable scales (as specified above).

Moreover, we consider the following *hydrological and meteorological attributes*:

k. Flow accumulation (on logarithmic scale)
l. Groundwater (value for 7 August 2023 relative to normal for the same day)
m. Precipitation (acc. 3 days, mm)
n. Relative precipitation (acc. 3-day relative to normal August precipitation, %)
o. Maximum precipitation intensity (in mm/ 3h)

Flow accumulation is calculated with the D8 flow modelling algorithm (O'Callaghan and Mark, 1984), and the maximum values in an area with 30-meter radius surrounding landslide and control points are allocated. The Topographic Wetness Index (TWI) and the Stream Power Index (SPI) were also calculated (Saleem et al., 2019) but since both indices are derivatives of flow accumulation and slope, they naturally correlate strongly with these variables, and are excluded from the analysis. Pre-'Hans' relative groundwater level is the HBV-modelled groundwater for 7 August 2023 in mm relative to the 1981-2010 normal groundwater level for the same day and a spatial resolution of 1 km (Beldring et al., 2003; Bergström and Lindström, 2015). The total 'Hans' precipitation is derived from seNorge v2018, a gridded observation-based dataset with 1 km spatial resolution and accumulated over three hydrological days (from 7 August 2023, 8 am to 10 August 2023, 8 am). The relative precipitation is the accumulated precipitation for the above-mentioned 'Hans' time period relative to normal precipitation for the entire month of August (for the standard normal period 1991-2020). Further, precipitation intensity was considered, and two raster datasets with maximum one-hourly and three-hourly precipitation values based on the MET Nordic dataset were produced for the same time period. Since these were strongly correlated, we included only the maximum three-hourly precipitation in the models as it had a more even dispersion.

Finally, we included information on *landuse and forest.* Because the forest parameters are not applicable in agricultural land and open terrain, we considered full machine learning models including explanatory factors a. to o. and forest-only models including the whole range of attributes a. to t.

p. Tree canopy (%)
q. Tree type
r. Mean tree height (dm)
s. Tree volume with bark (m^3/ha)
t. Number of trees per hectare

Landuse (a combination of AR5 and AR50) was considered as an explanatory factor with the simplified categories agricultural land, forest or open terrain. Only seven control points but as many as 119 landslide points occur in agricultural land. Therefore, landuse was excluded from the final models as it showed that agricultural land increases the likeliness for shallow landslide occurrence on thin statistical ground. In recent years, many new forest attribute maps have been produced in the framework of the Sentinel-2-derived dataset SR16 (Astrup et al., 2019; Breidenbach et al., 2021; Hauglin et al., 2021;

https://doi.org/10.5194/egusphere-2026-579
Preprint. Discussion started: 2 March 2026

Schumacher et al., 2020). While tree canopy can be considered close to a direct measurement, linear mixed regression is used to predict tree type (with categories 1: spruce, 2: pine and 3: deciduous), mean tree height, tree volume and number of trees per hectare (Breidenbach et al. 2021). Greater uncertainty is associated with the estimation of tree diameter and above/below ground biomass which are therefore not included in this analysis.

4 Machine learning methods

All analyses were performed in the R programming environment (R Core Team, 2024). To model the probability of shallow landslide occurrence against environmental covariates, we chose a gradient boosted decision tree model approach, as this machine learning (ML) method has been found to outperform other ML methods (Cao et al. 2020; Wei et al. 2022). In addition, this method was chosen due to the availability of a recent algorithm, the GPBoost Framework (Gaussian Process Boosting; Sigrist, 2022), that combines gradient boosting with mixed effects modelling, allowing the spatial structure of the landslide data to be explicitly modelled. Preliminary testing of this method on our dataset revealed that the degree to which spatial autocorrelation is accounted for depends on the cross-validation method chosen during hyperparameter selection. Therefore, to formally assess the impact of cross-validation strategy on model performance and spatial structure, we implemented a systematic comparison of 32 models varying in three key dimensions: (1) hyperparameter optimisation approach (termed *nested CV* vs. *simple CV* approaches), (2) spatial structure incorporation (*spatial* vs. *non-spatial*) and (3) cross-validation strategy (*spatial clustering*, *large block*, *small block* and *random CV*). These models were compared for both the *full dataset* (reduced to 570 landslide and 568 control points due to NAs in the predictors), and a subset of data in which all landslide and control points occur within forests (termed *forest-only dataset*, comprising 339 landslides and 434 control points).

4.1 Hyperparameter optimisation

To obtain unbiased estimates of model performance while accounting for spatial structure, we first employed a nested cross-validation strategy (Cawley and Talbot, 2010; Hastie, 2009). This *nested CV* approach consists of two levels of CV: an inner CV loop for hyperparameter optimisation and an outer CV loop for model evaluation (Fig. 2). Initially, the data were assigned to k-folds (where k = 5 or 4, depending on data availability and landslide/control point balance) and in the first iteration of the outer loop, one complete fold (approximately 20% of data) was held out as an independent test set (Step 2 in Fig. 2) while the remaining 80% was passed to the inner loop. Within the inner loop, the 80% training data were reassigned to 4 or 5 folds for cross-validation, where all combinations of candidate hyperparameters (see below) were evaluated across all inner CV folds (Steps 3 and 4 in Fig. 2). Hyperparameters were optimised using the built-in *gpb.grid.search.tune.parameters* function of the *GPBoost* package (Sigrist, 2022), which allows the user to provide a grid of parameters across which to search and to provide pre-defined folds for CV (Step 4 in Fig. 2). We started with the following grid:

https://doi.org/10.5194/egusphere-2026-579
Preprint. Discussion started: 2 March 2026

- Learning rate: 0.01, 0.05
- Minimum data in leaf: 10, 20, 50
- Number of leaves: 8, 16, 32
- L2 regularisation (lambda): 10, 50, 100

The hyperparameter combination resulting in the best performance (highest mean AUC value) was then used to retrain the model on all 80% of the training data (Step 5) and the predictions tested on the 20% held out data (Step 6), providing the Test AUC metric for that outer fold. This process was repeated for each outer fold, allowing every original fold to serve once as the test fold. This is important, because the choice of test fold can impact the evaluation, and the mean Test AUC across the 4 or 5 resulting inner loop models provides a more robust measure of model performance than a single test AUC. Other metrics calculated from these models (proportion of variance explained by spatial component, Moran's I index of residual spatial autocorrelation) were also averaged across folds. Finally, to provide a model for interpretation purposes, a final model is constructed using 100% of the data with the median hyperparameter values identified through the nested CV process (Hastie, 2009).

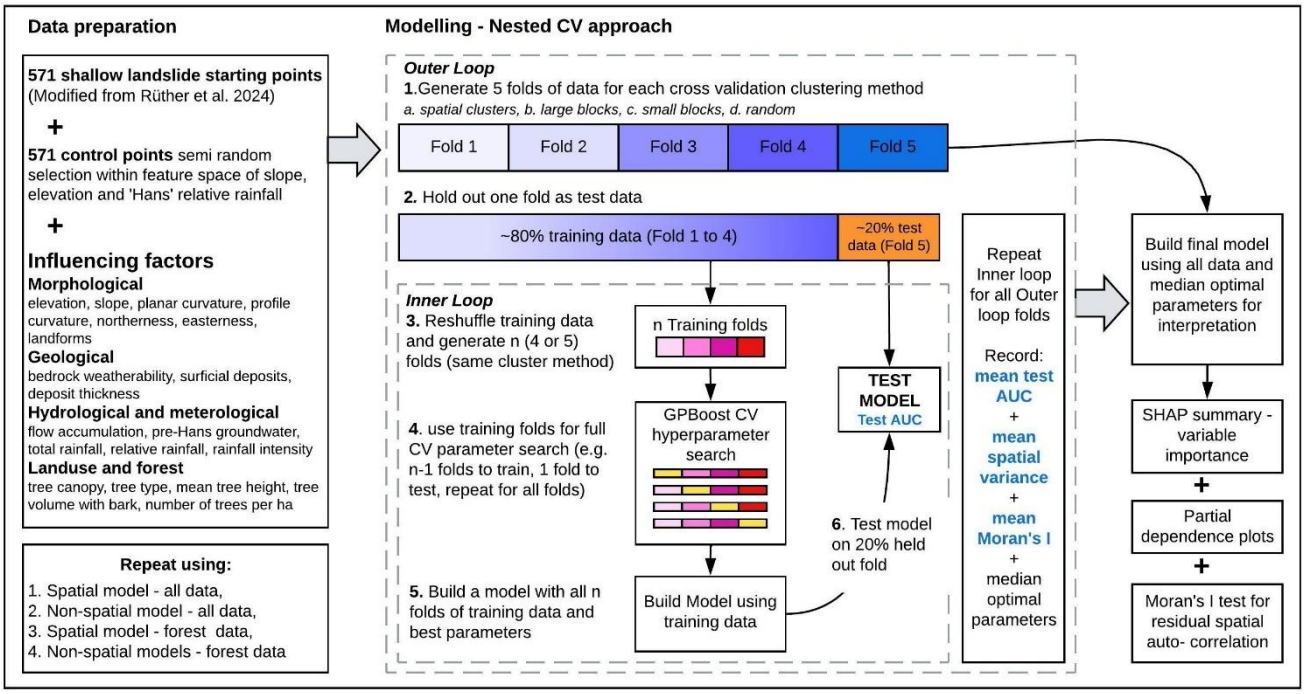


Figure 2 Flowchart summarising data and methods for the nested cross-validation (CV) approach presented in this study. For each of the four model types listed (1. to 4., bottom left), the nested CV workflow was completed using four different CV clustering methods (a. to d., specified under step 1). The methodology for additional modelling runs following a simpler approach is presented in Appendix B.

The alternative, non-nested *simple CV* approach was also tested as a less computationally intensive method and because it appears to resemble approaches commonly employed in the recent landslide literature. In this approach, hyperparameter

https://doi.org/10.5194/egusphere-2026-579
Preprint. Discussion started: 2 March 2026

optimisation was performed using CV on the full dataset and the final model was trained using the entire dataset with optimised hyperparameters (Fig B1). The results are presented in Appendix B for comparison.

4.2 Spatial structure

Spatial models were implemented using the *GPBoost* package (Sigrist, 2022), which combines gradient boosted decision trees with Gaussian process (GP) random effects to account for spatial autocorrelation. The model takes the form:

$$P(landslide_i = 1) = \Phi(f(x_i) + \xi_i)$$

where Φ is the standard normal cumulative distribution function (probit link), $f(x_i)$ represents the ensemble of boosted trees as a function of environmental covariates x_i and ξ_i is a spatially structured random effect modelled as a Gaussian process:

$$\xi \sim GP(0, C(\cdot,\cdot))$$

The spatial covariance function C(·,·) was specified as a Matérn function with smoothness parameter ν = 1.5, which allows for moderately smooth spatial transitions in landslide susceptibility. For comparison, *non-spatial* models were developed using only the gradient boosting component (setting gp_model = NULL in *GPBoost*), relying exclusively on environmental covariates without spatial random effects.

4.3 Cross-validation strategy

To evaluate how spatial dependence in training and validation data affects model performance and residual spatial autocorrelation, we implemented the above modelling procedures using four cross-validation strategies with increasing degrees of spatial separation. The most stringent spatial separation was achieved using k-means clustering on point coordinates via the *spatialsample* package (Mahoney et al., 2023). Points were clustered into five spatially cohesive groups (k=5, algorithm = "Hartigan-Wong"), with each cluster serving as a fold. This approach ensures geographic separation between folds while maintaining relatively balanced fold sizes by adaptively grouping points regardless of regular grid constraints.

However, while this method of CV often resulted in models that successfully accounted for spatial autocorrelation, models were also difficult to interpret because the spatial element of the model explained the majority of variance in landslide occurrence. To relax this constraint, we tested two types of block CV using the *blockCV* package (Valavi et al., 2018). For large blocks, we forced the *spatial_block_cv* function to identify 5 blocks in a single column (3 blocks for the forest-only data). This resulted in strong spatial separation between folds but at the cost of some fold size imbalance, reflecting the natural heterogeneity in shallow landslide distribution across the study area. To relax constraints further, we selected a smaller block size to create a checkerboard pattern of 8 x 8 blocks across the study area. This pattern allowed for more balanced folds as blocks were systematically assigned to five folds such that adjacent blocks were placed in different folds when possible, providing moderate spatial separation between training and validation data.

https://doi.org/10.5194/egusphere-2026-579
Preprint. Discussion started: 2 March 2026

Finally, we used random CV (using the caret package; Kuhn, 2008) to randomly partition the observations into five equally sized folds using stratified sampling to maintain equal proportions of landslide and non-landslide observations in each fold. This represents the most common approach in the literature but does not account for spatial autocorrelation (see Sect. 5).

4.4 Model evaluation

For spatial models, the GP model parameters (spatial variance and range) were estimated jointly with tree parameters at each boosting iteration, and the GP component was used for validation set predictions. Optimal parameters were selected based on mean cross-validation AUC across folds. For nested CV models, we report both inner loop CV performance (Mean CV AUC and Mean Training AUC), evaluating generalisation to held-out spatial regions, and outer loop test set performance (Mean Test AUC), evaluating transferability to unseen areas. For simple CV models, cross-validation performance represents expected predictive accuracy, while performance on the full training dataset reflects within-study-area interpolation accuracy.

To evaluate whether CV strategies successfully removed spatial dependence, we calculated Moran's I statistic on model residuals (observed - predicted probability) using the *spdep* package (Bivand and Wong, 2018). A spatial weights matrix was constructed based on inverse distance weighting with a maximum distance threshold of 12 km. Statistically significant positive spatial autocorrelation ($p < 0.05$) indicates remaining spatial structure in residuals, suggesting the CV strategy did not adequately separate spatially dependent observations. Moran's I was calculated for both inner and outer loop models.

We quantified variable importance using SHAP (SHapley Additive exPlanations) values via the *shapforxgboost* package (Liu and Just, 2019). SHAP values are calculated for each covariate and provide a robust estimate of the contribution of covariate values (feature values) to individual predictions of landslide probability. The mean absolute SHAP values provide an indication of a variable's overall importance in predicting landslides.

5 Results

In the following, results from 16 models following a nested CV approach for hyperparameter optimisation (Fig. 2) are presented, divided into the account of the full models (Sect. 5.1) and forest models (Sect. 5.2). Additional 16 models following a simple CV approach for hyperparameter optimisation (Fig. B1) are summarised in Appendix B. The relative importance of explanatory factors is highly sensitive to the structure of the models. While reporting model performance metrics for all models (Tables 1, 2, B1 and B2), for brevity, we opt to focus the presentation of relative importance of explanatory factors (Figs 3, 5, B2 and B4) on three informative cases: (1) the most rigorous, but restrictive extreme - spatial clustering CV with spatial model (models 1a and 3a), (2) the least rigorous and relaxed extreme - random CV with non-spatial model (models 2d and 4d) and (3) the practical compromise with acceptable performance and moderate autocorrelation reduction - small block CV with spatial model (models 1c and 3c). For the compromise models based on the full dataset (model 1c) and forest-only dataset (model 3c), we further explore the top six predictors in greater detail

https://doi.org/10.5194/egusphere-2026-579
Preprint. Discussion started: 2 March 2026

providing partial dependence plots relating feature values to their model contributions and showing a second features with strongest interaction (Figs 4, 6, B3 and B5). The selection of the top six predictors is somewhat arbitrary since there is no guidance on what characterises a top predictor or on how many to pick.

5.1 Full models

The only models to remove residual spatial autocorrelation were the spatial models parameterised with small or large block CV, although the spatial clustering CV was also close (Table 1) where 4 of 5 of the spatial CV inner loop models successfully accounted for spatial structure (not shown). This suggests that random CV and non-spatial approaches to ML landslide modelling do not adequately account for spatial structure in the data. Out of the models accounting for spatial structure, the small block CV approach provides the better predictive performance with a test AUC of 0.77. Our systematic exploration of model performance revealed a general tendency for the more restrictive CV methods to better account for spatial autocorrelation and for the model's spatial structure to absorb more of the variance in landslide occurrence, but at the expense of predictive power. At the other extreme, the random CV method provided high performance metrics, but these are likely to be over-optimistic, even when spatial structure is incorporated into the model. Thus, a conservative balance between accounting for spatial structure and retaining interpretability is most useful, provided that conclusions are accompanied by a cautionary warning. The small block CV strategy with a spatially explicit model appears to strike such a balance.

Table 1 Mean (and standard deviation) model performance metrics of two Gaussian Process Gradient Boosted classification models using four different cross-validation (CV) methods on the full landslide dataset. AUC = area under the curve, where values close to 1 indicate near perfect accuracy of landslide/control point classification. Mean CV AUC = mean value from hyperparameter search (Step 4 in Fig. 2); Mean training AUC = mean value from inner loop training models (Step 5 in Fig. 2); Mean test AUC = mean value from testing inner loop models (Step 6 in Fig. 2). This value is in bold as it is the most useful metric for evaluating performance. Spatial variance indicates the amount of variance explained by the spatial structure of the model (absent from non-spatial models). The significance of the Moran's I test for spatial autocorrelation is indicated by asterisks (*** = $p < 0.001$, ** = $p < 0.01$, * = $p < 0.05$, ns = non-significant). A significant Moran's I indicates that residual spatial autocorrelation is likely.

Model & CV type	Mean CV AUC	Mean training AUC	Mean test AUC	Spatial variance	Moran's I
Nested CV approach					
1. Spatial models					
1a. Spatially clustered CV	0.64 (0.05)	0.98 (0.01)	**0.64 (0.06)**	0.70	0.02 *
1b. Large block CV	0.61 (0.03)	0.97 (0.01)	**0.57 (0.10)**	0.71	0.01 (ns)
1c. Small block CV	0.67 (0.05)	0.98 (0.01)	**0.77 (0.06)**	0.68	0.02 (ns)
1d. Random CV	0.86 (0.01)	1 (0)	**0.90 (0.01)**	0.47	0.18 ***
2. Non-spatial models					
2a. Spatially clustered CV	0.67 (0.02)	0.99 (0.01)	**0.68 (0.08)**	-	0.11 ***
2b. Large block CV	0.69 (0.03)	0.97 (0.03)	**0.67 (0.09)**	-	0.13 ***

https://doi.org/10.5194/egusphere-2026-579
Preprint. Discussion started: 2 March 2026

2c. Small block CV	0.75 (0.03)	0.96 (0.07)	**0.71 (0.13)**	-	0.15 ***
2d. Random CV	0.88 (0.01)	0.99 (0.01)	**0.87 (0.02)**	-	0.08 ***

Across all the tested full models, directional effects of explanatory factors and threshold values for specific variables seem to be consistent. What changes significantly across the different tested models is the relative importance of explanatory variables (Fig. 3, Fig. B2).

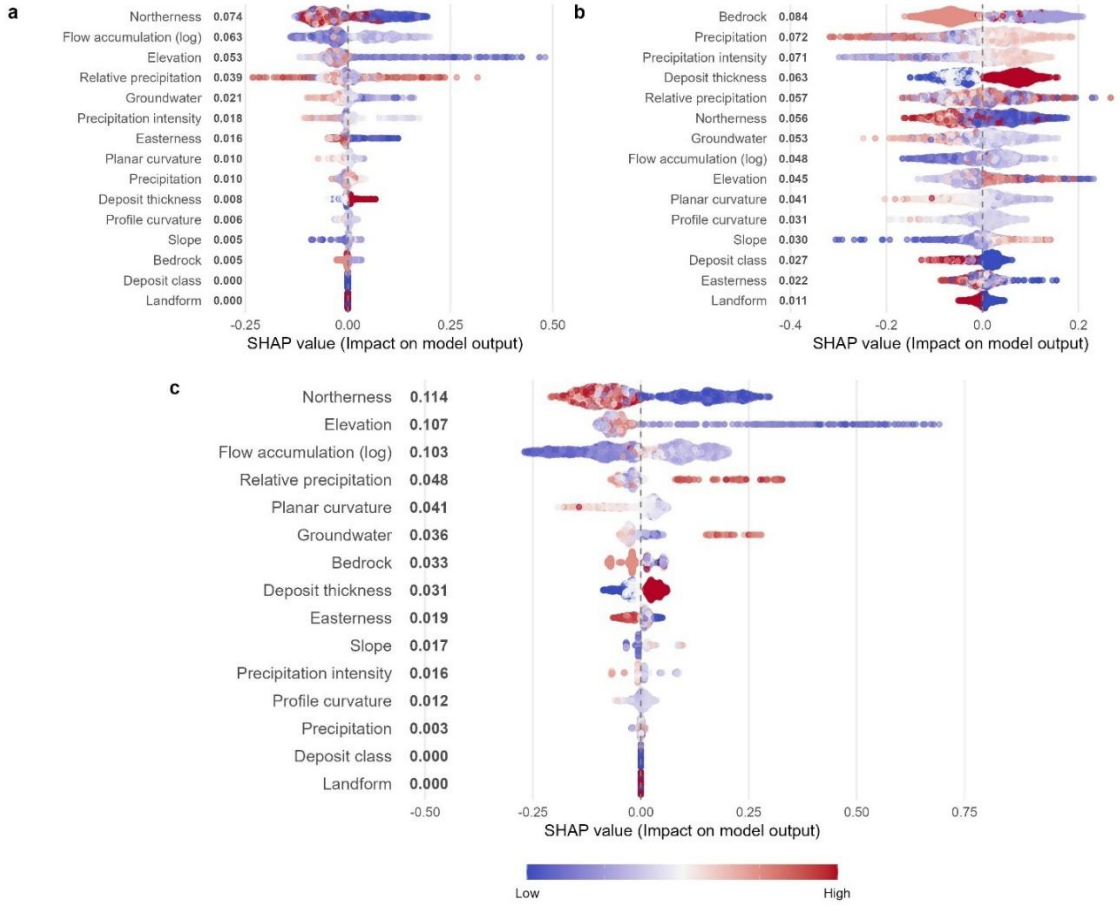


Figure 3 Shapley value summary plots for three models: a) spatial model with spatially clustered CV (model 1a), b) non-spatial model with random CV (model 2d), c) spatial model with small block CV (model 1c). Plots a) and b) represent the two extremes of underfitting and overfitting while c) represents a compromise. The bold numbers after variable names are mean absolute SHAP values, representing relative importance. Points are coloured by normalised feature value.

In the spatial model with spatially clustered CV (model 1a), the spatial structure explains most of the observed variance (70%; Table 1). Further, the included explanatory variables have generally low mean SHAP values suggesting that their ability to discriminate between landslide and control points is limited (Fig. 3a). This is likely because the large spatial component removed by the model leaves little variance to be explained or "learned" by the model. Here the top three

https://doi.org/10.5194/egusphere-2026-579
Preprint. Discussion started: 2 March 2026

variables are northerness, flow accumulation and elevation. On the contrary, the non-spatial model with random CV (model 2d) is overly optimistic and appears to overfit the data, which can be seen by the detailed pattern of SHAP values for all explanatory variables included in the model (Fig. 3b). The top three variables of this model are bedrock weatherability, precipitation (accumulated over 3 days in mm) and precipitation intensity (max 3h precipitation during 'Hans'), showing the high sensitivity of variable importance to model selection.

In model 1c, the spatially explicit model with small block CV and our preferred model for interpretation, five of the top six variables are the same as model 1a, with elevation moving to second most important for explaining landslide occurrences (Fig. 3c). The partial dependence plot suggests the strong influence of elevation is due to large positive contributions to the probability of landslide occurrence at elevations lower than ca. 480 meters a.s.l., while elevation does not contribute to explaining shallow landslide starting points above this elevation (upper mid panel in Fig. 4). An interaction with precipitation intensity further suggests that some of these landslide occurrences at lower elevations may be driven by particularly high 3-hour rain intensities, but not all of them. An underlying sampling bias in the landslide point dataset due to availability of post-event orthophotos at lower elevations was considered as explanation but can be rejected as the shallow landslide starting points are normally distributed across the observed elevation range. Northerness and relative precipitation are two additional important explanatory factors in model 1c, ranking 1st and 4th respectively (Fig. 3c), and they also show the strongest interaction with each other (upper and lower left panel in Fig. 4). South-facing slopes are more likely to give rise to shallow landslides than slopes with slightly southern and northern aspects, this effect is however less pronounced for south-facing slopes which experience highest relative precipitation during 'Hans' (upper left panel in Fig. 4). Relative precipitation does not make clear contributions to the model before reaching double the amount of normal August precipitation when it contributes positively to the probability of shallow landslides on north-facing slopes and less positively on south-facing slopes (lower left panel in Fig. 4). Flow accumulation values between 10 000 and 5 000 000 m^2 contribute positively to landslide probability, in particular on north-facing slopes (upper right panel in Fig. 4). Negative to neutral planar curvature values, corresponding to channelised and planar terrain, contribute to higher shallow landslides probability, in particular on north-facing slopes (lower mid panel in Fig. 4). Pre-'Hans' groundwater levels do not explain landslide occurrence before reaching double the amount of normal groundwater levels, especially for low-lying areas (lower right panel in Fig. 4).

Several of the included categorical variables are less important in this compromise model 1c, in particular surficial deposits and landforms which score lowest, in 14th and 15th place, respectively. Bedrock weatherability and deposit thickness are the categorical variables which came highest on the list over explanatory factors in model 1c (7th and 8th place, respectively). Slope is relatively unimportant in this full model (in 10th place), but maybe most surprisingly, the meteorological variables 'Hans' precipitation intensity (max 3h, mm) and total 'Hans' precipitation (acc. 3 days, mm) do not contribute significantly to explaining shallow landslide occurrence (11th and 13th place, respectively). Model S1c, based on the simple CV approach and presented in Appendix B, has similar results. Deposit thickness is, however, more important than relative precipitation in

https://doi.org/10.5194/egusphere-2026-579
Preprint. Discussion started: 2 March 2026

this model (Fig B2), with thickness of more than 0.5 meter contributing positively to landslide probability, especially at low elevations (Fig B3).

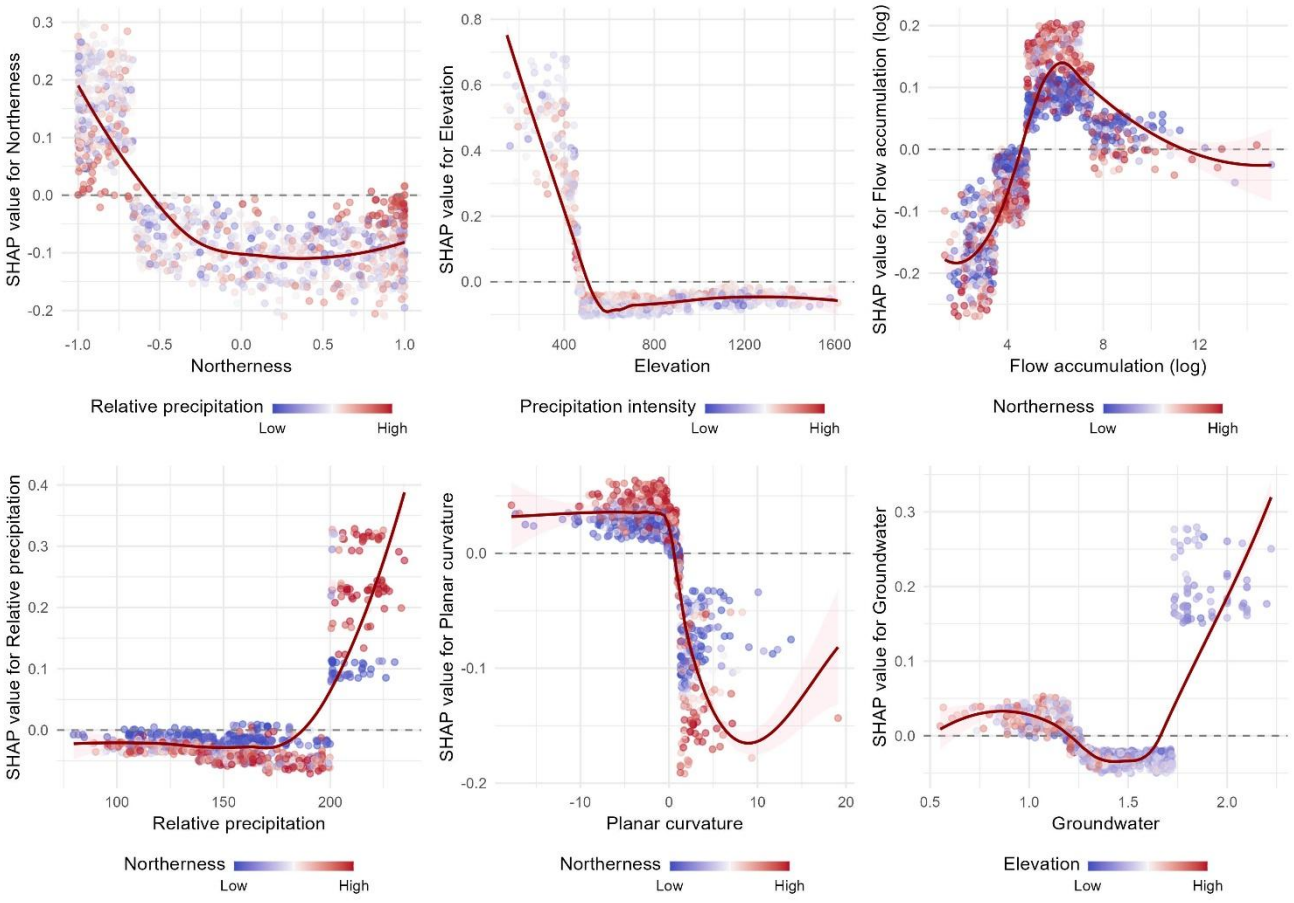


Figure 4 Partial dependence plots for the top six variables of the spatial model with small block nested CV (model 1c). Each point represents an observation and is plotted with its feature value on the x-axis and SHAP value (contribution to model output) on the y-axis. SHAP values above 0 indicate that a feature value is likely to make a positive contribution to the probability of a landslide. Points are coloured by a second feature chosen to minimise the variance of the SHAP value given both features and represents the strongest interaction. Solid red lines are LOESS smoothed lines and shading represents 95% confidence intervals.

5.2 Forest models

The same modelling runs as presented above for the full dataset were repeated based on the subsampled forest-only dataset consisting of 339 shallow landslide starting points and 434 control points. The same tendencies in model performance were found as with the full data models, although only the large block nested CV spatially explicit model fully accounted for spatial autocorrelation. However, the small block nested CV spatial model was close to having a non-significant Moran's I test result (Table 2) and could have had a lower Moran's I values with some manipulation of the block size and number. For

https://doi.org/10.5194/egusphere-2026-579
Preprint. Discussion started: 2 March 2026

this dataset, the relatively low number of landslide points makes it difficult to create spatial blocks with balanced and sufficient test landslides, suggesting that a dataset of this size may be on the borderline of successfully applying this approach. However, the test AUC of the large block CV model is barely better than random guessing and even though the small block CV model does not fully account for spatial autocorrelation, we again suggest that this may present the best trade-off between predictive power and over-optimistic performance.

Table 2 Mean (and standard deviation) model performance metrics of two Gaussian Process Gradient Boosted classification models using four different cross-validation (CV) methods on the forest-only landslide dataset. Mean CV AUC = mean value from hyperparameter search (Step 4 in Fig. 2); Mean training AUC = mean value from inner loop training models (Step 5 in Fig. 2); Mean test AUC = mean value from testing inner loop models (Step 6 in Fig. 2). This value is in bold as it is the most useful metric for evaluating performance. Spatial variance indicates the amount of variance explained by the spatial structure of the model (absent from non-spatial models). The significance of the Moran's I test for spatial autocorrelation is indicated by asterisks (*** = $p < 0.001$, ** = $p < 0.01$, * = $p < 0.05$, ns = non-significant). A significant Moran's I indicates that residual spatial autocorrelation is likely.

Model & CV type	Mean CV AUC	Mean training AUC	Mean test AUC	Spatial variance	Moran's I
Nested CV approach					
3. Spatial models					
3a. Spatially clustered CV	0.64 (0.04)	0.99 (0.01)	**0.71 (0.10)**	0.63	0.08 ***
3b. Large block CV	0.61 (0.04)	0.96 (0.01)	**0.56 (0.05)**	0.70	0.01 (ns)
3c. Small block CV	0.68 (0.05)	0.98 (0.01)	**0.73 (0.10)**	0.68	0.03 **
3d. Random CV	0.86 (0.02)	0.99 (0.00)	**0.89 (0.03)**	0.43	0.16 ***
4. Non-spatial models					
4a. Spatially clustered CV	0.67 (0.04)	0.99 (0.01)	**0.70 (0.08)**	-	0.13 ***
4b. Large block CV	0.68 (0.02)	0.94 (0.06)	**0.62 (0.02)**	-	0.23 ***
4c. Small block CV	0.75 (0.03)	0.99 (0.00)	**0.79 (0.02)**	-	0.11 ***
4d. Random CV	0.88 (0.02)	0.99 (0.01)	**0.88 (0.02)**	-	0.10 ***

Also, across all the tested forest-only models, directional effects of explanatory factors and threshold values for specific variables are consistent, while the ranking of explanatory variables differs between the models (Fig. 5, Fig. B4). As with the full model, we show the shapely value summary plots for the spatial forest-only model with spatially clustered CV (model 3a; Fig. 5a) and non-spatial forest-only model with random CV (model 4d; Fig. 5b) as two extremes and the compromise spatial model with small block CV as the preferred compromise (Fig. 5c).

https://doi.org/10.5194/egusphere-2026-579
Preprint. Discussion started: 2 March 2026

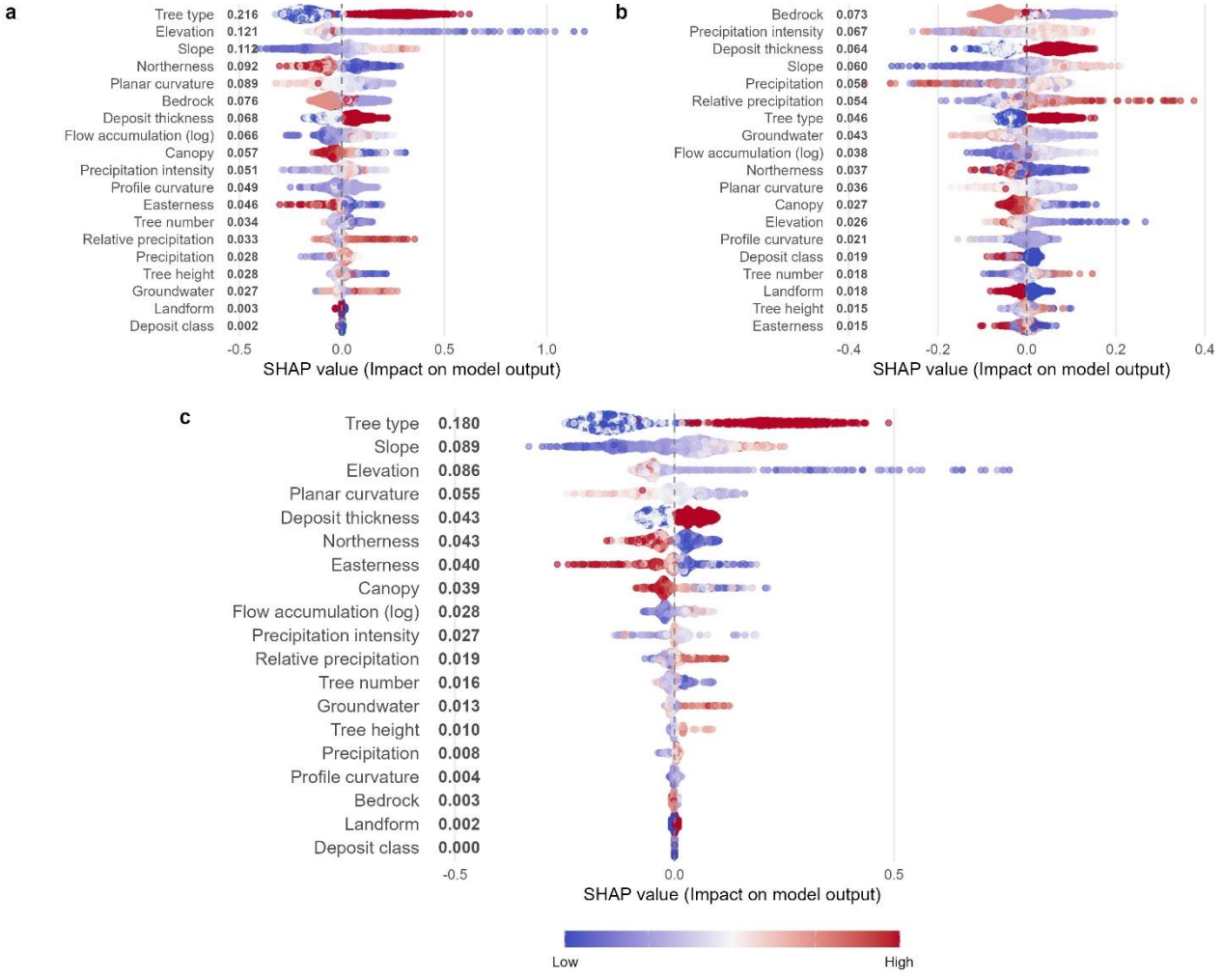
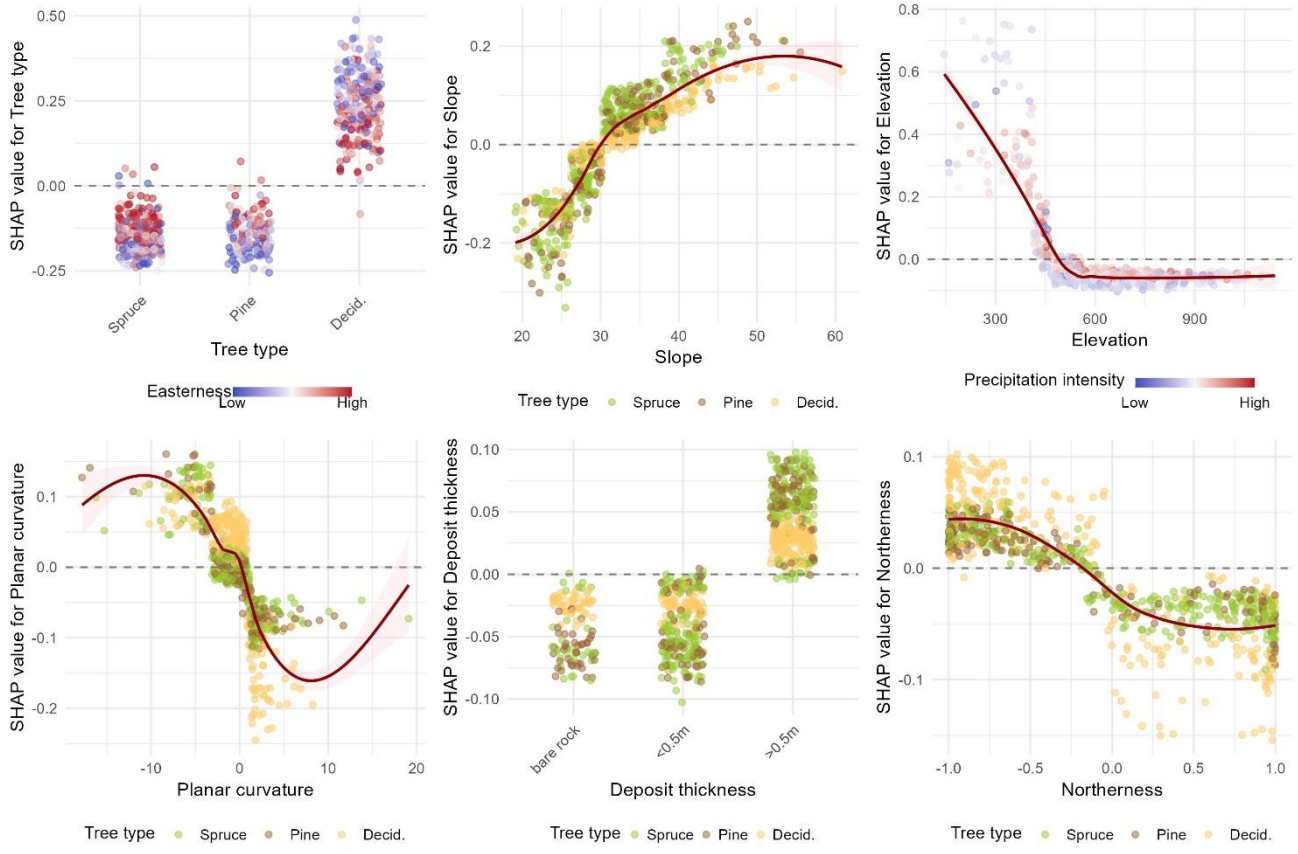


Figure 5 Shapley value summary plots for three models with forest-only data: a) spatial model with spatially clustered CV (model 3a.), b) non-spatial model with random CV (model 4d.), c) spatial model with small block CV (model 3c.). Plots a) and b) represent the two extremes of underfitting and overfitting while c) represents a compromise. The bold numbers after variable names are mean absolute SHAP values, representing relative importance. Points are coloured by normalised feature value.

Interestingly, both spatial models (3a and 3c) have tree type as the most important variable, but this variable does not make the top 6 in the non-spatial random CV model, again highlighting the importance of model design and selection to interpretation. As with the full dataset, five of the top six important variables are the same for the two spatial models, tree type, elevation and slope being the top three variables for both, while bedrock weatherability, precipitation intensity and deposit thickness are most important in model 4d. In model 3c, the spatially explicit forest model with small block CV, the importance of tree type is also outlined by its strong interactions with several of the other top variables (slope, planar curvature, deposit thickness and northerness in Fig. 6). Forested areas dominated by spruce and pine stands are less likely to contribute to predictions of landslide occurrence, particularly on west-facing slopes. By contrast, deciduous forest stands

https://doi.org/10.5194/egusphere-2026-579
Preprint. Discussion started: 2 March 2026

490 contribute positively to landslide probability, again especially on west-facing slopes (upper left panel in Fig. 6). Slope is second most important in this model and the interaction with tree type suggests that the importance of slope is less pronounced in deciduous forest compared to spruce and pine (upper mid panel in Fig. 6). While shallow landslides released during the 'Hans' storm included relatively low slope angles outside of the forest, slope in the forest must be above a threshold of 30 degrees to raise landslide probability (upper mid panel in Fig. 6).



495 **Figure 6** Partial dependence plots for the top six variables of the spatial model with small block nested CV for the forest-only data (model 3c). Each point represents an observation and is plotted with its feature value on the x-axis and SHAP value (contribution to model output) on the y-axis. SHAP values above 0 indicate that a feature value is likely to make a positive contribution to the probability of a landslide. Points are coloured by a second feature chosen to minimise the variance of the SHAP value given both features and represents the 500 strongest interaction. Solid red lines are LOESS smoothed lines.

Landslide occurrences in forest are also most likely at elevations below 480 meters a.s.l. independently of rainfall intensity while low rainfall intensities at higher elevations make landslides less likely (upper right panel in Fig. 6). Further, deposit thickness plays an important role and thick deposits (> 0.5 m) increase landslide susceptibility in all forest types, an effect most pronounced in spruce and pine forest (lower mid panel in Fig. 6). Compared to the full model 1c, where northerness is 505 the top variable, it ranks 6th place. Greater shallow landslide susceptibility on south-facing slopes is most pronounced in deciduous forest but is still present in spruce and pine forests (lower right panel in Fig. 6). Surprisingly, in the compromise

https://doi.org/10.5194/egusphere-2026-579
Preprint. Discussion started: 2 March 2026

forest-only model 3c bedrock weatherability is unimportant and ranks 17th place, with only landforms and surficial deposits scoring lower (18th and 19th place respectively). Also, hydrometeorological variables including flow accumulation, relative 'Hans' precipitation and groundwater which had high scores in the full model 1c (3rd, 4th and 6th place respectively) are relatively unimportant in the forest model 3c (9th, 11th and 13th place respectively). Apart from tree type, the only other forest variable to explain shallow landslide starting points reasonably well is canopy cover (8th place in model 3c), while tree number and tree height have relatively low mean SHAP values and rank 12th and 14th place respectively. While forest model S3c, based on the simple CV approach and presented in Appendix B, has again similar results, it also has one major difference as bedrock weatherability ranks 2nd place (Fig. B4). For bedrock, the first three weatherability classes (1: carbonous rocks, 2: shale and phyllite, 3: mafic plutonic rocks) are likely to contribute to shallow landslides, while the last two weatherability classes (4: sandstones, conglomerates, felsic volcanic and metamorphic rocks and 5: quartz-rich rocks) contribute to stability, in particular on south-facing slopes (Fig. B5).

6 Discussion

In the following, we will discuss our objectives. To start with, we advocate for our approach in addressing the problem of overfitting in landslide modelling with machine learning methods, acknowledging the potential tension between quantitative performance metrics and geomorphologic plausibility (first objective O1; Sect. 6.1). Then, we discuss the impact of important variables in the compromise full model (second objective O2; Sect. 6.2) including the hydrometeorological conditions as well as elevation, soil thickness and aspect as soil indicators. Further, we demonstrate the important role of forest characteristics in influencing landslide susceptibility based on the compromise forest-only model (third objective O3; Sect. 6.3), before providing a synthesis and outlook (Sect. 6.4).

6.1 Quantitative model evaluation versus plausibility

Landslide occurrences and control points exhibited strong spatial autocorrelation as both raw data (test not shown) and as model residuals, at least from those models not adequately addressing the issue. As a result, models that failed to or only partially account for spatial autocorrelation had highest evaluation metrics (test AUC of more than 0.88), comparable to recent studies of landslide modelling with gradient boosting (Choubin et al. 2025; Dai and Huang, 2025; Liu et al. 2021; Cao et al. 2020). However, we argue that such models should be viewed with caution because the results may be biased, over-optimistic and may suggest important predictor variables that are capturing spatial covariance rather than mechanistic relationships (Roberts et al. 2017; Urdangarin et al., 2023). This is also consistent with investigations exploring potential conflicts between quantitative performance evaluation and so-called geomorphic plausibility (Steger et al., 2016; Steger et al., 2021). Variables such as distance to roads, distance to settlements and landuse often have high explanatory power but may represent underlying sampling biases in landslide databases. This is likely because landslides which pose greater risk to society have higher chances of being registered than events far from settlements, e.g. in forested areas. While the systematic

https://doi.org/10.5194/egusphere-2026-579
Preprint. Discussion started: 2 March 2026

mapping approach of the 'Hans' event database (Rüther et al., 2024) minimises the sampling bias, we deliberately chose to not include distance to roads or settlements as explanatory factors. Nonetheless, the observed differences in variable importance in the spatial models versus non-spatial models can be discussed in the light of discrepancies between quantitative evaluation metrics and geomorphic plausibility (Steger et al., 2016; Steger et al., 2021).

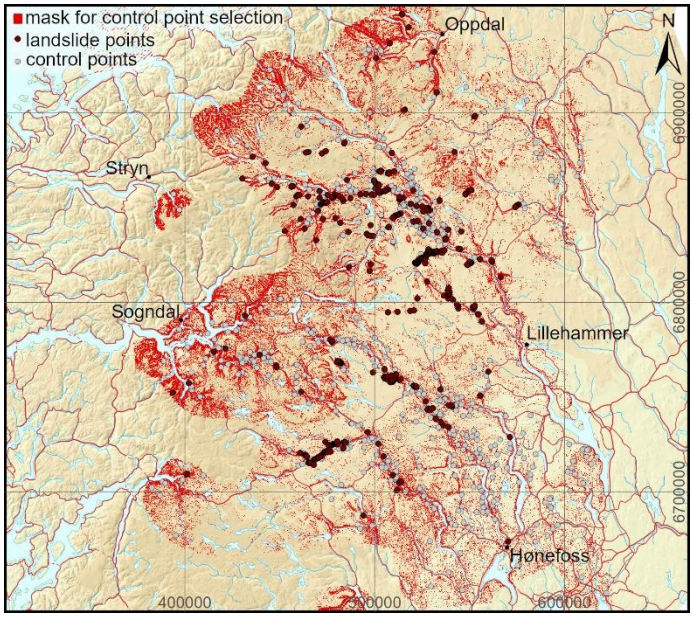


Figure 7 Shallow landslide starting points and control points together with the mask used for control point selection. © Kartverket.

As the 'Hans' inventory consists of rainfall-induced shallow landslides, the influence of precipitation is inevitable. Our control point sampling approach ensures parity in distribution across relative precipitation, slope and elevation (red areas in Fig. 7). This means that the contribution of these variables is reduced to effects within the observed attribute ranges and will generally be lower than for approaches with completely randomised selection of non-landslide points (Dou et al., 2023). Nevertheless, 'Hans' precipitation relative to normal August precipitation is the most important meteorological variable, outperforming precipitation magnitude (total 'Hans' accumulated precipitation) and intensity (maximum 3h 'Hans' precipitation) for several of the spatial full models (4th place in models 1a and 1c) as well as in several of the tested Random Forest models in a recent master thesis (Peeters, 2024). While being unimportant in spatial models with small block CV (13^{h} and 15th place in models 1c and 3c, respectively), precipitation magnitude is highly important in the non-spatial models with random CV (2nd and 5th place in models 2d and 4d, respectively). The same is observed for precipitation intensity, which receives low scores in the spatial models with small block CV (11th and 10th place in models 1c and 3c, respectively), while being amongst the top predictors in the non-spatial models with random CV (3rd and 2nd place in models 2d and 4d, respectively). Based on simple visual examination, it seems plausible that relative precipitation (Fig. 1b) has greater explanatory power than precipitation magnitude (Fig. 1a) because several parts of western Norway received precipitation with comparable magnitude during 'Hans' as the heaviest effected areas in the east without consequences like flooding and

https://doi.org/10.5194/egusphere-2026-579
Preprint. Discussion started: 2 March 2026

shallow landslide occurrence. The lack of consequences in western Norway is likely related to smaller and steeper catchment areas and soil that is better adapted to receiving high amounts of rainfall in western Norway (Ekholt, 2023). The spatial distribution of precipitation intensity during the 'Hans' storm (Fig 8a), suggests that shallow landslides did not occur in the areas with highest rainfall intensities. This may be due to imperfections in the dataset as the MET Nordic hourly precipitation dataset is more reliable in lower-lying areas where it is based on a denser net of direct observations. In this study, the operational gridded dataset is included while a new ensemble-based interpolation method (Lussana et al., 2025) may improve the reliability of rainfall intensity data.

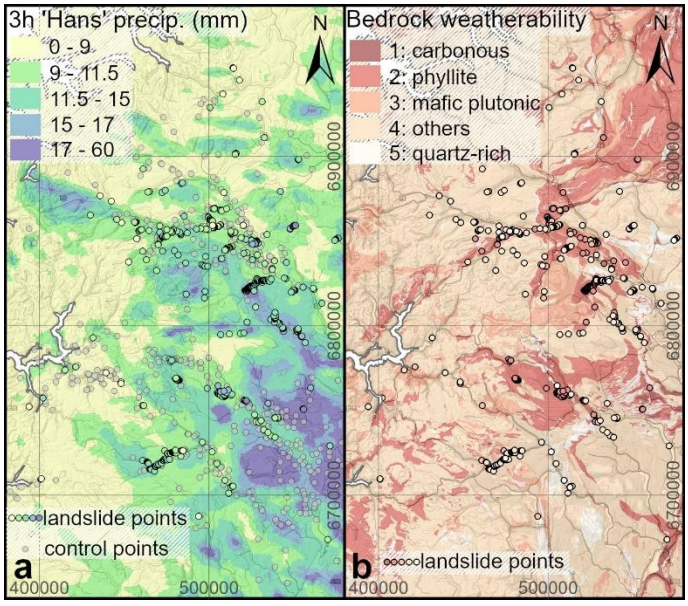


Figure 8 a) Maximum 3h precipitation summed up for the 'Hans' time period based on the MET Nordic dataset, shallow landslide starting points symbolised with colours corresponding to rainfall intensity and control points in transparent grey for reference b) Bedrock weatherability classes and shallow landslide starting points symbolised with colours corresponding to weatherability classes. © Kartverket, SeNorge, NGU.

Areas in the southeast with particularly high rainfall intensities during the 'Hans' storm have steep enough slopes to give rise to landslides as illustrated by the placement of numerous control points (Fig. 8a). Bedrock weatherability (Fig. 8b) can potentially explain why very few shallow landslides were triggered in these southeastern areas despite for favourable meteorological conditions. Since the national bedrock map differentiates the main bedrock type into 121 categories (NGU, 2025a), the dataset was reclassified into five weatherability classes based on the anticipated effect of physical and chemical weathering in a polar climate (Hakim et al., 2021; Momeni et al., 2015). In 7th place, bedrock weatherability is the highest ranked categorical variable in the spatial model with small block CV based on the full dataset (model 1c). Weatherability classes 1 (carbonous rocks), 2 (shales, phyllites), 3 (mafic plutonic rocks) and 5 (quartz-rich rocks) all contribute positively to landslide probability while class 4 (sandstones, conglomerates, felsic volcanic and metamorphic rocks) reduces chances of landslide occurrence; these effects being most pronounced on south-facing slopes. There is a striking co-occurrence of

https://doi.org/10.5194/egusphere-2026-579
Preprint. Discussion started: 2 March 2026

weatherability class 4 (Fig. 8b) and areas which experienced more than 17 mm of 3h rainfall without triggering any landslides (Fig. 8a), suggesting that soil in these areas with slowly weathering bedrock might be generally thinner, less stratified (Cámara et al., 2017) and thus less susceptible to shallow landslides. Notably, as with precipitation magnitude and intensity, bedrock weatherability is more important in the non-spatial models with random CV based on the full dataset (model 2d) and forest-only dataset (model 4d) for both of which it is ranked in 1st place. This detection of alternative key variables in non-spatial models could be explained with spatial confounding (Urdangarin et al. 2023), suggesting that rainfall magnitude, intensity and bedrock weatherability may act as spatial structure proxies in non-spatial models, capturing spatial patterns in landslide occurrence in the absence of explicitly modelled structure. If this is the case, the apparent importance of these variables in non-spatial models may be inflated. However, as the influences of rainfall magnitude, intensity and bedrock weatherability make physical sense or are plausible, the importance of these variables may also be suppressed in the spatial models in the absence of higher quality datasets. Further landslide susceptibility studies that compare spatial and non-spatial ML models are recommended to resolve this question.

6.2 Hydrometeorological conditions and soil indicators explain landslide susceptibility

Rainfall-induced shallow landslides release when a critical porewater pressure in the soil is reached, i.e. rainfall intensity and surface runoff in combination exceed the hydraulic drainage capacity of the subsurface (Bogaard and Greco, 2016; Guzzetti et al., 2008). Therefore, water supply is key to determining susceptibility to shallow landslides for a given event. While precipitation amount and intensity as well as precedent groundwater level are event-specific, flow accumulation is a DEM derivative and a constant measure of the upstream drainage areas into each grid cell giving good indications of the amount of expected surface runoff whilst neglecting the effect of water infiltration. Flow accumulation is one of three topographic variables considered in the national shallow landslide susceptibility map (Fischer et al. 2012; Pullarello et al. 2025). Norway's land surface is categorised into three susceptibility zones depending on shallow landslide activity levels and applied threshold values for flow accumulation vary from 3500 to 9000 m^2 within susceptibility zones. Our results suggest a consistent flow accumulation threshold of 5000 m^2 throughout the study area (upper right panel in Fig. 4). Another factor which is known to contribute to shallow landslide release are pre-event groundwater conditions as saturated soils prior to rainstorms increase chances of shallow landslide triggering (Bondevik and Sorteberg, 2021; Johnson and Sitar, 1990; Montgomery et al., 2009). In most of the tested models, pre-'Hans' groundwater level relative to normal for the same day (modelled for 7 August 2023, 8 am) only increased landslide probability where it exceeded 170%, a relative value which is only attained at low elevations (Fig. 9a). As such, these high values may not be representative of the shallow landslide release areas, but rather an effect of the coarse nature of the modelled groundwater data (1 km resolution). Relative groundwater levels between 125 and 170% even reduce landslide probability in the spatial full model with small block CV (model 1c), which is probably due to its spatially extensive occurrence in eastern Norway (Fig. 9a), making it an imprecise indicator.

https://doi.org/10.5194/egusphere-2026-579
Preprint. Discussion started: 2 March 2026

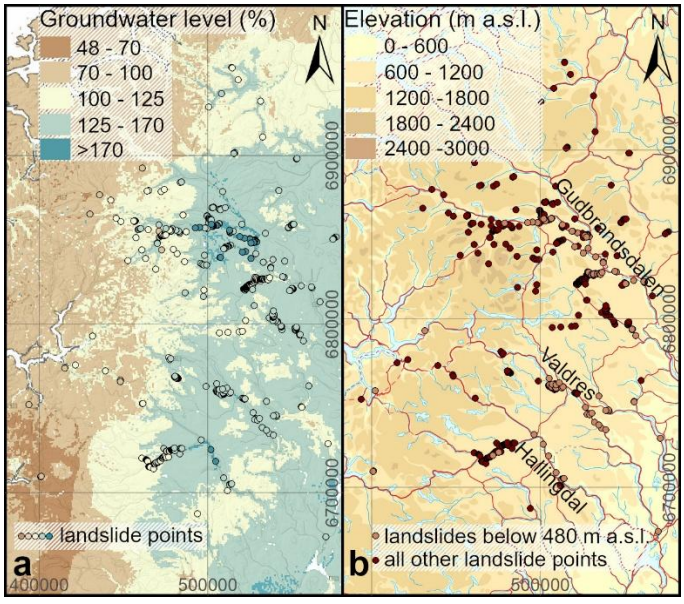


Figure 9 a) Modelled pre-'Hans' groundwater level compared to normal (1981-2010) for 7 August 2023, 8 am and shallow landslide starting points symbolised with colours corresponding to groundwater level.b) Shallow landslide starting points in the study area with those under 480 meters a.s.l. highlighted in lighter brown. © SeNorge, Kartverket.

Despite the chosen control point sampling method, elevation is amongst the top predictors in several of the tested models (e.g. in the nested CV models 1a, 1c, 3a and 3c, as well as in the simple CV models S1a, S1c, S3a and S3c). Elevation explains shallow landslide occurrence second best in the full model 1c and third best in the forest model 3c, in both models driven by the positive contribution of elevation under 480 meters a.s.l., interacting most strongly with precipitation intensity (Figs 4 and 6). Elevation indirectly impacts landslide occurrence due to variations in environmental variables like temperature, rainfall regime and vegetation (Catani et al., 2013; Quevedo et al. 2022). While meteorological, vegetational and soil parameters have been considered as explanatory factors in this study, the included datasets may be too coarse to adequately represent important variations. Numerous debris slides and small debris flows were released in low-lying agricultural hillslopes of Hallingdal, Valdres and Gudbrandsdalen (Figs 9b and 10). The categorical variable landuse has been excluded from the machine learning models due to a large imbalance as only seven control points fell in the category of agricultural land. Thus, the strong influence of elevation may suggest that it works as a proxy for agricultural land. However, the higher probability of landslide occurrence at low elevation is not restricted to agricultural land alone because elevation was also important in the forest-only model 3c. Impressions from fieldwork in the valleys of Hallingdal, Valdres and Gudbrandsdalen (Figs 9b and 10) further suggest that the majority of shallow landslides were released in connection with excessive surface runoff, often redirected along forest roads and sometimes in anthropogenically influenced surficial deposits. Ideally, a variable considering distance to forest roads should have been included in the model, but there is no complete and updated dataset over Norwegian forest roads. In the absence of a direct variable, it is likely that elevation functions as a proxy representing anthropogenic terrain interventions as they are likely to be more extensive with decreasing

https://doi.org/10.5194/egusphere-2026-579
Preprint. Discussion started: 2 March 2026

635 elevation. In addition, the effect of elevation under 480 meters a.s.l. may also correspond to areas with particularly thick and stratified soils. Numerous studies suggest increased shallow landslide susceptibility with soil thickness (e.g. Thiery et al., 2014; Cascini et al., 2017; Guzzetti et al., 1999). Thin soils have been shown to correlate with greater scree cover and exhibit higher infiltration capacity and lower erodibility (Fu et al., 2011). Moreover, soil depth as well as soil characteristics including stratification have been shown to be crucial in modelling shallow landslide occurrences (Kim et al., 2015; Tofani
640 et al., 2017). Thicker soils are thus more likely to be heterogeneous. Heterogeneity in the subsurface includes impermeable soil layers and spatial variability of layer thicknesses which both contribute to the build-up of locally elevated porewater pressures and triggering of shallow landslides (Johnson and Sitar, 1990; Tufano et al. 2021; Oguz et al., 2022). We therefore suggest that the effect of elevation interacting with precipitation intensity (Fig. 4) is a mixture of anthropogenic influence and thicker, more stratified soils at elevations below 480 meters a.s.l. that experienced particularly high 3h precipitation
645 during the 'Hans'-storm.

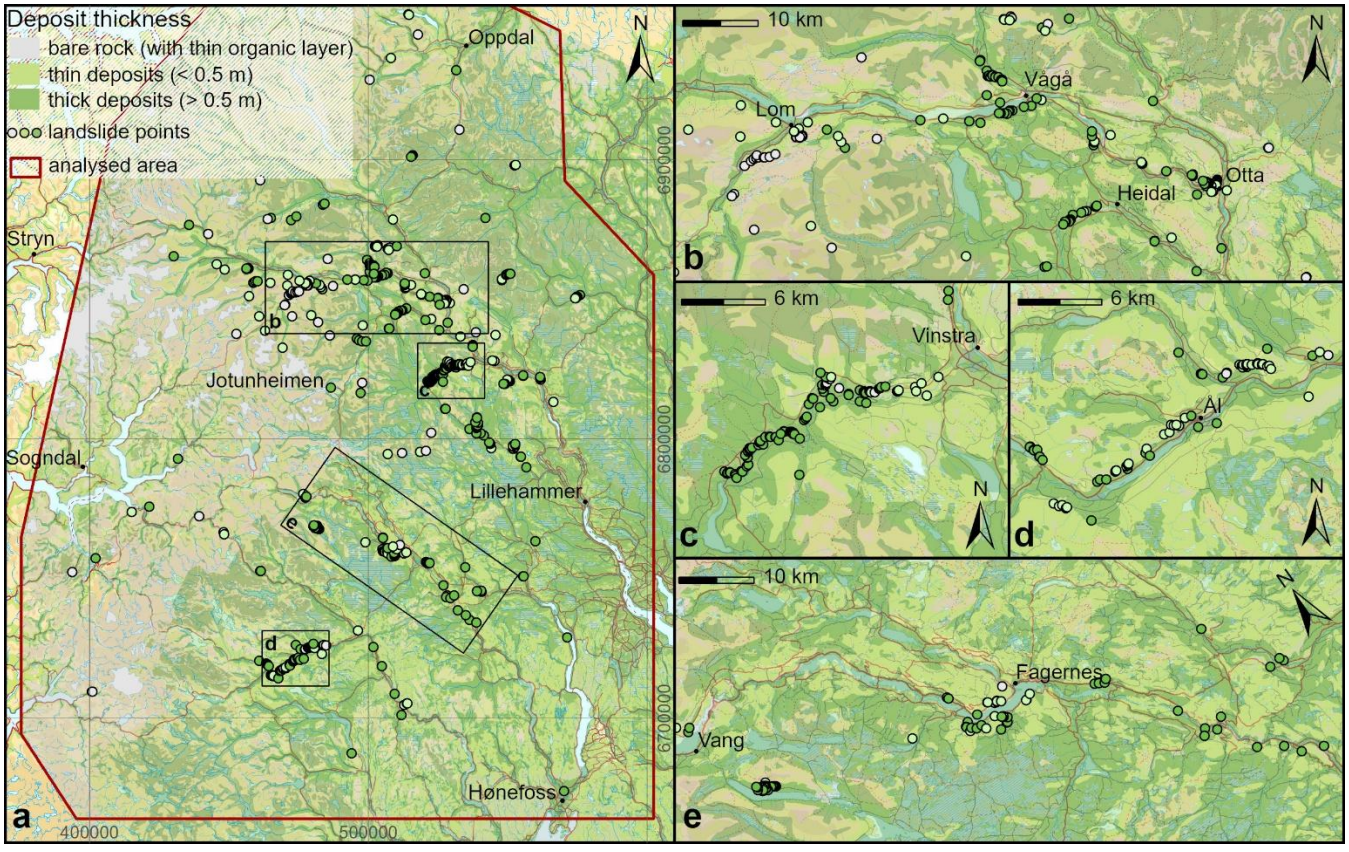


Figure 10 Spatial distribution of the categorical variable deposit thickness and shallow landslide starting points symbolised with corresponding colours for a) the entire study area, b) Ottadalen, c) Skåbudalen, d) inner Hallingdal and e) Valdres. © Kartverket, NGU.

650

https://doi.org/10.5194/egusphere-2026-579
Preprint. Discussion started: 2 March 2026

Deposit thickness, included as explanatory factor in all models, is a categorical variable with only three categories and is likely to not reflect variations in soil thickness, let alone soil characteristics, in great enough detail (Fig. 10). Despite these restrictions, it was ranked in 8th and 5th place in the full model 1c and forest model 3c, respectively. As summarised in the above paragraph, the influence of deposit thickness on landslide probability is well documented. Yet, data availability is often limited and most studies rely on soil thickness modelling approaches (Dietrich et al., 1995; Moore et al., 1993; D'Odorico, 2000) and more lately supported by machine learning (Baltensweiler et al., 2021; Gomes et al., 2023; Wadoux et al., 2020). The deposit thickness dataset included in this study is a biproduct of the national surficial deposit map and based on direct observations but at highly variable spatial scale (NGU, 2025c; Fig. 10).

Northerness is the most important factor in the full model 1c and ranked 6th place in the forest model 3c, which are both sampled with small blocks for CV. Aspect is commonly considered in studies of landslide susceptibility, mostly as it influences evapotranspiration and weathering processes and consequently vegetation and root development (Sidle and Ochiai, 2006). As the 'Hans' low-pressure system approached Norway from the southeast, one may also expect more precipitation on the windward mountain slopes, with southern and eastern aspects. This orographic effect may not be adequately reflected in the available gridded precipitation data with 1 km resolution. While southern aspects clearly increased the probability of shallow landslides during the 'Hans'-storm, the variable easterness is less influential and ranked 9th and 7th place for the full model 1c and forest model 3c, respectively. Moreover, the models suggest west-facing slopes were more susceptible to shallow landslide occurrence during 'Hans' than east-facing slopes, making an orographic effect less likely. Whilst the observed effects of aspect on landslide probability may be related to differences in vegetation, northerness explains landslides less well in the forest models than in the full models and, therefore, we argue that aspect can also be understood as a proxy for soil thickness and characteristics. Due to reduced weathering rates soils at comparable elevations will likely be thinner and less stratified in northern and eastern aspects compared to south- and west-facing slopes. While this directional effect is well-known for mechanical weathering in Norway due to fewer freeze-thaw cycles on north- and east-facing slopes, it may also include biological weathering, as smaller species diversities on north-facing slopes (Heikkinen, 1991; Parker, 1988) may reduce organic and regolith production (Phillips et al., 2021; Sidle and Bogaard, 2016). The observed interaction of northerness with relative precipitation is interesting in this context, as it suggests that north-facing slopes are most prone to the occurrence of shallow landslide when exposed to close to double the amount of normal August precipitation (Fig. 4). South-facing slopes on the other hand, raise landslide probability in general while contributing neutrally to landslide probability for areas which received double the amount of normal August precipitation (Fig. 4). In relatively thick and stratified soil cover on southern and western slopes, chances of building up sufficient pore pressure over an impermeable soil or bedrock layer are higher than over a thinner, more homogeneous soil cover which often have higher infiltration capacity (e.g. Tufano et al., 2021) on northern and eastern slopes. Observations from the Jølster event on 30 July 2019 in western Norway indicate that thin soil may be more susceptible to shallow landslide releases than thicker soil given extraordinary rainfall intensities (Rüther et al., 2022).

https://doi.org/10.5194/egusphere-2026-579
Preprint. Discussion started: 2 March 2026

6.3 Forest effects on shallow landslide susceptibility

Vegetation, and most importantly trees, are drivers of slope stability through mechanical effects (Waldron, 1977; Wu et al., 1979) and hydrological effects, including interception of precipitation, root water uptake and evapotranspiration (Kim et al. 2017). Amongst the mechanical effects root networks, in particular fine roots with diameters smaller than 2 mm, provide tensile strength to the soil (Schwarz et al., 2010b), while the surcharge of trees as well as drag forces due to wind may influence slope stability slightly negatively (Fan and Lai, 2014; Kim et al. 2020). Due to the combined hydrological and mechanical reinforcement effects provided by vegetation and forests, they are a part of what is coined Nature-based Solutions (NbS; Capobianco et al., 2025; De Jesús Arce-Mojica et al., 2019); meaning that when managed correctly forests may reduce shallow landslide risk.

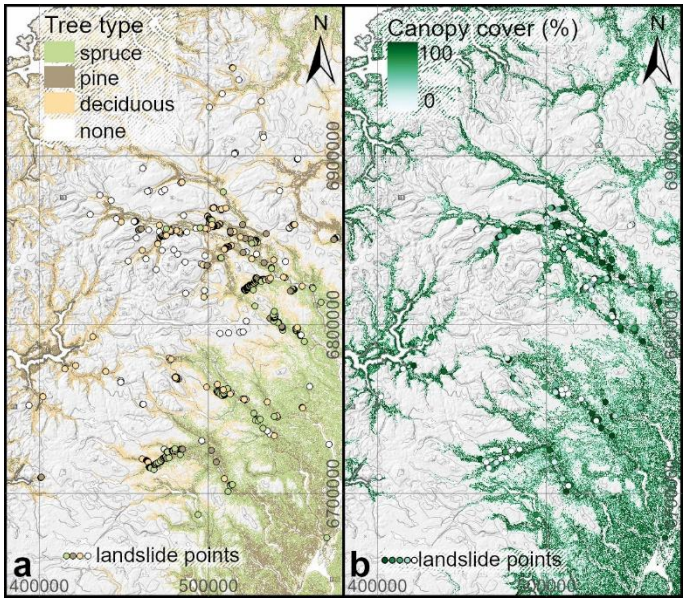


Figure 11 a) Tree type and shallow landslide starting points symbolised according to tree type. b) Canopy cover and shallow landslide starting points symbolised according to canopy cover. © Kartverket, NIBIO.

In the spatial forest model with small block CV (model 3c), tree type in 1st place and canopy cover in 8th are the most influential forest variables (Figs 11a and 11b). Tree type is known to influence slope stability due to differences in their root systems which determine how much tensile strength they provide within the soil, across potential soil boundaries and potentially across the soil-bedrock-interface. Both spruce and birch, the most common deciduous tree in Norway, have relatively shallow root systems, while pine has a more substantial root system and anchors the soil to the bedrock substrate (Kutschera and Lichtenegger, 2002). However, the growth of the important fine root system has also been shown to depend on nutrient availability in the soil and competition from other species (Schmid, 2002). Planted spruce on steep slopes is seen as a particular threat to slope stability after clear-cutting or due to uprooting of older stands (Sidle, 1992). Therefore, the fact that deciduous forest shows the greatest susceptibility to shallow landslide occurrence, while spruce and pine forests reduce

https://doi.org/10.5194/egusphere-2026-579
Preprint. Discussion started: 2 March 2026

landslide susceptibility, may seem surprising at first as their provided mechanical reinforcement is not expected to be any lower than in spruce forest. While the hydrological effect of root water uptake cannot easily be estimated from aggregated forest data as root water uptake has been shown to depend on both under- and overground species interactions (Kinzinger et al., 2023), canopy cover may provide an estimate to the effect of interception. Amongst landslide starting points in deciduous forest mean canopy cover (70.4%) is comparable to spruce forest canopy (70.5%), while pine canopy is less dense on average (61.9%). Thus, the hydrological effect of interception probably does not explain the observed higher landslide susceptibility in deciduous forest either. It is more likely that greater susceptibility of deciduous forest to shallow landslides under 'Hans' is an indication of the type of terrain and soils where this forest typically grows, i.e. in steep terrain and poor soils not suited for production forests and in relatively thin soils towards the tree line (Fig. 11a). Canopy cover (Fig. 11b) is not as influential as tree type but comes in 8th place in the compromise model 3c, suggesting that forests in general and deciduous forest in particular, contribute to stability when canopy cover exceeds 75%. Other forest-specific variables like tree number and tree height have relatively small influences on both the compromise and other forest-models (Fig. 5). This may suggest that the satellite-derived forest SR16 datasets at 16 m resolution (Astrup et al. 2019; Breidenbach et al. 2021; Hauglin et al. 2021) are too coarse to adequately represent forest structure. Model 3c further shows that shallow landslide release on forested slopes is most likely in thick soils. This effect is particularly driven by occurrences in spruce and pine forests while deciduous forest on thick soil increases landslide probability only slightly (Fig. 6). Again, soil thickness may be related to soil heterogeneity which is likely the real driver of instability.

Across both spatial and non-spatial models with different CV sampling methods, slope angle is more important in the forest-only models (e.g. models 3a, 3c and 4d) compared to their full model counterparts (e.g. models 1a, 1c and 2d). It is important to note that the sampling of control points within observed landslide slope ranges reduced the significance of slope in the full models. Typically, slope is the most powerful morphological variable and particularly effective in explaining the spatial distribution of landslides (Reichenbach et al. 2018; Corominas et al. 2014; Malamud et al., 2004). Slope ranges for the 'Hans'-induced landslide starting points are surprisingly wide ranging from 19 to 61 degrees. This includes gentler and steeper terrain than typically considered; e.g. the Norwegian shallow landslide susceptibility map considers slope angles of 22 to 45 degrees for the zones with highest historic activity levels (Pullarello et al. 2025). The mapping approach of the event inventory (Rüther et al. 2024) implies that shallow landslides in the most remote areas are only detected based on satellite data where mass movement processes cannot be determined unequivocally. The inventory may therefore include debris floods which can start in steeper terrain and have longer runouts than debris flows (Church and Jakob, 2020; Hungr et al., 2014) and, in some cases, starting points could have been mapped falsely where debris spreads out from a stream channel at lower slope angles. Despite these mapping-related constraints, 'Hans' starting points which were studied in greater detail also suggest an unusual wide range of slope angles. The mean slope angle for landslide starting points in the forest is 33.8 degrees as compared to 28.1 degrees for landslide starting points in agricultural and other open land (the mean slope for all landslide starting points is 31.8 degrees). These observations are in line with the pioneering Swiss study suggesting fewer landslides and higher mean slope angles on forested slopes compared to open land (Rickli and Graf, 2009). 65% of starting

https://doi.org/10.5194/egusphere-2026-579
Preprint. Discussion started: 2 March 2026

points occurred in the forest, while only 50% of the mask used for control point selection (Fig. 7) is forest-covered, suggesting a slight overrepresentation of landslide starting points in the forest. However, the ML results clearly suggest that slope angles below 30 degrees contribute negatively to landslide susceptibility while slope angles above 30 and up to 50 degrees increase landslide probability in the forest. This key difference to the full models shows that trees efficiently hinder initiation of shallow landslides in gentler terrain.

Aspect has a slightly different effect in the forest than in the full model. While the variable northerness is less influential in the forest model 3c than full model 1c, it is opposite for the variable easterness. As argued above, south- and to a smaller degree west-facing slopes may be more susceptible to shallow landslide initiation due to more active weathering processes and thicker, more stratified soils. It seems that the existence of tree cover makes south-facing slopes slightly less susceptible and west-facing slopes slightly more susceptible to landslide release compared to non-forested slopes. The study area is located in the southern polar region where temperature is often the limiting factor for growth in the forest ecosystem and where highest species diversities have been found on south-facing slopes (Heikkinen, 1991; Parker, 1988). Therefore, forests with higher species diversities may counteract the weathering effects on soil thickness and characteristics by providing mechanical and hydrological reinforcement. Future analyses of non-event specific inventories are recommended to strengthen the argument that higher thresholds for slope and a less pronounced weathering-related effect of aspect can generally be expected in forests in polar climates. If so, the existence of forest and selected forest-specific variables (e.g. tree type, canopy cover) and threshold values (e.g. slope) could become an integral part of an updated Norwegian shallow landslide susceptibility map, potentially transitioning from a knowledge-based (Fischer et al. 2012; Pullarello et al. 2025) to data-driven shallow landslide susceptibility assessment. On a general note, this study demonstrates that forests can reduce shallow landslide risk, highlighting the need for an integral forest management as part of nature-based solutions for mitigation.

6.4 Synthesis and outlook

The ranking of landslide explanatory factors in the compromise spatial models with small block CV (the full model 1c and forest-only model 3c) is comparable to the majority of data-driven susceptibility assessments (Lima et al. 2022). However, huge differences in availability, resolution and potential biases of the included datasets make a direct comparison extremely difficult. Moreover, model performances have been shown to vary greatly depending on landslide point sampling strategies, ranging from distributed points or cells throughout the entire landslide path or the release area, centroids for whole landslide or release area to points at landslide vicinity (Alvioli et al., 2016; Bordoni et al., 2020; Hong et al., 2019). In contrast, the sampling of non-landslide points does not commonly receive much attention but has been shown to have significant influence on model outcome as well (Dou et al. 2020; Hong et al. 2019). While 88% of recent publications on data-driven susceptibility maps find slope angle to be amongst the top predictors (Lima et al. 2022), in this study the influence of slope is moderately low for the full models as control points were sampled within observed slope ranges for landslide starting points. Nevertheless, slope plays a more decisive role in the forest-only models illustrating the importance of including forest

https://doi.org/10.5194/egusphere-2026-579
Preprint. Discussion started: 2 March 2026

variables in data-driven shallow landslide susceptibility modelling (Moos et al. 2016) and, more broadly, underscoring the potential in managing forest to mitigate shallow landslides risk (Teich et al., 2022). Aspect and rock type are commonly considered in data-driven landslide susceptibility studies and amongst the top 15 predictors in 73% of recent studies, while elevation is only used and found to be important by 47% of recent studies (Lima et al. 2022). While the significance of bedrock can clearly be attributed to differences in weatherability and regolith production rates (Hakim et al. 2021; Momeni et al. 2015), the significance of elevation and aspect is more open to interpretation. In different contexts, they may be proxies for a range of variables which influence slope stability, including distribution of temperature, precipitation, weathering and vegetation (Bordoni et al. 2020; Fabbri et al., 2003; Catani et al. 2013; Quevedo et al. 2022; Sidle and Ochiai, 2006). In this study, elevation, aspect and bedrock weatherability are suggested to be indicative of soil characteristics which are not sufficiently represented in the existing categorical dataset of deposit thickness. Soil thickness and heterogeneity exert a key control on shallow landslide initiation (Bogaard and Greco, 2016; Sidle and Bogaard, 2016), but are generally difficult to access and map.

Furthermore, we have shown that ML results are highly sensitive to cross-validation techniques and spatial modelling choices, and that testing of autocorrelation of input data and model residuals should become an integral part of landslide susceptibility modelling approaches. The spatial autocorrelation in landslide inventories most likely reflects incorrect, incomplete and missing datasets for explanatory factors. However, even given perfect representation of known drivers of slope instability, it is likely that there will remain a spatial component in the data that needs to be accounted for. In this study, the most stringent spatial clustering approach (models 1a and 3b) and the more relaxed compromise models (model 1c and model 3c) gave similar results in terms of predictor importance ranking, but the former were barely better than random guessing, probably because too much of the spatial variance was explained by the spatial component in the model. We therefore conclude that although the compromise models with small block nested CV did not always fully account for spatial autocorrelation, they provide a satisfactory middle ground between overfitting and underfitting. Interestingly, it was not sufficient to only model the spatial structure with random effects to address the spatial autocorrelation issue, but hyperparameter tuning also required the spatial structure to be accounted for through spatial clustering of the folds for cross-validation. This process is gaining traction in shallow landslide susceptibility mapping, particularly with the use of Generalised Additive Modelling (GAM; Brenning, 2005; Goetz et al., 2015; Knevels et al., 2023; Steger et al., 2016; Steger et al., 2021). To our knowledge, fewer studies have modelled the spatial structure directly (Knevels et al., 2023; Steger et al., 2021), also this in the GAM framework. The novelty in our study lies in addressing the popularity of extreme gradient boosting ML methods and suggesting a similarly robust approach in this space. However, this analysis is based on one dataset collected during an unusual weather event, and we recommend that future studies consider exploring, adapting and testing this methodological approach further afield. With further testing, future studies can then begin to develop a more reliable machine learning protocol for landslide susceptibility. We recommend that future landslide susceptibility studies employing ML methods: (1) test for initial and residual spatial autocorrelation, (2) explicitly model spatial structure, (3) explore and employ several spatial cross-validation techniques for both hyperparameter optimisation and model evaluation,

https://doi.org/10.5194/egusphere-2026-579
Preprint. Discussion started: 2 March 2026

and (4) report complete methodological details to enable reproducibility and proper interpretation of results. Failure to account for spatial autocorrelation through appropriate ML protocols may lead to over-optimistic model performance and misleading interpretations of predictor importance rankings due to spatial confounding. These findings do not invalidate previous endeavours but highlight that performance evaluation and predictor importance from ML studies should be interpreted cautiously if spatial structure has not been explicitly addressed in method descriptions, particularly regarding model generalisability to new spatial locations or future landslide probabilities.

7 Conclusions

In this study, we explore which explanatory factors best explain shallow landslides triggered during the 'Hans' storm on 7 to 10 August 2023. 571 shallow landslide starting points, from the refined version of the published event inventory (Rüther et al. 2024) and 571 control points sampled in the attribute space for slope, elevation and relative precipitation were compiled with data for 20 attributes considered as explanatory factors. Based on this dataset (reduced to 570 landslide and 568 control points due to missing data), we implemented a systematic comparison of 32 gradient boosted decision tree model runs to assess the impact of: (1) the hyperparameter optimisation approach (nested CV presented in this article vs. simple CV in Appendix B), (2) spatial structure incorporation (spatial vs. non-spatial), (3) cross-validation strategy (spatial clustering, large block, small block and random) and (4) analyses performed on full dataset vs. forest-only dataset. We find that the spatial gradient boosted models with a small block sampling strategy for cross-validation (models 1c and 3c) strike a suitable balance between predictive power and how well the spatial structure in the landslide data is accounted for. Based on analyses of the results of the full model 1c and forest-only model 3c, we conclude with the following:

- The importance of elevation and aspect shows that the current seamlessly available soil thickness dataset does not resolve soil thickness, let alone soil heterogeneity sufficiently. Rather, elevations under 480 meters a.s.l. as well as south- and west-facing slopes increase landslide probability and are probably related to anthropogenic influences at low elevations, increased weathering rates and thicker and/or more heterogeneous soils.
- The contributions of the various hydrometeorological factors are in line with expectations, showing that the combined effect of water received from precipitation and runoff (flow accumulation above 5000 m^2) determines whether critical pore pressure for shallow landslide release are exceeded. Whilst the analysis suffers from the coarse nature of the meteorological datasets (1 km), it also shows somewhat unexpected results as relative precipitation ('Hans' precipitation relative to normal August precipitation) is more influential than both total 'Hans' precipitation and precipitation intensity.
- Bedrock weatherability explains why few landslides were released in the southeastern part of the study area, which received amongst the highest amounts of precipitation and exceptionally high precipitation intensities, as it consists mostly of slowly weathering bedrock (weatherability class 4: sandstones, conglomerates and felsic volcanic rocks and metamorphic rocks).
- Shallow landslide release in forested areas requires higher slope angles compared to all landslide starting points in the full models suggesting that forested area should be treated differently from non-forested areas in the national landslide susceptibility map. These results also demonstrate forest's capability as nature-based solution in mitigating against shallow landslide release.

https://doi.org/10.5194/egusphere-2026-579
Preprint. Discussion started: 2 March 2026

- In tree-covered areas, shallow landslide susceptibility is highest in deciduous forest, which is probably related to the occurrence of birch-dominated forest with relatively low canopy cover towards the tree line and on slopes which are not suited as production forest. Finally, the effect of aspect with increased landslide probabilities on south-facing slopes is less pronounced on forested slopes suggesting that increased species diversity and growth rates counteract the weathering-related negative effects on soils.

Overall, whilst this study explores event-specific explanatory factors for shallow landslide release, the employed methodology may lead a way forward from the current Norwegian knowledge-based towards a data-driven susceptibility assessment for shallow landslides which may also be applicable in other countries with similar climate and topography.

https://doi.org/10.5194/egusphere-2026-579
Preprint. Discussion started: 2 March 2026

Appendix A

Table A1 Reclassification of bedrock type codes from the National bedrock map into five bedrock weatherability classes. The English names for the main bedrock type codes (*Kode*, fourth column) can be looked up under *Rocktype* (the last column) in the following table: presentasjonsregler-berggrunn-hovedbergart.pdf

Bedrock weatherability class	Main bedrock type code (hovedbergart)
1: very easily (carbonous rocks)	1, 144, 320, 321, 322, 405, 406, 415, 416
2: easily (shale, phyllite)	301, 302, 303, 401, 402, 403, 404, 407, 408, 410, 411, 412, 427, 470
3: relatively easily (mafic plutonic rocks)	108, 109, 112, 113, 114, 115, 116, 129, 130, 131, 132, 133, 134, 135, 136, 137, 220, 221, 222, 240, 242, 243, 340, 450, 451, 452, 453, 454, 456, 457
4: slowly (sandstones, conglomerates, felsic volcanic rocks, i.e. granites and metamorphic rocks, i.e. gneiss)	100, 101, 102, 103, 104, 105, 106, 107, 110, 111, 120, 139, 140, 141, 142, 143, 150, 151, 160, 200, 201, 202, 203, 204, 210, 211, 212, 213, 214, 230, 241, 300, 304, 305, 306, 307, 310, 311, 312, 350, 400, 409, 420, 421, 422, 424, 426, 428, 429, 430, 431, 432, 433, 434, 435, 437, 440, 441, 442, 455, 471, 472, 473
5: very slowly (quartzite and quartz sandstones)	308, 330, 423, 425, 460, 461

Table A2 Reclassification of surficial deposit codes from the National surficial deposit map into five classes based on lithological similarities and class occurrences in the dataset. *Code names* and *Definition/Description* for the surficial deposit codes can be found in this table: SuperficialDepositSurfaceType_codelist.pdf.

Surficial deposit (Losmassetype) reclassification	Surficial deposit codes (losmassetype)
1: till	11,12, 13, 14, 90
2: (glaci)fluvial, beach	20, 30, 43, 50
3: weathered material and anthropogenic	70, 71, 73, 120
4: slope processes	81, 82
8: thin organic, bare rock	100, 130

https://doi.org/10.5194/egusphere-2026-579
Preprint. Discussion started: 2 March 2026

Appendix B

4 Machine learning methods

Since the *simple CV* approach is a less computer intensive method of CV that may be employed in recent ML-supported shallow landslide susceptibility literature (it is often difficult to discern as methods are not always well described), we tested this approach to provide comparison to the *nested CV* approach presented in the main body of the article.

4.1 Data partitioning approach

In contrary to the nested CV approach, the simple CV approach maximises data utilisation and assesses within-study-area performance, because all 100% of observations are used for cross-validation without a separate hold-out test set (Fig. B1). In our analysis, hyperparameters were optimised using the full dataset with each CV fold clustering strategy as described in the main paper. Final models were then retrained on all available data using optimal parameters. Performance metrics from cross-validation represent expected generalisation performance, while evaluation on the complete training dataset reflects interpolation accuracy within the study area. The lack of a final "Test data AUC" metric, however, prevents an assessment of model performance in other areas.

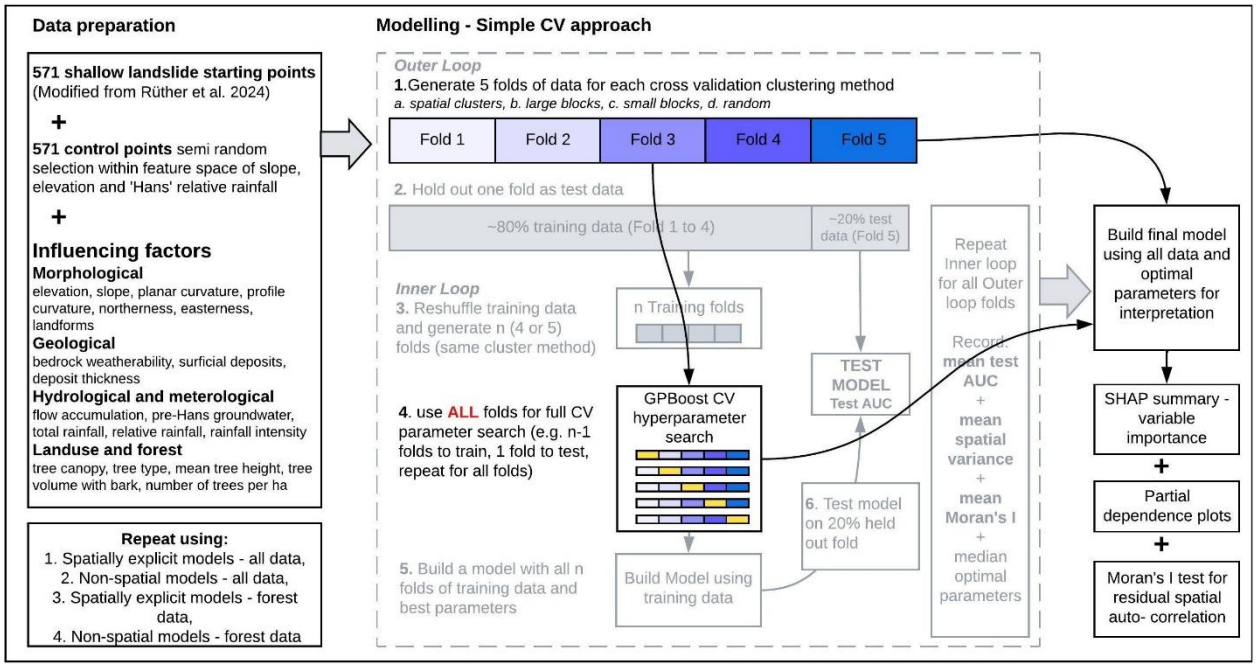


Figure B1 Flowchart summarising data and methods for the additional simple CV approach presented here in Appendix B

https://doi.org/10.5194/egusphere-2026-579
Preprint. Discussion started: 2 March 2026

Table B1 Mean (and standard deviation) model performance metrics of two Gaussian Process Gradient Boosted classification models using four different cross-validation (CV) methods and two CV strategies (nested CV and simple CV) on the full landslide dataset. AUC = area under the curve, where values close to 1 indicate near perfect accuracy of landslide/control point classification. Mean CV AUC = mean value from hyperparameter search (Step 4 in Fig. 2 and B1); Mean training AUC = mean value from inner loop training models (Step 5 in Fig. 2 and Final model in Fig. B1); Mean test AUC = mean value from testing inner loop models (Step 6 in Fig. 2). This value is in bold as it is the most useful metric for evaluating performance, but it is absent in the simple CV approach. Spatial variance indicates the amount of variance explained by the spatial structure of the model (absent from non-spatial models). The significance of the Moran's I test for spatial autocorrelation is indicated by asterisks (*** = $p < 0.001$, ** = $p < 0.01$, * = $p < 0.05$, ns = non-significant). A significant Moran's I indicates that residual spatial autocorrelation is likely. Note there are no standard deviation values for the simple CV, as the hyperparameter optimisation was completely internally and only one final model was constructed.

Model & CV type	Mean CV AUC	Mean training AUC	Mean test AUC	Spatial variance	Moran's I
Nested CV approach (presented in main article)					
1. Spatial models					
1a. Spatially clustered CV	0.64 (0.05)	0.98 (0.01)	**0.64 (0.06)**	0.70	0.02 *
1b. Large block CV	0.61 (0.03)	0.97 (0.01)	**0.57 (0.10)**	0.71	0.01 (ns)
1c. Small block CV	0.67 (0.05)	0.98 (0.01)	**0.77 (0.06)**	0.68	0.02 (ns)
1d. Random CV	0.86 (0.01)	1 (0)	**0.90 (0.01)**	0.47	0.18 ***
2. Non-spatial models					
2a. Spatially clustered CV	0.67 (0.02)	0.99 (0.01)	**0.68 (0.08)**	-	0.11 ***
2b. Large block CV	0.69 (0.03)	0.97 (0.03)	**0.67 (0.09)**	-	0.13 ***
2c. Small block CV	0.75 (0.03)	0.96 (0.07)	**0.71 (0.13)**	-	0.15 ***
2d. Random CV	0.88 (0.01)	0.99 (0.01)	**0.87 (0.02)**	-	0.08 ***
Simple CV approach (presented in Appendix B)					
S1. Spatial models					
S1a.Spatially clustered CV	0.631	0.969	-	0.73	0.01 (ns)
S1b. Large block CV	0.630	0.978	-	0.68	0.03 **
S1c. Small block CV	0.667	0.977	-	0.69	0.02 *
S1d. Random CV	0.845	0.999	-	0.53	0.13 ***
S2. Non-spatial models					
S2a. Spatially clustered CV	0.687	0.999	-	-	0.10 ***
S2b. Large block CV	0.708	0.999	-	-	0.11 ***
S2c. Small block CV	0.772	0.999	-	-	0.11 ***
S2d. Random CV	0.883	0.999	-	-	0.11 ***

https://doi.org/10.5194/egusphere-2026-579
Preprint. Discussion started: 2 March 2026

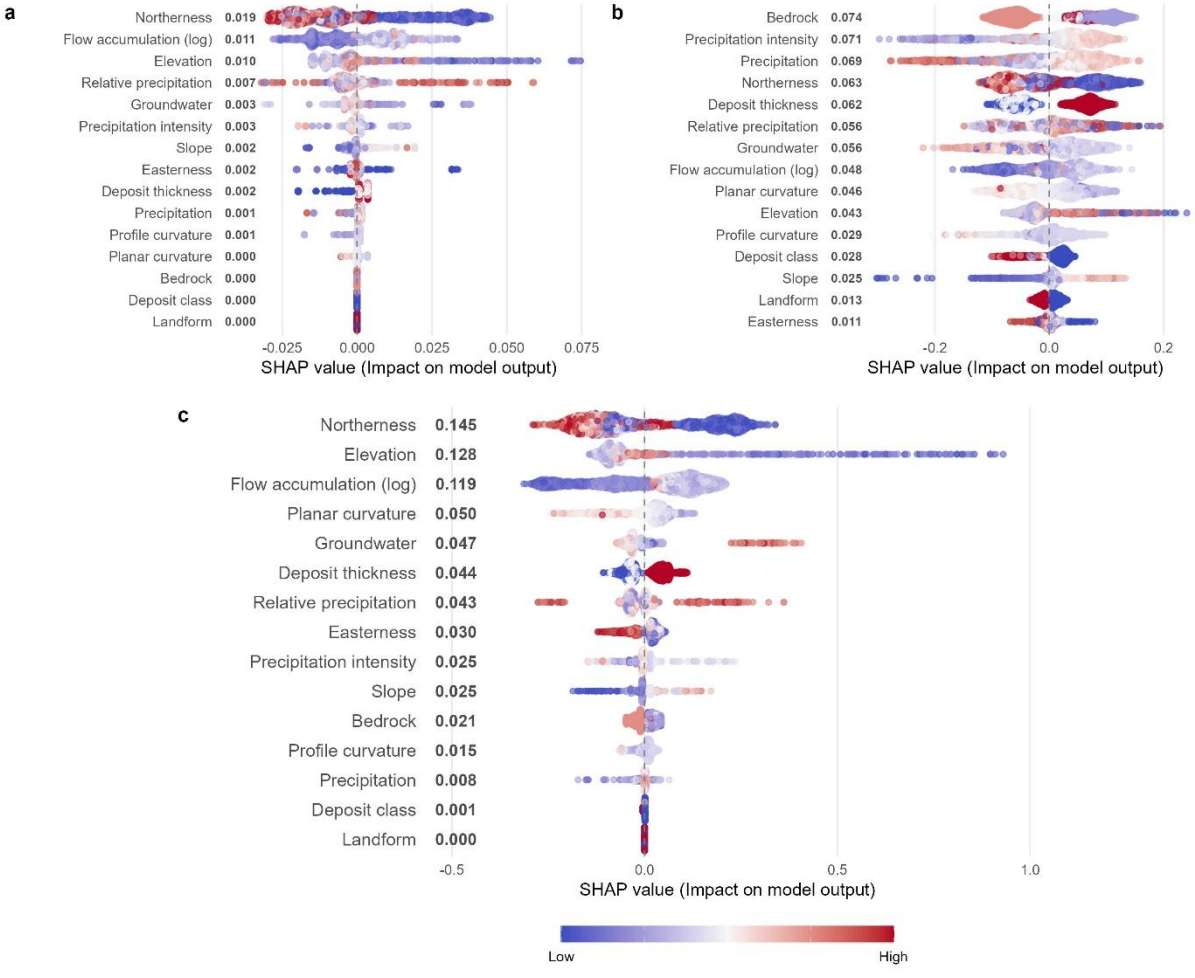


Figure B2 Shapley value summary plots for three simple CV models on the full dataset: a) spatial model S1a with spatially clustered CV, b) non-spatial model S2d with random CV, c) spatial model S1c with small block CV. Plots a) and b) represent the two extremes of underfitting and overfitting while c) represents a compromise. The bold numbers after variable names are mean absolute SHAP values, representing relative importance. Points are coloured by normalised feature value.

https://doi.org/10.5194/egusphere-2026-579
Preprint. Discussion started: 2 March 2026

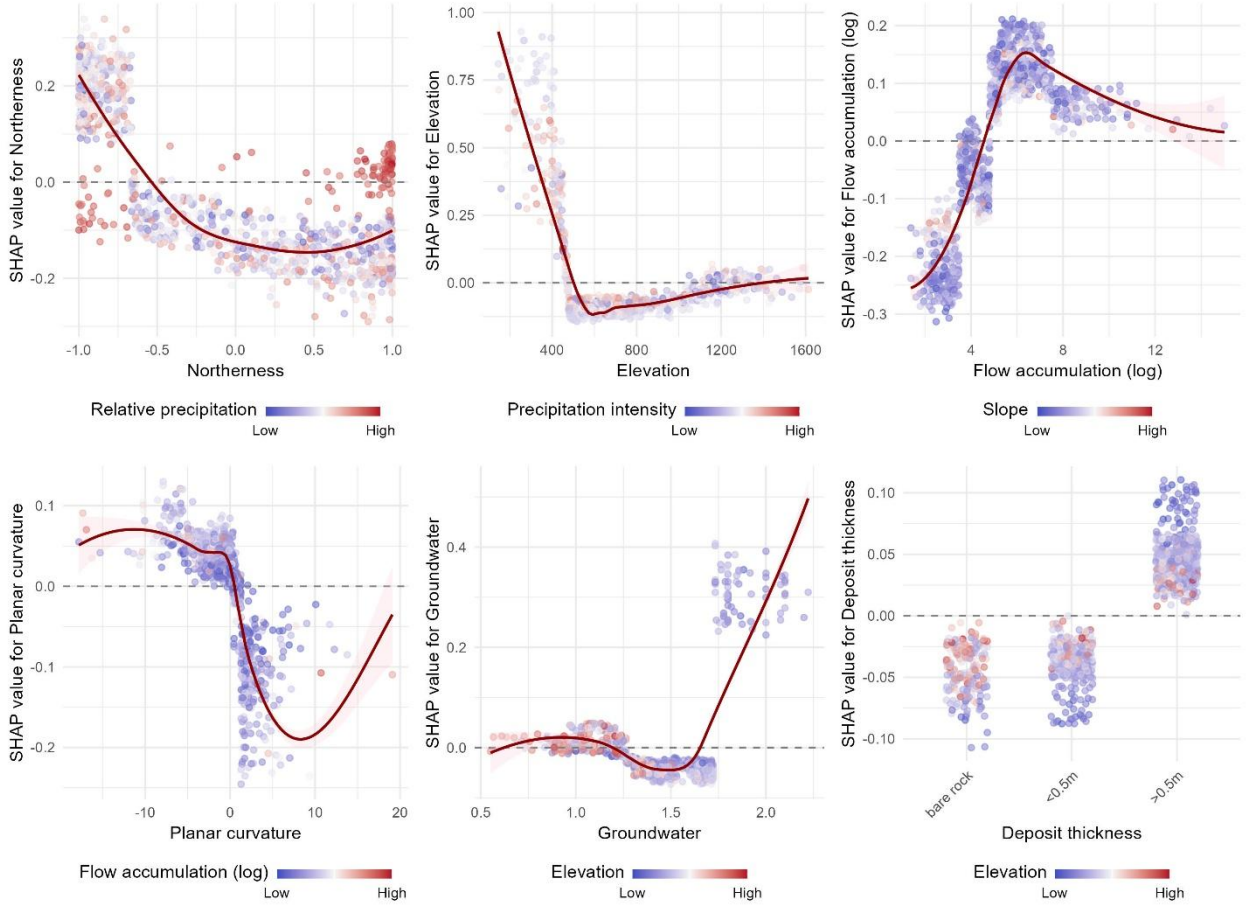


Figure B3 Partial dependence plots for the top six variables of the spatial model S1c with small block CV on the full dataset, using the simple CV approach. Each point represents an observation and is plotted with its feature value on the x-axis, and SHAP value (contribution to model output) on the y-axis. SHAP values above 0 indicate that a feature value is likely to make a positive contribution to the probability of a landslide. Points are coloured by a second feature chosen to minimise the variance of the SHAP value given both features and represents the strongest interaction. Solid red lines are LOESS smoothed lines.

https://doi.org/10.5194/egusphere-2026-579
Preprint. Discussion started: 2 March 2026

Table B2 Mean (and standard deviation) model performance metrics of two Gaussian Process Gradient Boosted classification models using four different cross-validation (CV) methods and two CV strategies (nested CV and simple CV) on the forest-only landslide dataset. AUC = area under the curve, where values close to 1 indicate near perfect accuracy of landslide/control point classification. Mean CV AUC = mean value from hyperparameter search (Step 4 in Fig. 2 and B1); Mean training AUC = mean value from inner loop training models (Step 5 in Fig. 2 and Final model in Fig. B1); Mean test AUC = mean value from testing inner loop models (Step 6 in Fig. 2). This value is in bold as it is the most useful metric for evaluating performance, but it is absent in the simple CV approach. Spatial variance indicates the amount of variance explained by the spatial structure of the model (absent from non-spatial models). The significance of the Moran's I test for spatial autocorrelation is indicated by asterisks (*** = $p < 0.001$, ** = $p < 0.01$, * = $p < 0.05$, ns = non-significant). A significant Moran's I indicates that residual spatial autocorrelation is likely. Note there are no standard deviation values for the simple CV, as the hyperparameter optimisation was completely internally and only one final model was constructed.

Model & CV type	Mean CV AUC	Mean training AUC	Mean test AUC	Spatial variance	Moran's I
Nested CV approach (presented in main article)					
3. Spatial models					
3a. Spatially clustered CV	0.64 (0.04)	0.99 (0.01)	**0.71 (0.10)**	0.63	0.08 ***
3b. Large block CV	0.61 (0.04)	0.96 (0.01)	**0.56 (0.05)**	0.70	0.01 (ns)
3c. Small block CV	0.68 (0.05)	0.98 (0.01)	**0.73 (0.10)**	0.68	0.03 **
3d. Random CV	0.86 (0.02)	0.99 (0.00)	**0.89 (0.03)**	0.43	0.16 ***
4. Non-spatial models					
4a. Spatially clustered CV	0.67 (0.04)	0.99 (0.01)	**0.70 (0.08)**		0.13 ***
4b. Large block CV	0.68 (0.02)	0.94 (0.06)	**0.62 (0.02)**		0.23 ***
4c. Small block CV	0.75 (0.03)	0.99 (0.00)	**0.79 (0.02)**		0.11 ***
4d. Random CV	0.88 (0.02)	0.99 (0.01)	**0.88 (0.02)**		0.10 ***
Simple CV approach (presented in Appendix B)					
S3. Spatial models					
S3a. Spatially clustered CV	0.649	0.997	-	0.62	0.10 ***
S3b. Large block CV	0.637	0.976	-	0.70	0.02 *
S3c. Small block CV	0.727	0.993	-	0.63	0.08 ***
S3d. Random CV	0.861	0.999	-	0.54	0.16 ***
S4. Non-spatial models					
S4a. Spatially clustered CV	0.716	0.999	-	-	0.11 ***
S4b. Large block CV	0.707	0.849	-	-	0.42 ***
S4c. Small block CV	0.776	0.999	-	-	0.11 ***
S4d. Random CV	0.880	0.992	-	-	0.13 ***

https://doi.org/10.5194/egusphere-2026-579
Preprint. Discussion started: 2 March 2026

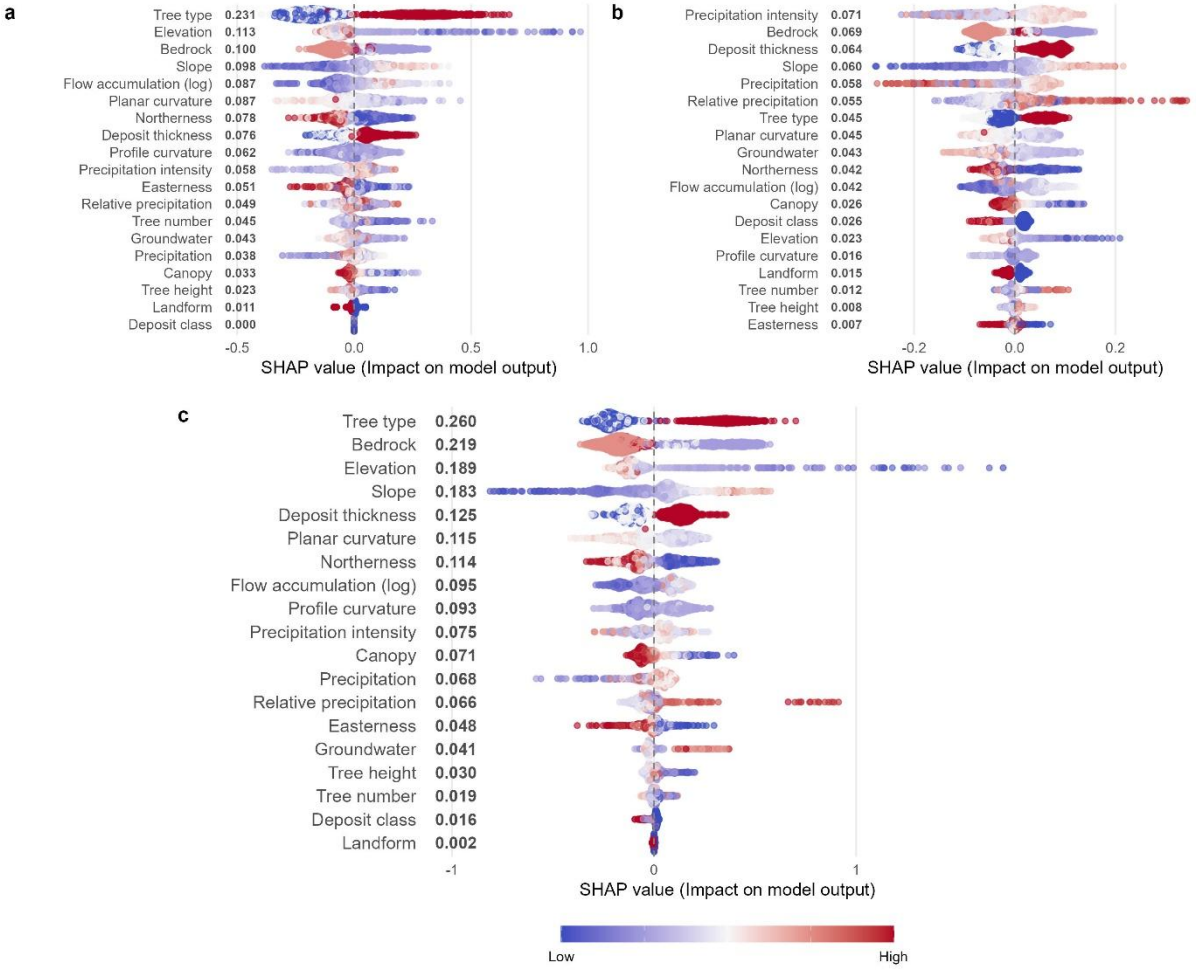


Figure B4 Shapley value summary plots for three simple CV models on the forest-only dataset: a) spatial model S3a with spatially clustered CV, b) non-spatial model S4d with random CV, c) spatial model S3c with small block CV. Plots a) and b) represent the two extremes of underfitting and overfitting while c) represents a compromise. The bold numbers after variable names are mean absolute SHAP values, representing relative importance. Points are coloured by normalised feature value.

https://doi.org/10.5194/egusphere-2026-579
Preprint. Discussion started: 2 March 2026

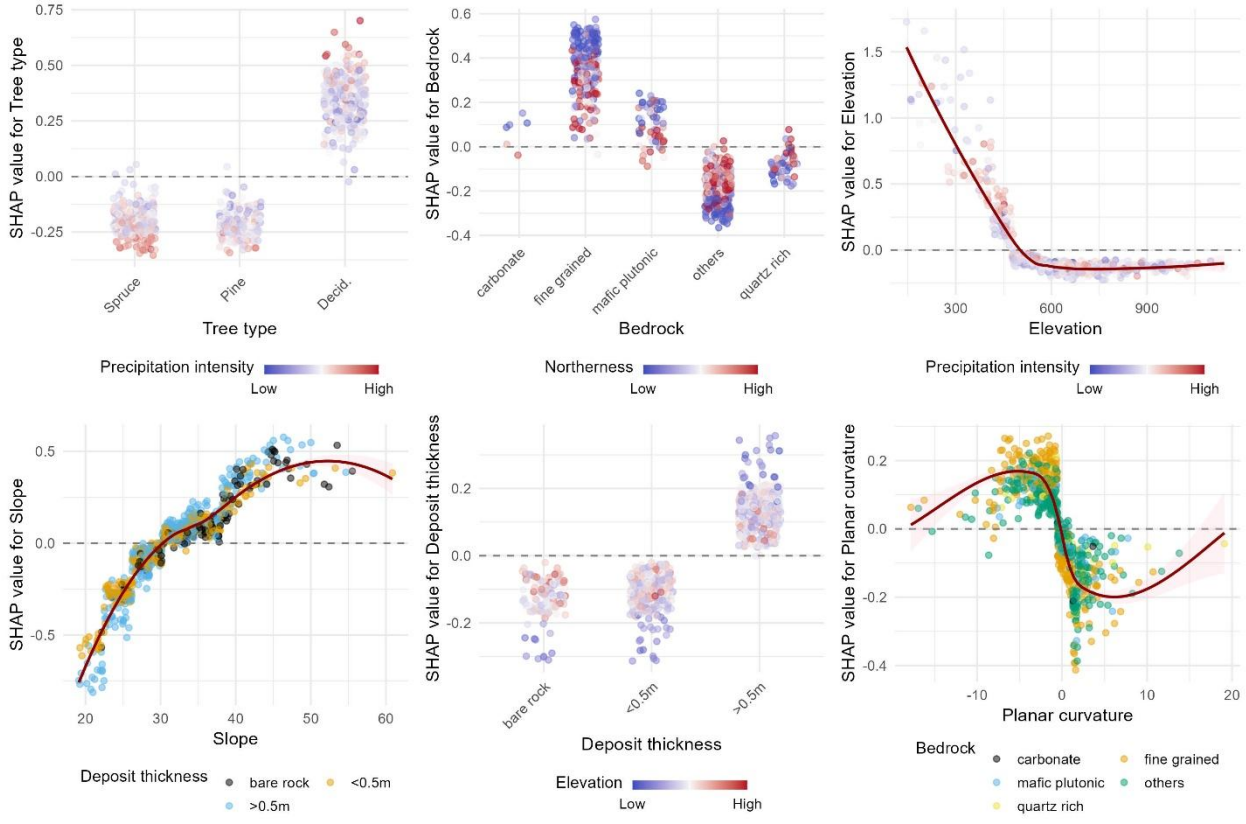


Figure B5 Partial dependence plots for the top six variables of the spatial model S3c with small block CV on the forest-only dataset, using the simple CV approach. Each point represents an observation and is plotted with its feature value on the x-axis, and SHAP value (contribution to model output) on the y-axis. SHAP values above 0 indicate that a feature value is likely to make a positive contribution to the probability of a landslide. Points are coloured by a second feature chosen to minimise the variance of the SHAP value given both features and represents the strongest interaction. Solid red lines are LOESS smoothed lines.

https://doi.org/10.5194/egusphere-2026-579
Preprint. Discussion started: 2 March 2026

Code and data availability

The analysed dataset can be obtained by contacting the corresponding author. The Python code used for data preparation and R code used for machine learning modelling will be published and available if the paper gets accepted for publication.

Author contributions

DR: Conceptualisation, Data collection & curation, Analysis, Funding acquisition, Methodology, Validation, Visualisation, Writing - original draft, review and editing. **KH:** Conceptualisation, Data collection & curation, Methodology, Validation, Writing - review and editing. **IP:** Data collection, Writing - review and editing. **MG:** Conceptualisation, Analysis, Methodology, Validation, Visualisation, Writing – original draft, review and editing.

Competing interest

The authors declare that they have no conflict of interest.

Disclaimer

Copernicus Publications remains neutral with regard to jurisdictional claims made in the text, published maps, institutional affiliations, or any other geographical representation in this paper. While Copernicus Publications makes every effort to include appropriate place names, the final responsibility lies with the authors. Views expressed in the text are those of the authors and do not necessarily reflect the views of the publisher.

Acknowledgements

Prof. Deta Gasser has assisted in the categorisation into bedrock weatherability classes. Jess Joar Andersen (NVE) kindly provided the relevant modelled groundwater data and Tuomo Mikael Saloranta (NVE) helped finding the relevant precipitation intensity dataset from MET Nordic. Further, John Paul McLean (NIBIO) provided guidance in the evaluation of SR16 datasets, while Åsbjørn Flacké Haualand contributed to the development of the control point sampling strategy. Parts of the data analysis were performed at the Erlangen National High Performance Computing Center of the Friedrich-Alexander-Universität Erlangen-Nürnberg.

Financial support

This work is supported by a grant from the Norwegian Research Centre on Sustainable Climate Change Adaptation (Noradapt).

https://doi.org/10.5194/egusphere-2026-579
Preprint. Discussion started: 2 March 2026

Review statement

The review statement will be added by Copernicus Publications listing the handling editor as well as all contributing referees according to their status anonymous or identified.

https://doi.org/10.5194/egusphere-2026-579
Preprint. Discussion started: 2 March 2026

References

Ageenko, A., Hansen, L. C., Lyng, K. L., Bodum, L., and Arsanjani, J. J.: Landslide susceptibility mapping using machine learning: a Danish case study, ISPRS International Journal of Geo-Information, 11, 324, 2022.

Albano, R. and Sole, A.: Geospatial methods and tools for natural risk management and communications, ISPRS International Journal of Geo-Information, 7, 470, 2018.

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

Arabameri, A., Chandra Pal, S., Rezaie, F., Chakrabortty, R., Saha, A., Blaschke, T., Di Napoli, M., Ghorbanzadeh, O., and Thi Ngo, P. T.: Decision tree based ensemble machine learning approaches for landslide susceptibility mapping, Geocarto International, 37, 4594–4627, 10.1080/10106049.2021.1892210, 2022.

Astrup, R., Rahlf, J., Bjørkelo, K., Debella-Gilo, M., Gjertsen, A.-K., and Breidenbach, J.: Forest information at multiple scales: development, evaluation and application of the Norwegian forest resources map SR16, Scandinavian Journal of Forest Research, 34, 484–496, 10.1080/02827581.2019.1588989, 2019.

Baltensweiler, A., Walthert, L., Hanewinkel, M., Zimmermann, S., and Nussbaum, M.: Machine learning based soil maps for a wide range of soil properties for the forested area of Switzerland, Geoderma Regional, 27, e00437, https://doi.org/10.1016/j.geodrs.2021.e00437, 2021.

Beldring, S., Engeland, K., Roald, L. A., Sælthun, N. R., and Voksø, A.: Estimation of parameters in a distributed precipitation-runoff model for Norway, Hydrology and Earth System Sciences, 7, 304–316, 2003.

Bergström, S. and Lindström, G.: Interpretation of runoff processes in hydrological modelling—experience from the HBV approach, Hydrological Processes, 29, 3535–3545, https://doi.org/10.1002/hyp.10510, 2015.

Bischl, B., Mersmann, O., Trautmann, H., and Weihs, C.: Resampling methods for meta-model validation with recommendations for evolutionary computation, Evolutionary Computation, 20, 249–275, 10.1162/EVCO_a_00069, 2012.

Bivand, R. S. and Wong, D. W. S.: Comparing implementations of global and local indicators of spatial association, TEST, 27, 716–748, 10.1007/s11749-018-0599-x, 2018.

Bogaard, T. A. and Greco, R.: Landslide hydrology: from hydrology to pore pressure, Wiley Interdisciplinary Reviews: Water, 3, 439–459, https://doi.org/10.1002/wat2.1126, 2016.

Bondevik, S. and Sorteberg, A.: Groundwater fluctuations during a debris flow event in western Norway–triggered by rain and snowmelt, Hydrology and Earth System Sciences, 25, 4147–4158, https://doi.org/10.5194/hess-25-4147-2021, 2021.

Bordoni, M., Galanti, Y., Bartelletti, C., Persichillo, M. G., Barsanti, M., Giannecchini, R., Avanzi, G. D. A., Cevasco, A., Brandolini, P., Galve, J. P., and Meisina, C.: The influence of the inventory on the determination of the rainfall-induced shallow landslides susceptibility using generalized additive models, CATENA, 193, 104630, https://doi.org/10.1016/j.catena.2020.104630, 2020.

Breidenbach, J., Waser, L. T., Debella-Gilo, M., Schumacher, J., Rahlf, J., Hauglin, M., Puliti, S., and Astrup, R.: National mapping and estimation of forest area by dominant tree species using Sentinel-2 data, Canadian Journal of Forest Research, 51, 365–379, 10.1139/cjfr-2020-0170, 2021.

Brenning, A.: Spatial prediction models for landslide hazards: review, comparison and evaluation, Nat. Hazards Earth Syst. Sci., 5, 853–862, 10.5194/nhess-5-853-2005, 2005.

Cao, J., Zhang, Z., Du, J., Zhang, L., Song, Y., and Sun, G.: Multi-geohazards susceptibility mapping based on machine learning—a case study in Jiuzhaigou, China, Natural Hazards, 102, 851–871, 10.1007/s11069-020-03927-8, 2020.

Capobianco, V., Fraccica, A., Anselmucci, F., and Tagarelli, V.: Nature- and bio-based solutions for ecosystem restoration, landslide hazard mitigation, and ground improvement: Research and application novelties, Ecological Engineering, 219, 107710, https://doi.org/10.1016/j.ecoleng.2025.107710, 2025.

Cascini, L., Ciurleo, M., and Di Nocera, S.: Soil depth reconstruction for the assessment of the susceptibility to shallow landslides in fine-grained slopes, Landslides, 14, 459–471, 10.1007/s10346-016-0720-8, 2017.

https://doi.org/10.5194/egusphere-2026-579
Preprint. Discussion started: 2 March 2026

Catani, F., Lagomarsino, D., Segoni, S., and Tofani, V.: Landslide susceptibility estimation by random forests technique: sensitivity and scaling issues, Nat. Hazards Earth Syst. Sci., 13, 2815–2831, 10.5194/nhess-13-2815-2013, 2013.

Cawley, G. C. and Talbot, N. L.: On over-fitting in model selection and subsequent selection bias in performance evaluation, The Journal of Machine Learning Research, 11, 2079–2107, 2010.

Choubin, B., Jaafari, A., and Mafi-Gholami, D.: A spatially explicit multi-hazard framework for assessing flood, landslide, wildfire, and drought susceptibilities, Advances in Space Research, 75, 2569–2583, https://doi.org/10.1016/j.asr.2024.11.005, 2025.

Church, M. and Jakob, M.: What is a debris flood?, Water Resources Research, 56, e2020WR027144, https://doi.org/10.1029/2020WR027144, 2020.

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

Crozier, M. J.: Deciphering the effect of climate change on landslide activity: A review, Geomorphology, 124, 260–267, 2010.

Cámara, J., Gómez-Miguel, V., and Martín, M. Á.: Lithologic control on soil texture heterogeneity, Geoderma, 287, 157–163, https://doi.org/10.1016/j.geoderma.2016.09.006, 2017.

D'Odorico, P.: A possible bistable evolution of soil thickness, Journal of Geophysical Research: Solid Earth, 105, 25927–25935, https://doi.org/10.1029/2000JB900253, 2000.

Dai, Z. and Huang, W.: Examining the role of machine learning methods for ANN prediction of landslide coverage area, GeoJournal, 90, 134, 10.1007/s10708-025-11384-8, 2025.

de Jesús Arce-Mojica, T., Nehren, U., Sudmeier-Rieux, K., Miranda, P. J., and Anhuf, D.: Nature-based solutions (NbS) for reducing the risk of shallow landslides: Where do we stand?, International Journal of Disaster Risk Reduction, 41, 101293, https://doi.org/10.1016/j.ijdrr.2019.101293, 2019.

De Reu, J., Bourgeois, J., Bats, M., Zwertvaegher, A., Gelorini, V., De Smedt, P., Chu, W., Antrop, M., De Maeyer, P., Finke, P., Van Meirvenne, M., Verniers, J., and Crombé, P.: Application of the topographic position index to heterogeneous landscapes, Geomorphology, 186, 39–49, https://doi.org/10.1016/j.geomorph.2012.12.015, 2013.

Dietrich, W. E., Reiss, R., Hsu, M. L., and Montgomery, D. R.: A process-based model for colluvial soil depth and shallow landsliding using digital elevation data, Hydrological processes, 9, 383–400, 1995.

Dou, H., He, J., Huang, S., Jian, W., and Guo, C.: Influences of non-landslide sample selection strategies on landslide susceptibility mapping by machine learning, Geomatics, Natural Hazards and Risk, 14, 2285719, 10.1080/19475705.2023.2285719, 2023.

Dou, J., Yunus, A. P., Merghadi, A., Shirzadi, A., Nguyen, H., Hussain, Y., Avtar, R., Chen, Y., Pham, B. T., and Yamagishi, H.: Different sampling strategies for predicting landslide susceptibilities are deemed less consequential with deep learning, Science of The Total Environment, 720, 137320, https://doi.org/10.1016/j.scitotenv.2020.137320, 2020.

DSB: Evaluering av ekstremværet Hans - forebygging, beredskap og håndtering, Direktoratet for samfunnssikkerhet og beredskap (DSB), ISBN 978-82-7768-549-6, 2024.

Ekholt, C. T.: Derfor skaper ekstremværet Hans ekstreme tilstander på Østlandet, 2023.

Fabbri, A. G., Chung, C.-J. F., Cendrero, A., and Remondo, J.: Is prediction of future landslides possible with a GIS?, Natural Hazards, 30, 487–503, 10.1023/B:NHAZ.0000007282.62071.75, 2003.

Fan, C.-C. and Lai, Y.-F.: Influence of the spatial layout of vegetation on the stability of slopes, Plant and Soil, 377, 83–95, 10.1007/s11104-012-1569-9, 2014.

Fischer, L., Rubensdotter, L., Sletten, K., Stalsberg, K., Melchiorre, C., Horton, P., and Jaboyedoff, M.: Debris flow modeling for susceptibility mapping at regional to national scale in Norway, Proceedings of the 11th International and 2nd North American Symposium on Landslides, 3–8,

Fu, Z., Li, Z., Cai, C., Shi, Z., Xu, Q., and Wang, X.: Soil thickness effect on hydrological and erosion characteristics under sloping lands: A hydropedological perspective, Geoderma, 167-168, 41–53, https://doi.org/10.1016/j.geoderma.2011.08.013, 2011.

https://doi.org/10.5194/egusphere-2026-579
Preprint. Discussion started: 2 March 2026

Gariano, S. L. and Guzzetti, F.: 5.32 - Mass-movements and climate change, in: Treatise on Geomorphology (Second Edition), edited by: Shroder, J. F., Academic Press, Oxford, 546–558, https://doi.org/10.1016/B978-0-12-818234-5.00043-2, 2022.

Ghosh, S., van Westen, C. J., Carranza, E. J. M., and Jetten, V. G.: Integrating spatial, temporal, and magnitude probabilities for medium-scale landslide risk analysis in Darjeeling Himalayas, India, Landslides, 9, 371–384, 10.1007/s10346-011-0304-6, 2012.

Goetz, J. N., Brenning, A., Petschko, H., and Leopold, P.: Evaluating machine learning and statistical prediction techniques for landslide susceptibility modeling, Computers & Geosciences, 81, 1–11, https://doi.org/10.1016/j.cageo.2015.04.007, 2015.

Gomes, L. C., Beucher, A. M., Møller, A. B., Iversen, B. V., Børgesen, C. D., Adetsu, D. V., Sechu, G. L., Heckrath, G. J., Koch, J., Adhikari, K., Knadel, M., Lamandé, M., Greve, M. B., Jensen, N. H., Gutierrez, S., Balstrøm, T., Koganti, T., Roell, Y., Peng, Y., and Greve, M. H.: Soil assessment in Denmark: Towards soil functional mapping and beyond, Frontiers in Soil Science, Volume 3 - 2023, 10.3389/fsoil.2023.1090145, 2023.

Granerød, M., Stabell, D., Mjelstad, H., and Tajet, H. T. T.: Ekstremværet "Hans", ekstremt mye nedbør i deler av Sør-Norge 07.-09. august 2023, Oslo, 2023.

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

Guzzetti, F., Peruccacci, S., Rossi, M., and Stark, C. P.: The rainfall intensity–duration control of shallow landslides and debris flows: an update, Landslides, 5, 3–17, https://doi.org/10.1007/s10346-007-0112-1, 2008.

Hakim, K., Bower, D. J., Tian, M., Deitrick, R., Auclair-Desrotour, P., Kitzmann, D., Dorn, C., Mezger, K., and Heng, K.: Lithologic controls on silicate weathering regimes of temperate planets, The Planetary Science Journal, 2, 49, 10.3847/PSJ/abe1b8, 2021.

Han, Y. and Semnani, S. J.: Integration of physics-based and data-driven approaches for landslide susceptibility assessment, International Journal for Numerical and Analytical Methods in Geomechanics, 49, 3060–3097, https://doi.org/10.1002/nag.4016, 2025.

Haque, U., Blum, P., da Silva, P. F., Andersen, P., Pilz, J., Chalov, S. R., Malet, J.-P., Auflič, M. J., Andres, N., Poyiadji, E., Lamas, P. C., Zhang, W., Peshevski, I., Pétursson, H. G., Kurt, T., Dobrev, N., García-Davalillo, J. C., Halkia, M., Ferri, S., Gaprindashvili, G., Engström, J., and Keellings, D.: Fatal landslides in Europe, Landslides, 13, 1545–1554, 10.1007/s10346-016-0689-3, 2016.

Hastie, T.: The elements of statistical learning: data mining, inference, and prediction, 2009.

Haualand, K. F.: Code repository "Hans landslides": https://github.com/krifla/hans_landslide, 2025.

Hauglin, M., Rahlf, J., Schumacher, J., Astrup, R., and Breidenbach, J.: Large scale mapping of forest attributes using heterogeneous sets of airborne laser scanning and National Forest Inventory data, Forest Ecosystems, 8, 65, https://doi.org/10.1186/s40663-021-00338-4, 2021.

Heikkinen, R. K.: Multivariate analysis of esker vegetation in southern Häme, S Finland, Annales Botanici Fennici, 28, 201–224, 1991.

Hong, H., Miao, Y., Liu, J., and Zhu, A. X.: Exploring the effects of the design and quantity of absence data on the performance of random forest-based landslide susceptibility mapping, CATENA, 176, 45–64, https://doi.org/10.1016/j.catena.2018.12.035, 2019.

Huang, Y. and Zhao, L.: Review on landslide susceptibility mapping using support vector machines, CATENA, 165, 520–529, https://doi.org/10.1016/j.catena.2018.03.003, 2018.

Hungr, O., Leroueil, S., and Picarelli, L.: The Varnes classification of landslide types, an update, Landslides, 11, 167–194, 2014.

Jakob, M. and Owen, T.: Projected effects of climate change on shallow landslides, North Shore Mountains, Vancouver, Canada, Geomorphology, 393, 107921, https://doi.org/10.1016/j.geomorph.2021.107921, 2021.

Jiang, H., Zou, Q., Zhou, B., Jiang, Y., Cui, J., Yao, H., and Zhou, W.: Estimation of shallow landslide susceptibility incorporating the impacts of vegetation on slope stability, International Journal of Disaster Risk Science, 14, 618–635, 10.1007/s13753-023-00507-9, 2023.

https://doi.org/10.5194/egusphere-2026-579
Preprint. Discussion started: 2 March 2026

Johnson, K. A. and Sitar, N.: Hydrologic conditions leading to debris-flow initiation, Canadian Geotechnical Journal, 27, 789–801, https://doi.org/10.1139/t90-092, 1990.

Kim, J. H., Fourcaud, T., Jourdan, C., Maeght, J.-L., Mao, Z., Metayer, J., Meylan, L., Pierret, A., Rapidel, B., Roupsard, O., de Rouw, A., Sanchez, M. V., Wang, Y., and Stokes, A.: Vegetation as a driver of temporal variations in slope stability: The impact of hydrological processes, Geophysical Research Letters, 44, 4897–4907, https://doi.org/10.1002/2017GL073174, 2017.

Kim, M. S., Onda, Y., Kim, J. K., and Kim, S. W.: Effect of topography and soil parameterisation representing soil thicknesses on shallow landslide modelling, Quaternary International, 384, 91–106, https://doi.org/10.1016/j.quaint.2015.03.057, 2015.

Kim, Y., Rahardjo, H., and Tsen-Tieng, D. L.: Stability analysis of laterally loaded trees based on tree-root-soil interaction, Urban Forestry & Urban Greening, 49, 126639, https://doi.org/10.1016/j.ufug.2020.126639, 2020.

Kinzinger, L., Mach, J., Haberstroh, S., Schindler, Z., Frey, J., Dubbert, M., Seeger, S., Seifert, T., Weiler, M., Orlowski, N., and Werner, C.: Interaction between beech and spruce trees in temperate forests affects water use, root water uptake pattern and canopy structure, Tree Physiology, 44, 10.1093/treephys/tpad144, 2023.

Knevels, R., Petschko, H., Proske, H., Leopold, P., Mishra, A. N., Maraun, D., and Brenning, A.: Assessing uncertainties in landslide susceptibility predictions in a changing environment (Styrian Basin, Austria), Nat. Hazards Earth Syst. Sci., 23, 205–229, 10.5194/nhess-23-205-2023, 2023.

Kuhn, M.: Building predictive models in R using the caret package, Journal of Statistical Software, 28, 1 – 26, 10.18637/jss.v028.i05, 2008.

Kutschera, L. and Lichtenegger, E.: Wurzelatlas mitteleuropäischer Waldbäume und Sträucher, Leopold Stocker Verlag, Graz2002.

Li, R., Luo, J., Yanqiao, Z., Duan, Z., and Xuan, C.: The influence of non-landslide sample selection on susceptibility result: a case study in the Hanbing District, Shaanxi Province, China, Stochastic Environmental Research and Risk Assessment, 39, 5011–5028, 10.1007/s00477-025-03062-5, 2025.

Lima, P., Steger, S., and Glade, T.: Counteracting flawed landslide data in statistically based landslide susceptibility modelling for very large areas: a national-scale assessment for Austria, Landslides, 18, 3531–3546, 10.1007/s10346-021-01693-7, 2021.

Lima, P., Steger, S., Glade, T., and Murillo-García, F. G.: Literature review and bibliometric analysis on data-driven assessment of landslide susceptibility, Journal of Mountain Science, 19, 1670–1698, 10.1007/s11629-021-7254-9, 2022.

Lindsay, E., Frauenfelder, R., Rüther, D., Nava, L., Rubensdotter, L., Strout, J., and Nordal, S.: Multi-temporal satellite image composites in Google Earth Engine for improved landslide visibility: a case study of a glacial landscape, Remote Sensing, 14, 2301, https://doi.org/10.3390/rs14102301, 2022.

Linnenbrink, J., Milà, C., Ludwig, M., and Meyer, H.: kNNDM: k-fold Nearest Neighbour Distance Matching cross-validation for map accuracy estimation, EGUsphere, 2023, 1–16, 10.5194/egusphere-2023-1308, 2023.

Liu, L., Li, S., Li, X., Jiang, Y., Wei, W., Wang, Z., and Bai, Y.: An integrated approach for landslide susceptibility mapping by considering spatial correlation and fractal distribution of clustered landslide data, Landslides, 16, 715–728, 10.1007/s10346-018-01122-2, 2019.

Liu, Y. and Just, A.: SHAPforxgboost: SHAP Plots for'XGBoost' [code], 10.32614/CRAN.package.SHAPforxgboost, 2019.

Liu, Z., Gilbert, G., Cepeda, J. M., Lysdahl, A. O. K., Piciullo, L., Hefre, H., and Lacasse, S.: Modelling of shallow landslides with machine learning algorithms, Geoscience Frontiers, 12, 385–393, https://doi.org/10.1016/j.gsf.2020.04.014, 2021.

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

Lussana, C., Nipen, T. N., Ménétrier, B., and Seierstad, I. A.: Ensemble-based statistical interpolation of atmospheric variables near the surface, Quarterly Journal of the Royal Meteorological Society, 151, e5046, https://doi.org/10.1002/qj.5046, 2025.

Mahoney, M. J., Johnson, L. K., Silge, J., Frick, H., Kuhn, M., and Beier, C. M.: Assessing the performance of spatial cross-validation approaches for models of spatially structured data, arXiv preprint arXiv:2303.07334, 2023.

https://doi.org/10.5194/egusphere-2026-579
Preprint. Discussion started: 2 March 2026

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

Marín-Rodríguez, N. J., Vega, J., Zanabria, O. B., González-Ruiz, J. D., and Botero, S.: Towards an understanding of landslide risk assessment and its economic losses: a scientometric analysis, Landslides, 21, 1865–1881, 10.1007/s10346-024-02272-2, 2024.

Medina, V., Hürlimann, M., Guo, Z., Lloret, A., and Vaunat, J.: Fast physically-based model for rainfall-induced landslide susceptibility assessment at regional scale, CATENA, 201, 105213, https://doi.org/10.1016/j.catena.2021.105213, 2021.

Meld. St. 27: Tryggere framtid - førebudd på flaum og skred, 2023–2024.

Mets, K. D., Armenteras, D., and Dávalos, L. M.: Spatial autocorrelation reduces model precision and predictive power in deforestation analyses, Ecosphere, 8, e01824, https://doi.org/10.1002/ecs2.1824, 2017.

Meyer, H. and Pebesma, E.: Machine learning-based global maps of ecological variables and the challenge of assessing them, Nature Communications, 13, 2208, 10.1038/s41467-022-29838-9, 2022.

Meyer, H., Ludwig, M., Milà, C., Linnenbrink, J., and Schumacher, F.: The CAST package for training and assessment of spatial prediction models in R, arXiv preprint arXiv:2404.06978, 2024.

Mizutori, M.: SRSG statement for the international landslide consortium conference Kyoto, Japan, Proceedings of the 2018 IPL Symposium on Landslides,

Momeni, A. A., Khanlari, G. R., Heidari, M., Sepahi, A. A., and Bazvand, E.: New engineering geological weathering classifications for granitoid rocks, Engineering Geology, 185, 43–51, https://doi.org/10.1016/j.enggeo.2014.11.012, 2015.

Montgomery, D. R., Schmidt, K. M., Dietrich, W. E., and McKean, J.: Instrumental record of debris flow initiation during natural rainfall: Implications for modeling slope stability, Journal of Geophysical Research: Earth Surface, 114, https://doi.org/10.1029/2008JF001078, 2009.

Moore, I. D., Gessler, P. E., Nielsen, G., and Peterson, G.: Soil attribute prediction using terrain analysis, Soil science society of america journal, 57, 443–452, 1993.

Moos, C., Bebi, P., Graf, F., Mattli, J., Rickli, C., and Schwarz, M.: How does forest structure affect root reinforcement and susceptibility to shallow landslides?, Earth Surface Processes and Landforms, 41, 951–960, https://doi.org/10.1002/esp.3887, 2016.

NGU: 1:250 000 harmonized bedrock dataset "berggrunn N250" [dataset], 2025a.

NGU: 1:50 000 harmonized superficial deposits dataset "løsmasse harmonisert N50" [dataset], 2025b.

NGU: 1:250 000 deposit thickness dataset "løsmassemektighet N50" [dataset], 2025c.

O'Callaghan, J. F. and Mark, D. M.: The extraction of drainage networks from digital elevation data, Computer Vision, Graphics, and Image Processing, 28, 323–344, https://doi.org/10.1016/S0734-189X(84)80011-0, 1984.

Oguz, E. A., Depina, I., and Thakur, V.: Effects of soil heterogeneity on susceptibility of shallow landslides, Landslides, 19, 67–83, 10.1007/s10346-021-01738-x, 2022.

Parker, K. C.: Environmental relationships and vegetation associates of columnar cacti in the northern Sonoran Desert, Vegetatio, 78, 125–140, 10.1007/BF00033422, 1988.

Peeters, I. L. J.: Testing the significance of vegetational parameters on shallow landslide occurrence for landslides triggered in August 2023 during the extreme weather event Hans, Western Norway University of Applied Sciences, Sogndal, 2024.

Petschko, H., Brenning, A., Bell, R., Goetz, J., and Glade, T.: Assessing the quality of landslide susceptibility maps – case study Lower Austria, Nat. Hazards Earth Syst. Sci., 14, 95–118, 10.5194/nhess-14-95-2014, 2014.

Phillips, C., Hales, T., Smith, H., and Basher, L.: Shallow landslides and vegetation at the catchment scale: A perspective, Ecological Engineering, 173, 106436, https://doi.org/10.1016/j.ecoleng.2021.106436, 2021.

Probst, P., Boulesteix, A.-L., and Bischl, B.: Tunability: importance of hyperparameters of machine learning algorithms, J. Mach. Learn. Res., 20, 1934–1965, 2019.

Pullarello, J.: Improved and curated National DEM raster dataset, based on Norwegian national mapping agency 2023 data tiles [dataset], 2024.

Pullarello, J., Rubensdotter, L., and Stalsberg, K.: Technical report explaining the development of the National Susceptibility map for shallow landslides in Norway, version 2025, Trondheim, 2025.

https://doi.org/10.5194/egusphere-2026-579
Preprint. Discussion started: 2 March 2026

Qin, B., Zou, J., Wang, X., Yang, L., and Lv, P.: Landslide susceptibility evaluation using multiple machine learning models: a case study of Wanzhou district in the Three Gorges reservoir area, Geological Journal, 1–16, https://doi.org/10.1002/gj.70104, 2025.

Quevedo, R. P., Maciel, D. A., Uehara, T. D. T., Vojtek, M., Rennó, C. D., Pradhan, B., Vojteková, J., and Pham, Q. B.: Consideration of spatial heterogeneity in landslide susceptibility mapping using geographical random forest model, Geocarto International, 37, 8190–8213, 10.1080/10106049.2021.1996637, 2022.

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

Rickli, C. and Graf, F.: Effects of forests on shallow landslides–case studies in Switzerland, Forest Snow and Landscape Research, 82, 33–44, 2009.

Riley, S. J., DeGloria, S. D., and Elliot, R.: Index that quantifies topographic heterogeneity, Intermountain Journal of Sciences, 5, 23–27, 1999.

Roberts, D. R., Bahn, V., Ciuti, S., Boyce, M. S., Elith, J., Guillera-Arroita, G., Hauenstein, S., Lahoz-Monfort, J. J., Schröder, B., Thuiller, W., Warton, D. I., Wintle, B. A., Hartig, F., and Dormann, C. F.: Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure, Ecography, 40, 913–929, https://doi.org/10.1111/ecog.02881, 2017.

Rouai, M. and Jaaidi, E. B.: Scaling properties of landslides in the Rif mountains of Morocco, Engineering Geology, 68, 353–359, https://doi.org/10.1016/S0013-7952(02)00237-5, 2003.

Rüther, D. C., Hefre, H., and Rubensdotter, L.: Extreme precipitation induced landslide event on 30th July 2019 in Jølster, western Norway, Norwegian Journal of Geology, Manuscript submitted for publication, 2022.

Rüther, D. C., Lindsay, E., and Slåtten, M. S.: Landslide inventory: 'Hans' storm southern Norway, August 7–9, 2023, Landslides, 21, 1155–1159, 10.1007/s10346-024-02222-y, 2024.

Sachinthaka, R., Kalatehjari, R., and Brook, M. S.: Evolution and critical evaluation of deterministic physically based rainfall-induced landslide susceptibility mapping: a mixed review, Natural Hazards, 121, 20795–20818, 10.1007/s11069-025-07634-0, 2025.

Saleem, N., Huq, M. E., Twumasi, N. Y. D., Javed, A., and Sajjad, A.: Parameters derived from and/or used with Digital Elevation Models (DEMs) for landslide susceptibility mapping and landslide risk assessment: a review, ISPRS International Journal of Geo-Information, 8, 545, 2019.

Schmid, I.: The influence of soil type and interspecific competition on the fine root system of Norway spruce and European beech, Basic and Applied Ecology, 3, 339–346, https://doi.org/10.1078/1439-1791-00116, 2002.

Schumacher, J., Hauglin, M., Astrup, R., and Breidenbach, J.: Mapping forest age using National Forest Inventory, airborne laser scanning, and Sentinel-2 data, Forest Ecosystems, 7, 60, 10.1186/s40663-020-00274-9, 2020.

Schwarz, M., Cohen, D., and Or, D.: Root-soil mechanical interactions during pullout and failure of root bundles, Journal of Geophysical Research: Earth Surface, 115, https://doi.org/10.1029/2009JF001603, 2010a.

Schwarz, M., Lehmann, P., and Or, D.: Quantifying lateral root reinforcement in steep slopes – from a bundle of roots to tree stands, Earth Surface Processes and Landforms, 35, 354–367, https://doi.org/10.1002/esp.1927, 2010b.

Sidle, R. C.: A theoretical model of the effects of timber harvesting on slope stability, Water Resources Research, 28, 1897–1910, https://doi.org/10.1029/92WR00804, 1992.

Sidle, R. C. and Bogaard, T. A.: Dynamic earth system and ecological controls of rainfall-initiated landslides, Earth-science reviews, 159, 275–291, 2016.

Sidle, R. C. and Ochiai, H.: Landslides - Processes, prediction, and land use, Water resources monograph, Washington, 870 pp.2006.

Sigrist, F.: Gaussian process boosting, Journal of Machine Learning Research, 23, 1–46, 2022.

Sim, K. B., Lee, M. L., and Wong, S. Y.: A review of landslide acceptable risk and tolerable risk, Geoenvironmental Disasters, 9, 3, 10.1186/s40677-022-00205-6, 2022.

Sorbino, G., Sica, C., and Cascini, L.: Susceptibility analysis of shallow landslides source areas using physically based models, Natural Hazards, 53, 313–332, 10.1007/s11069-009-9431-y, 2010.

Steger, S., Brenning, A., Bell, R., Petschko, H., and Glade, T.: 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, 2016.

https://doi.org/10.5194/egusphere-2026-579
Preprint. Discussion started: 2 March 2026

Steger, S., Mair, V., Kofler, C., Pittore, M., Zebisch, M., and Schneiderbauer, S.: Correlation does not imply geomorphic causation in data-driven landslide susceptibility modelling – Benefits of exploring landslide data collection effects, Science of The Total Environment, 776, 145935, https://doi.org/10.1016/j.scitotenv.2021.145935, 2021.

Stoffel, M., Tiranti, D., and Huggel, C.: Climate change impacts on mass movements — Case studies from the European Alps, Science of The Total Environment, 493, 1255–1266, https://doi.org/10.1016/j.scitotenv.2014.02.102, 2014.

Sultana, N.: Analysis of landslide-induced fatalities and injuries in Bangladesh: 2000-2018, Cogent Social Sciences, 6, 1737402, 10.1080/23311886.2020.1737402, 2020.

Team, R. C.: R: A language and environment for statistical computing [code], 2024.

Teich, M., Accastello, C., Perzl, F., and Kleemayr, K.: Protective Forests as Ecosystem-based Solution for Disaster Risk Reduction (Eco-DRR), IntechOpen, London, 10.5772/intechopen.95014, 2022.

Thiery, Y., Maquaire, O., and Fressard, M.: Application of expert rules in indirect approaches for landslide susceptibility assessment, Landslides, 11, 411–424, 10.1007/s10346-013-0390-8, 2014.

Tofani, V., Bicocchi, G., Rossi, G., Segoni, S., D'Ambrosio, M., Casagli, N., and Catani, F.: Soil characterization for shallow landslides modeling: a case study in the Northern Apennines (Central Italy), Landslides, 14, 755–770, 10.1007/s10346-017-0809-8, 2017.

Tufano, R., Formetta, G., Calcaterra, D., and De Vita, P.: Hydrological control of soil thickness spatial variability on the initiation of rainfall-induced shallow landslides using a three-dimensional model, Landslides, 18, 3367–3380, 10.1007/s10346-021-01681-x, 2021.

Urdangarin, A., Goicoa, T., and Ugarte, M. D.: Evaluating recent methods to overcome spatial confounding, Revista Matemática Complutense, 36, 2023.

Utthasini, M., Ilampooranan, I., Singh, S. K., Kanga, S., Kumar, P., Halder, K., Pradhan, B., Srivastava, A. K., Chatterjee, R. S., Chakrabortty, R., Ali, T., and Meraj, G.: Enhancing landslide susceptibility mapping in the Himalayas: geospatial and machine learning with explainable AI (XAI), Gondwana Research, 149, 262–290, https://doi.org/10.1016/j.gr.2025.08.003, 2026.

Valavi, R., Elith, J., Lahoz-Monfort, J. J., and Guillera-Arroita, G.: blockCV: an R package for generating spatially or environmentally separated folds for k-fold cross-validation of species distribution models, bioRxiv, 357798, 10.1101/357798, 2018.

Wadoux, A. M. J. C., Minasny, B., and McBratney, A. B.: Machine learning for digital soil mapping: Applications, challenges and suggested solutions, Earth-Science Reviews, 210, 103359, https://doi.org/10.1016/j.earscirev.2020.103359, 2020.

Waldron, L. J.: The shear resistance of root-permeated homogeneous and stratified soil, Soil Science Society of America Journal, 41, 843–849, https://doi.org/10.2136/sssaj1977.03615995004100050005x, 1977.

Wang, P. and Deng, H.: The impact of different sampling strategies on landslide susceptibility assessment: an explainable hybrid BO-XGBoost model, Earth Science Informatics, 18, 440, 10.1007/s12145-025-01931-9, 2025.

Wei, A., Yu, K., Dai, F., Gu, F., Zhang, W., and Liu, Y.: Application of tree-based ensemble models to landslide susceptibility mapping: a comparative study, Sustainability, 14, 6330, 2022.

Weiss, A.: Topographic position and landforms analysis, ESRI user conference, San Diego, CA,

Wu, T. H., McKinnell, W. P., and Swanston, D. N.: Strength of tree roots and landslides on Prince of Wales Island, Alaska, Canadian Geotechnical Journal, 16, 19–33, 10.1139/t79-003, 1979.

Wu, Y., Ke, Y., Chen, Z., Liang, S., Zhao, H., and Hong, H.: Application of alternating decision tree with AdaBoost and bagging ensembles for landslide susceptibility mapping, Catena, 187, 104396, https://doi.org/10.1016/j.catena.2019.104396, 2020.

Ye, C., Wu, H., Oguchi, T., Tang, Y., Pei, X., and Wu, Y.: Physically based and data-driven models for landslide susceptibility assessment: principles, applications, and challenges, Remote Sensing, 17, 2280, 2025.