Landslide susceptibility mapping in southwest Sweden using an automated deep neural support vector regression model

Abbas Abbaszadeh Shahri[1,2], *Stefan* Larsson[2*]

[1] JohanLundberg AB, Uppsala, Sweden
[2] KTH Royal Institute of Technology, Stockholm, Sweden

Abstract. Landslides are a persistent geohazard in Sweden, especially in areas dominated by soft, clay-rich soils. This study develops a high-resolution landslide susceptibility map using a novel automated deep neural support vector regression (ADNSVR) model. Seventeen conditioning factors were included, grouped into remote-sensing-, geo-, land-, and distance-based parameters. Geo-based factors included soil depth, soil type, and soil type, highlighting the importance of sensitive clays. The model was trained and validated with documented Swedish landslides and showed strong predictive performance. Sensitivity analysis identified soil type, soil depth, slope, and soil type as the most influential controls. Glacial and postglacial clays, covering about 32% of the area, were the most significant sub-types. These results support regional landslide susceptibility assessment and risk-informed land-use planning.

1 Introduction

Landslides are among the most critical geohazards in Sweden, particularly in regions characterized by soft, fine-grained sediments such as clays and silts. Because areas underlain by soft and sensitive clays are widespread across Sweden, there is a strong need for large-scale landslide susceptibility mapping. Such mapping provides a basis for identifying areas where more detailed geotechnical investigations should be prioritized. Such regional-scale assessments provide an important first step for screening large territories and supporting early-stage planning and risk management.

Landslide susceptibility mapping approaches are commonly divided into conventional methods (qualitative, semi-quantitative, and quantitative) and more recent AI-based techniques [1]. Qualitative methods rely on expert interpretation and field observations but are often subjective and difficult to standardize [2]. Semi-quantitative models reduce subjectivity through GIS-based weighting schemes, although their reliability depends on assigned weights and linear assumptions [3]. Quantitative approaches such as logistic regression and frequency ratio use historical inventories to link landslides with conditioning factors, but they may struggle in complex, heterogeneous environments [4].

Recently, *AI*-based methods, including neural networks and support vector machines, have shown strong potential for capturing non-linear interactions among multiple factors, especially in clay-rich regions where slope instability depends on complex geotechnical and hydrological processes [5]. AI-based methods should therefore not be regarded as a substitute for geotechnical analysis, but rather as an efficient tool for early-stage hazard assessment. By identifying potentially critical areas, AI-based susceptibility mapping can help decision-makers determine where more advanced site-specific stability investigations are needed, particularly in areas planned for development or where slope instability problems may demand mitigation measures.

In this study, we have developed a high-resolution landslide susceptibility map for southwest Sweden using an automated deep neural support vector regression (*ADNSVR*) model, with particular emphasis on the role of soft clay-related geo-parameters and their contribution to landslide occurrence. The study demonstrates the applicability of advanced machine-learning techniques for improving regional-scale landslide susceptibility assessment and supporting risk-informed land-use planning in Sweden.

2 Study area and database

The study area covers a large region in southwest Sweden, extending from Gothenburg northwards to Munkedal over approximately 110 km. The area includes the sites of major landslides that affected the E6 motorway in Munkedal in 2006 and in Stenungsund in 2023. In this study, a total of 17 categorized conditioning factors were derived from multiple data sources (Fig.1). Their spatial distributions were prepared at a 10×10 m resolution following the approach presented by [5].

* Corresponding author: stefan.larsson@byv.kth.se

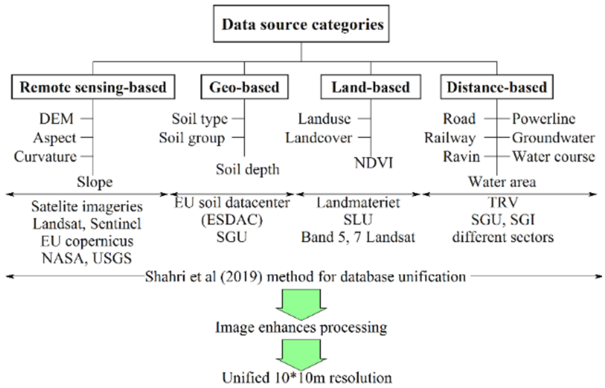


Fig. 1. Categorized and produced conditioning factors and data sources used.

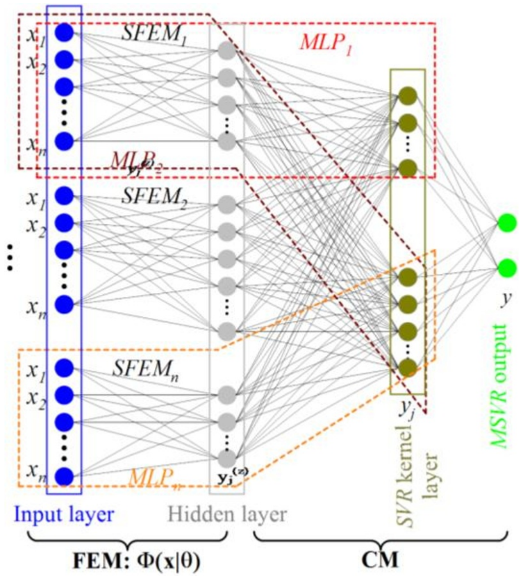
3 Modelling procedure and data

The model used in this study is ADNSVR (Fig. 2a), a hybrid machine-learning approach that combines a deep neural network with support vector regr human, ession. The neural network component extracts relevant features from the input data, whereas the support vector regression component produces the final prediction. The model consists of a feature extraction module, implemented as a multilayer perceptron (MLP), and a prediction module including the kernel-based SVR output stage. As presented in Fig. 2b, this architecture enables the model to capture complex relationships between conditioning factors and landslide occurrence. In addition, the automated tuning of hyperparameters improves reproducibility and reduces manual bias (Fig. 2b). ADNSVR is therefore well suited for landslide susceptibility mapping in complex and heterogeneous geospatial environments.

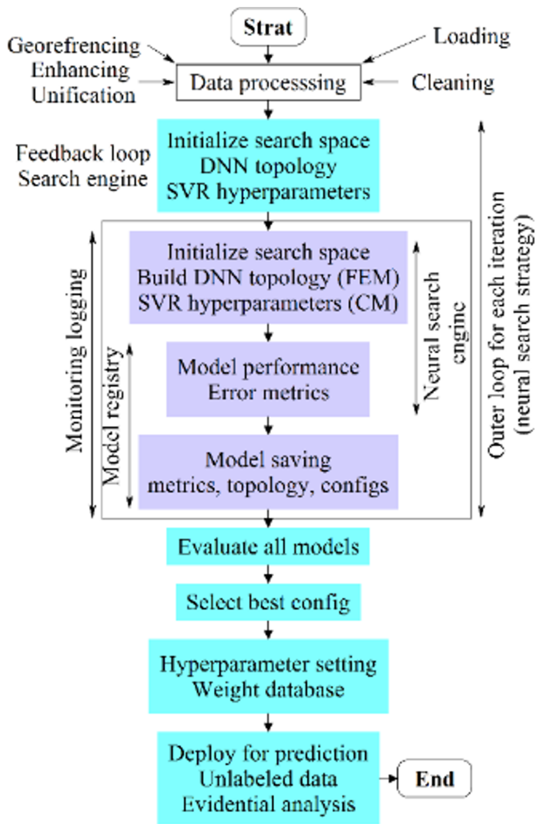
As the modelling framework is based on supervised learning using labelled landslide and non-landslide samples from the same inventory, the validation mainly evaluates internal model stability rather than independent prediction [6]. The dataset was therefore divided into training and validation subsets, where the training set was used for model development and the validation set for testing on unseen samples from the same population. Although model performance may vary depending on the random data partition, the similar values of area under the curve (AUC) [7], accuracy, sensitivity, and specificity for the training and validation datasets indicates internal consistency within the study area [8]. AUC reflects the threshold-independent discrimination ability of the model, while the confusion-matrix metrics reflect the classification performance at the selected threshold. The validation results should therefore be interpreted as evidence of internal generalization within the analysed inventory, rather than as proof of predictive transferability to other regions [9].

4 Spatial analysis of the distributed factors

The landslide inventory and susceptibility maps are shown in Fig. 3a and b, respectively, while the *ADNSVR* model performance, evaluated using AUC and confusion matrix metrics, is presented in Fig. 3c. In Sweden, landslides are strongly influenced by Quaternary glacial and postglacial processes. Clay sensitivity is associated with postglacial marine clays, which commonly have high water content, low permeability, weak mechanical properties, and a metastable structure caused by salt leaching [10]. These soils may fail catastrophically when triggered by erosion, rainfall, excavation, or loading.



(a)



(b)

Fig. 2. The overview of the designed *ADNSVR* topology (a) and developed automated procedure (b).

4.1 Soil type and anthropogenic influence

To evaluate the influence of the input factors, a sensitivity analysis based on the partial derivative method was applied. This method estimates how sensitive the model output is to changes in each input variable using the trained network weights and neuron outputs. For each input factor, a sensitivity value was calculated to represent its contribution to variations in the predicted output. To express these results as the variable importance percentages shown in Fig. 4, the absolute sensitivity values were normalized relative to the total sensitivity of all input variables. In this way, the importance values were scaled to sum to 100%, allowing direct comparison between the input factors. A higher percentage therefore indicates that the corresponding factor has a greater influence on the model prediction.

The sensitivity analysis of the ADNSVR inputs (Fig. 4) showed that soil type and soil type were among the most influential parameters. This likely reflects the model's ability to capture key contrasts between clay-rich and non-clay materials that are relevant to landslide occurrence. However, a meaningful interpretation of the model results also requires spatial comparison between the conditioning factors and the landslide inventory. In landslide susceptibility mapping, distance to roads and railways is generally considered an indicator of anthropogenic, activity and terrain modification. Such infrastructure may be associated with slope cutting, excavation, loading, and drainage alteration, rather than acting as a primary triggering factor [11].

They may therefore contribute locally to landslide occurrence. In the Swedish context, roads and railways are preferentially located in geomorphologically favourable settings (e.g., ravines, valleys, areas with thicker soil cover and gentler slopes), which are already represented by other variables in the model (elevation, soil type, soil depth, watercourses, etc.). As a result, the distance to road/railway may act as spatially correlated surrogates of these underlying environmental conditions rather than independent contributors [12].

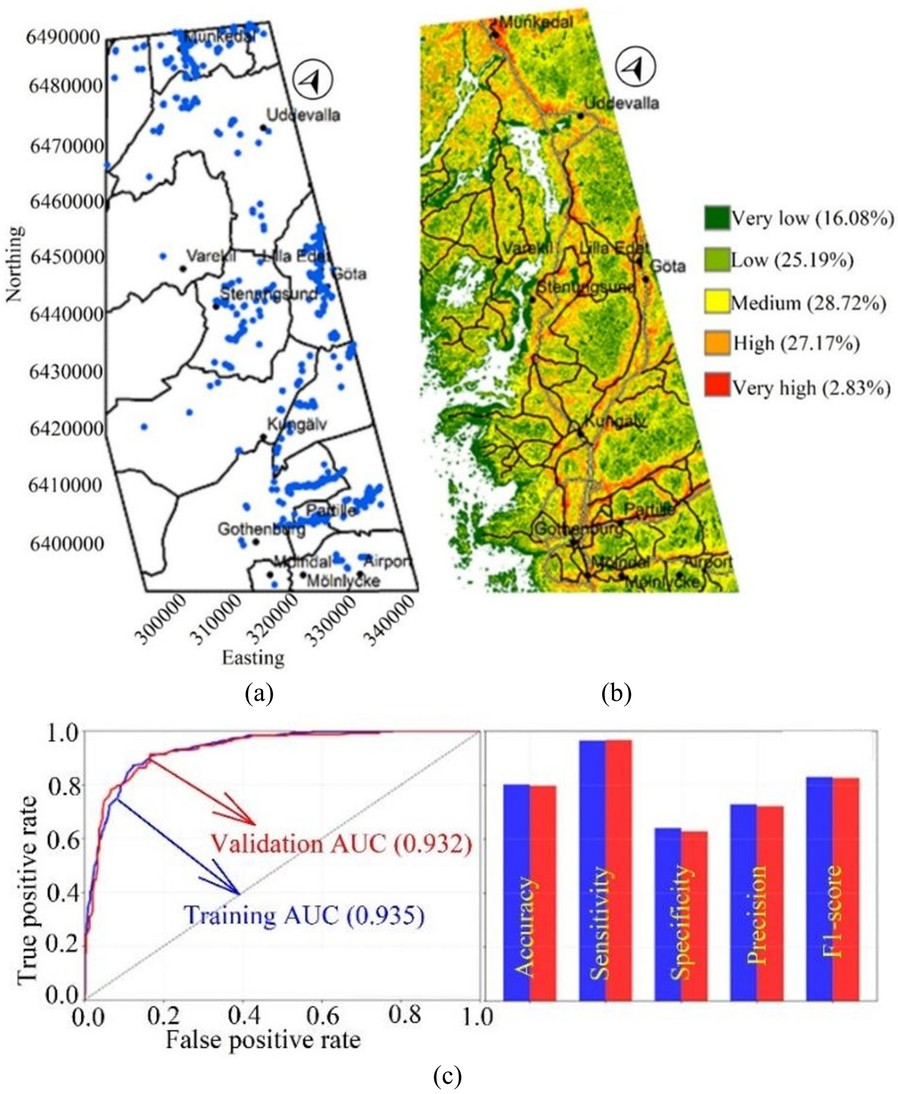


Fig. 3. The landslide inventory (a), generated landslide susceptibility map (b), and results of training stability (c).

 https://doi.org/10.1051/e3sconf/202673003020

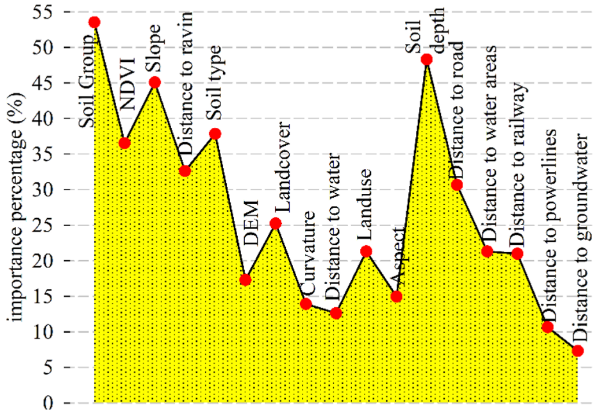


Fig. 4. Sensitivity analysis results obtained using the *ADNSVR* model.

4.2 Topography and soil conditions

Comparison of the landslide inventory with slope (Fig. 5a) and soil-type distribution (Fig. 5b) shows that in southwest Sweden, even moderate slopes can be unstable when underlain by soft clays with low shear strength, particularly under elevated pore-water pressures. As shown in Fig. 5a, most landslides occur on 3°–15° slopes in postglacial clay soils, consistent with [13].

4.3 Soil depth and slope stability

Soil depth is a critical parameter influencing slope stability, as it controls the thickness of potentially unstable material and the location of failure surfaces. Comparison of the soil depth distribution with the landslide inventory map showed that most documented landslides are associated with soil depths greater than 10 m and particularly within the range of 10 − 20 m (Fig. 6).

These results are consistent with [14], which showed that in clay-rich regions, deep soft soil layers overlying stronger bedrock or dense moraine can create favourable conditions for translational and rotational slides. Variations in soil depth also affect groundwater storage and pore-pressure development, further influencing slope stability [15].

4.4 Hydrological factors

Hydrological factors related to rivers, water bodies, and aquifers play an important role in landslide occurrence. The concentration of landslides near rivers and areas with fluctuating water levels indicates destabilization of clay-rich banks, while prolonged rainfall can reduce effective stress within sensitive clays (Fig. 7).

4.5 Distribution of clay-dominated soil types

The dominance of glacial and postglacial clays among influential factors (Fig. 8) underscores the necessity of incorporating clay sensitivity considerations into landslide hazard assessments [15]. Among the identified soil types, glacial and postglacial clays account for approximately 32% of the study area, making them the most widely distributed geomaterials associated with landslide occurrences.

4.6 Land use and infrastructure effects

In Swedish lowlands, urban development and infrastructure construction may disturb slope equilibrium through excavation and loading. Therefore, land-use and land-cover changes can partially modify slope stability due to infiltration rates, vegetation root reinforcement, and surface loading [16]

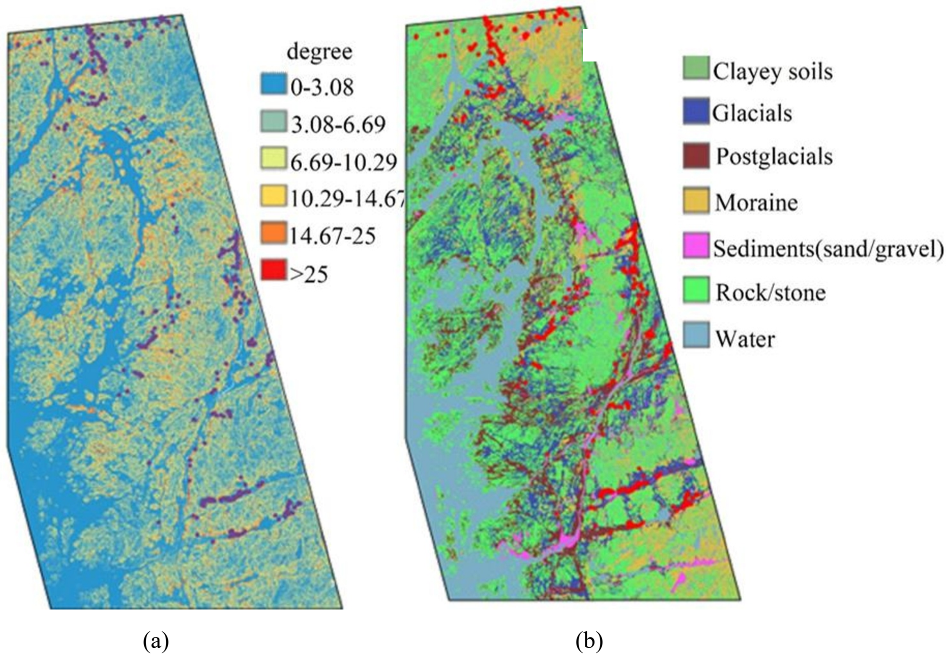


Fig. 5. Distribution of documented landslides in relation to slope (a) and identified soil types (b).

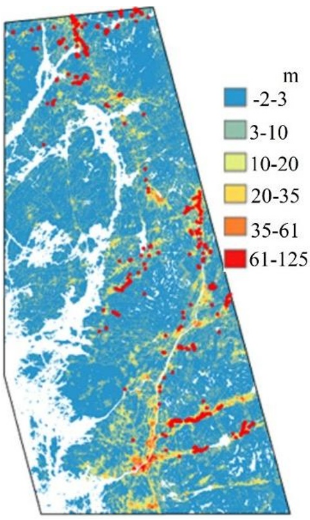
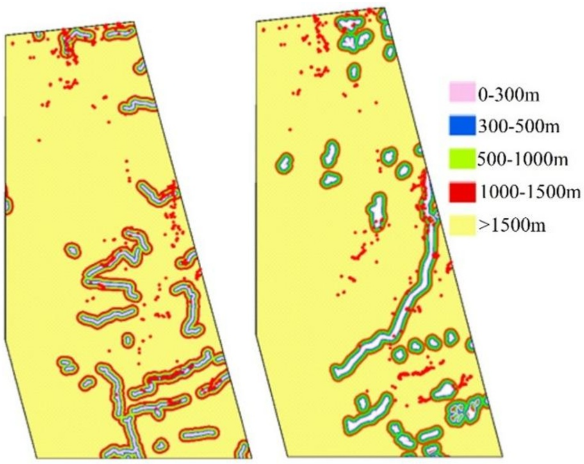


Fig. 6. Spatial distribution of the landslide inventory in relation to soil depth.



(a) (b)

Fig. 7. Comparison of the landslide inventory with distance to watercourses (a) and water bodies (b).

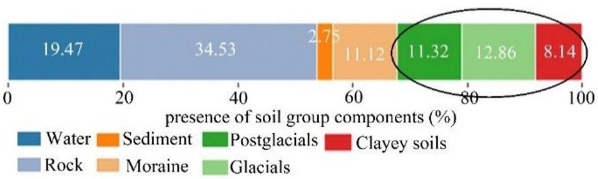


Fig. 8. Pixel-based distribution of the soil types in the study area.

5 Discussion

The results indicate that landslide susceptibility in southwest Sweden is strongly controlled by the distribution of soft clay deposits and related geomorphological conditions. The sensitivity analysis performed using the ADNSVR model indicates that soil type, soil depth, slope, and soil type are among the most influential conditioning factors. This result is consistent with previous studies showing that soft and sensitive clays play a key role in Swedish landslide processes [e.g. 11, 15, 16].

The importance of clay-related parameters reflects the geological setting of the region, where extensive deposits of glacial and postglacial marine clays occur. These soils are typically characterized by high water content, low permeability, and low shear strength, and they may develop high sensitivity due to leaching of salts. As discussed in previous studies on landslide susceptibility modelling [e.g. 12, 13], lithology and soil properties often represent the dominant explanatory variables in susceptibility analyses. The results are consistent with previous studies highlighting the importance of clay-rich soils for landslide occurrence in the Nordic countries.

Soil depth also emerges as an important factor in the model. This is in line with geotechnical interpretations of slope instability in clay-rich terrains, where thick deposits of soft soil overlying stronger material may favour rotational or translational failure mechanisms. Previous investigations of highly sensitive clay (quick clay) landslides in Sweden have similarly highlighted the importance of stratigraphy and the thickness of soft soil layers for slope stability [16]. Including such subsurface information in susceptibility models therefore contributes to a more realistic representation of landslide conditions.

In contrast, some human-induced factors, such as distance to roads and railways, appear to be of lower importance in the model. One possible explanation is that infrastructure corridors often follow geo-morphologically favourable terrain, such as valleys or areas with thicker soil cover. In this case, the variables partly reflect underlying environmental conditions rather than representing independent causal factors. Nevertheless, infrastructure development and excavation activities may locally disturb slope stability, particularly in areas with sensitive clays.

From a methodological perspective, the results suggest that hybrid machine-learning approaches can capture nonlinear relationships between conditioning factors and landslide occurrence. Similar advantages of machine learning approaches have been reported in previous susceptibility mapping studies, where models such as neural networks and support vector machines have shown good performance in complex geological environments.

However, some limitations should be noted. The validation results presented in this study are based on internal partitioning of the landslide inventory, and therefore primarily reflect internal model consistency rather than independent predictive capability. The reported performance metrics should therefore be interpreted as indicators of model robustness within the study area. Future work could include independent validation using temporally or spatially separate datasets, as well as testing the transferability of the modelling approach to other clay-dominated regions in Sweden.

Despite these limitations, the produced landslide susceptibility map provides useful information for regional hazard assessment and spatial planning.

Identifying areas where soft clay deposits coincide with unfavourable topographic and hydrological conditions can help guide more detailed geotechnical investigations and support risk-informed land-use planning. Such information is particularly relevant in Sweden, where ongoing infrastructure development and possible future climate-related changes in precipitation patterns may increase the likelihood of landslide occurrence.

This research was supported by the Swedish Transport Administration through the Maintenance R&I Portfolio.

References

1. L. Shano, T. K. Raghuvanshi, M. Meten, Landslide susceptibility evaluation and hazard zonation techniques – a review. Geoenviron. Disasters **7**, 18 (2020). https://doi.org/10.1186/s40677-020-00152-0
2. A. Carrara, M. Cardinali, R. Detti, F. Guzzetti, V. Pasqui, P. Reichenbach, GIS techniques and statistical models in evaluating landslide hazard. Earth Surf. Process. Landforms **16**, 427–445 (1991). https://doi.org/10.1002/esp.3290160505
3. N. Perilli, M. Lombardi, N. Squeglia, S. Stacul, S. Pagliara, Bridging gaps in landslide mapping: A semi-quantitative empirical framework for delineating key areas to improve collection of essential field-based and supplementary remote-based data. Infrastructures **11**(1), 11 (2026). https://doi.org/10.3390/infrastructures11010011
4. H. J. Park, J. H. Lee, A review of quantitative landslide susceptibility analysis methods using physically based modelling. J. Eng. Geol. **32**(1), 27–40 (2022). https://doi.org/10.9720/KSEG.2022.1.027
5. A. A. Shahri, J. Spross, F. Johansson, S. Larsson, Landslide susceptibility hazard map in southwest of Sweden using artificial neural network. CATENA **183**, 104225 (2019). https://doi.org/10.1016/j.catena.2019.104225
6. T. Hastie, R. Tibshirani, J. Friedman, The elements of statistical learning: Data mining, inference, and prediction. 2nd ed., Springer New York, NY, (2009). https://doi.org/10.1007/978-0-387-84858-7
7. A. P. Bradley, The use of the area under the ROC curve in the evaluation of machine learning algorithms. Pattern Recognition, **30**(7), 1145–1159 (1997). https://doi.org/10.1016/S0031-3203(96)00142-2
8. M. Kuhn, K. Johnson, Applied predictive modeling. Springer Science+Business Media, Springer Nature, New York, NY, (2013). https://doi.org/10.1007/978-1-4614-6849-3
9. D. M. W. Powers, Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation. Journal of Machine Learning Technologies, **2**(1):37–63 (2011).
10. H. Agrell, The quaternary of Sweden. Sveriges geologiska undersökning, Rep. CNR770 (1979). https://resource.sgu.se/dokument/publikation/c/c770rapport/c770-rapport.pdf
11. F. Guzzetti, A. Carrara, M. Cardinali, P. Reichenbach, Landslide hazard evaluation: a review of current techniques and their application in a multi-scale study. Geomorphology, **31**, 181–216, (1999). https://doi.org/10.1016/S0169-555X(99)00078-1
12. P. Reichenbach, M. Rossi, B. D. Malamud, M. Mihir, F. Guzzetti, A review of statistically-based landslide susceptibility models. Earth-Science Reviews, **180**, 60–91 (2018). https://doi.org/10.1016/j.earscirev.2018.03.001
13. SOU, Sweden facing climate change—threats and opportunities. Statens Offentliga Utredningar, SOU **2007**:60 (2007).
14. A. Ghaderi, A. A. Abbaszadeh Shahri, S. Larsson, A visualized hybrid intelligent model to delineate Swedish fine-grained soil layers using clay sensitivity. CATENA **214**, 106289 (2022). https://doi.org/10.1016/j.catena.2022.106289
15. A. A. Shahri, A. Malemir, C. Juhlin, Soil classification analysis based on piezocone penetration test data—A case study from a quick-clay landslide site in southwestern Sweden. Eng. Geol. **189**, 32–47 (2015). https://doi.org/10.1016/j.enggeo.2015.01.022
16. H. Löfroth, G. Göransson, J. Hedfors, K. Bergdahl, L. V. Well, Developing a methodology for characterizing cascading natural hazards associated with landslide risk in Sweden. Nat. Hazards **121**, 22915–22933 (2025). https://doi.org/10.1007/s11069-025-07706-1