

Evaluating disaster triggers and spatial susceptibility of retrogressive thawing slumps in permafrost regions Hoh Xil, Qinghai-Tibet Plateau

tanshuo li

3494236146@qq.com

Chengdu University of Technology <https://orcid.org/0009-0003-0650-5237>

Jiang Ye

Chengdu University of Technology <https://orcid.org/0000-0002-1289-4027>

Huiyuan Luo

Chengdu University of Technology

Qiang Xu

Chengdu University of Technology

Research Article

Keywords: Retrogressive thaw slumps, Permafrost region, Ensemble learning, Susceptibility mapping, Qinghai–Tibet Plateau

Posted Date: September 18th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-9891418/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

1 Evaluating disaster triggers and spatial
2 susceptibility of retrogressive thaw slumps in
3 permafrost regions Hoh Xil, Qinghai–Tibet
4 Plateau

5 Tanshuo Li^a, Jiang Ye^{a,*}, Huiyuan Luo^b, Qiang Xu^b

6
7 ^aSchool of Earth and Planetary Sciences, Chengdu University of
8 Technology, Chengdu, Sichuan 610059, China

9
10 ^bState Key Laboratory of Geohazard Prevention and
11 Geoenvironment Protection, Chengdu University of Technology,
12 Chengdu, Sichuan 610059, China

13
14 *Corresponding author.

15 E-mail address: yejiang@cdut.edu.cn

16 **Abstract**

17 The intensification of retrogressive thaw slumps (RTSs) across permafrost regions
18 in response to climate warming has become a major threat to ecological stabil-
19 ity and engineering safety on the Qinghai–Tibet Plateau. In this study, a novel
20 framework integrating environmental factor analysis, convolutional neural net-
21 work (CNN)-based feature extraction, and heterogeneous ensemble learning was
22 proposed to evaluate RTS susceptibility in the Hoh Xil region. Exploratory factor
23 analysis (EFA) was employed to identify dominant environmental controls and
24 reduce feature redundancy. An optimized negative sampling strategy combining
25 HDBSCAN clustering and environmental similarity constraints was developed to
26 improve sample reliability and reduce spatial bias. CNN was introduced to extract
27 latent spatial features from multiple environmental variables, and several machine
28 learning models were integrated to construct the susceptibility framework. The
29 results indicate that hydrothermal conditions, especially extreme rainfall and
30 thawing-freezing processes, are the primary drivers of RTS occurrence, while
31 ground-ice content and fine-grained sediments provide critical material condi-
32 tions for slope failure. The ensemble learning framework significantly improved

predictive performance, with the Boosting-integrated model achieving the highest accuracy and AUC values. The generated susceptibility maps reveal that high-susceptibility zones are mainly distributed along slope transition areas and regions with strong thermokarst activity. The proposed framework provides an effective approach for regional-scale RTS susceptibility assessment and offers scientific support for infrastructure planning and hazard prevention in permafrost regions under climate change.

Keywords: Retrogressive thaw slumps, Permafrost region, Ensemble learning, Susceptibility mapping, Qinghai–Tibet Plateau

1 Introduction

Permafrost is formally defined as lithospheric material (soil, sediment, or bedrock) sustaining cryotic conditions ($\leq 0^\circ\text{C}$) for at least two consecutive years, irrespective of volumetric ice content or seasonal thaw cycles [1]. Its global distribution spans approximately $3 \times 10^7 \text{ km}^2$, predominantly across circumpolar regions and high-altitude systems such as the Qinghai–Tibet Plateau (QTP), which represents the largest mid-latitude permafrost repository and stores approximately $1.06 \times 10^6 \text{ km}^2$ of ice-rich permafrost [2]. In recent decades, the QTP has experienced accelerated warming at a rate of approximately $0.43^\circ\text{C}/\text{decade}$ [3], triggering widespread permafrost degradation manifested by increasing ground temperatures, active-layer thickening, and the expansion of thaw-related geomorphic processes [4].

Among thermokarst landforms, retrogressive thaw slumps (RTSs) represent one of the most dynamically evolving geomorphic processes in permafrost regions. RTSs are typically initiated when ice-rich permafrost becomes exposed and undergoes rapid thawing [5]. Their initiation is commonly associated with warming-induced increases in surface heat flux, which destabilize ice-rich slopes and accelerate ground-ice melt [6]. Extreme rainfall events can further intensify thaw processes by increasing sediment saturation and reducing slope stability [7]. In addition, anthropogenic disturbances, including road construction and engineering excavation, may accelerate thermal erosion and headwall retreat through enhanced heat advection [8]. These combined climatic and human-induced drivers significantly increase the vulnerability of ice-rich permafrost landscapes under ongoing climate warming [9].

RTS occurrence is governed by complex interactions among climatic, topographic, and subsurface environmental factors. Previous studies have demonstrated that climatic variables, including mean annual air temperature, precipitation, and extreme rainfall events, interact synergistically with topographic parameters such as slope, aspect, and elevation, as well as geological characteristics including permafrost type, ground-ice content, and soil texture, to control RTS initiation and expansion [10]. Temperature and precipitation are generally regarded as dominant controls on RTS dynamics [11]. Regions characterized by annual temperatures between -6°C and -4°C and moderate precipitation conditions are particularly susceptible to freeze–thaw erosion and slope instability [12]. Ground-ice content and scarp morphology are

also critical factors controlling headwall retreat and long-term slump development [13]. Furthermore, RTS persistence and reactivation are strongly associated with retreat rate and sediment transport processes [14].

Empirical investigations have identified RTS hotspots primarily distributed at elevations of 4600–5100 m, commonly occurring on north- to northeast-facing slopes with gradients of approximately 3° – 8° [15–17]. Once activated, RTS headwalls may retreat at rates of approximately 3–5 m/year and remain active for decades until excess ground ice becomes depleted or protective debris accumulates over exposed surfaces [18]. As geomorphic products of climate forcing and terrain susceptibility, RTSs can substantially alter landscape structure and threaten engineering infrastructure, including highways, railways, and pipelines, as well as fragile alpine ecosystems. Therefore, understanding the mechanisms controlling RTS occurrence and evolution is essential for regional hazard assessment and early-warning system development [19].

Traditional RTS investigations have mainly relied on remote sensing interpretation and field surveys [20]. Although these approaches can provide relatively reliable RTS inventories, they are limited in characterizing complex nonlinear relationships among environmental variables and often suffer from low temporal resolution [21]. In recent years, machine learning (ML) techniques have demonstrated considerable potential in permafrost hazard assessment because of their ability to process heterogeneous geospatial datasets and model nonlinear environmental interactions [22]. Nevertheless, issues such as imbalanced training samples, feature redundancy, and overfitting remain major challenges that can reduce model robustness and predictive accuracy [23].

To address these limitations, this study proposes an integrated framework for RTS susceptibility assessment. First, Hierarchical Density-Based Spatial Clustering of Applications with Noise (HDBSCAN) combined with environmental similarity constraints was introduced to optimize negative sample selection and reduce spatial sampling bias. Second, exploratory factor analysis (EFA) was employed to identify dominant environmental controls and reduce feature redundancy. Third, convolutional neural networks (CNNs) were integrated with ensemble learning approaches to improve latent spatial feature extraction and predictive performance. The proposed framework was validated through a case study conducted in the Hoh Xil region of the central Qinghai–Tibet Plateau.

2 Study area and materials

2.1 Study area

Hoh Xil is situated in the northwestern part of the Qinghai–Tibet Plateau, between the Kunlun and Tanggula Mountains. Its coordinates range from $32^{\circ}52'$ to $36^{\circ}16'N$ and $89^{\circ}24'$ to $94^{\circ}18'E$ (Fig. 1). The region spans approximately 83,000 km², with an average elevation exceeding 5000 m. Hoh Xil features a cold plateau arid climate, characterized by annual mean temperatures ranging from $-1.7^{\circ}C$ to $5.0^{\circ}C$ and mean annual precipitation of approximately 300 mm [24].

Most areas of Hoh Xil are underlain by ice-rich permafrost and are characterized by widespread glaciers and periglacial landforms [25]. The mean annual ground temperature (MAGT) generally ranges from -5.7°C to 0°C and commonly remains above -3.0°C in many regions [26]. Under rapid climate warming, permafrost degradation in Hoh Xil has intensified, resulting in ground-ice melting, thermal erosion-induced surface deformation, and subsequent mass-wasting processes, including solifluction and active-layer detachment landslides. Long-term thermokarst activity has transformed the region into a landscape dominated by lakes and depressions.

The Hoh Xil region contains 337 lakes larger than 1 km^2 , accounting for nearly 30% of the total lake area on the Qinghai-Tibet Plateau [27]. Retrogressive thaw slumps (RTSs) in remote permafrost areas of Hoh Xil are increasingly threatening the stability of the Qinghai-Tibet Railway and Highway corridors, highlighting the urgent need for hazard assessment and susceptibility evaluation.

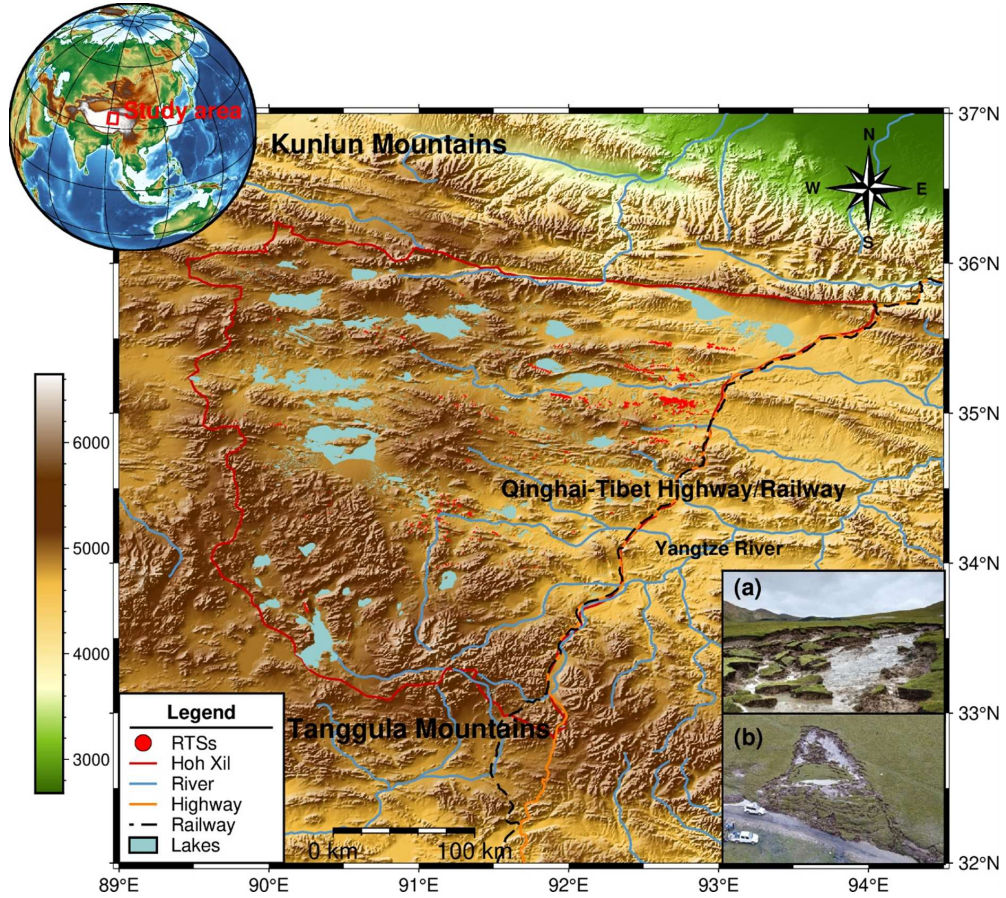


Fig. 1 Fig. 1. Geographical location of Hoh Xil and distribution of RTSs. (a) RTS typical cases. (b) RTSs destroy roads. (Image source: National Key Laboratory of Cryosphere Science, Chinese Academy of Sciences, Qinghai-Tibet Plateau Cryosphere Observation Research Station field research team.)

2.2 RTSs inventory

The RTS inventory used in this study was obtained from the open-access database of the National Tibetan Plateau Data Center and was originally interpreted from high-resolution remote sensing imagery acquired during 2018–2019 [28]. In addition, the interpretation results were validated through field investigations. Previous studies identified a total of 1734 RTSs in the Hoh Xil region, covering an area of approximately 30.82 km².

The mapped RTSs exhibit substantial variability in size and morphology. The largest slump covers approximately 0.2 km² and was formed through the coalescence of multiple adjacent slumps, whereas the smallest slump occupies only 307 m². Nearly 49.3% of all mapped RTSs are smaller than 0.01 km², indicating that small-scale thaw slumps dominate the regional distribution pattern.

Based on the existing inventory, the RTS dataset was further updated for the period from 2018 to 2021 using Google Earth (GE) imagery. Newly developed RTSs were manually interpreted and added to the inventory, while previously identified RTS boundaries were revised according to geomorphic characteristics, including surface color, trace extent, headwall morphology, and slump shape. Following refinement and quality control, a total of 1825 RTSs were ultimately confirmed within the Hoh Xil region.

2.3 Environmental factors

Climate and topography play key roles in controlling the physical retreat of RTS headwalls [29]. In addition, vegetation cover and soil properties strongly influence ground-ice content and permafrost stability [30]. In this study, potential environmental factors were grouped into three categories, including topographic, climatic, and soil-related variables. These variables were further screened to construct the environmental factor dataset used for susceptibility modeling. Detailed information on all selected factors is presented in Table 1, while their spatial distributions are shown in Fig. 2.

2.3.1 Terrain factors

The topographic factors included slope, topographic wetness index (TWI), and stream power index (SPI). These variables were derived from the TanDEM-X 30 m enhanced digital elevation model (EDEM) provided by EOC Geoservice [?]. Slope plays an important role in regulating slope stability by controlling runoff pathways and flow velocity, which subsequently influence soil moisture distribution [40]. Areas with steeper slopes generally exhibit stronger gravitational stress and surface erosion, thereby increasing the likelihood of slope failure.

The topographic wetness index (TWI) was used to characterize the spatial distribution of surface water accumulation. Higher TWI values indicate stronger water retention capacity and a greater tendency toward soil saturation. Under warming climatic conditions, these waterlogged areas are more susceptible to ground-ice thaw and thermokarst development, which can further promote RTS initiation.

The stream power index (SPI) reflects the erosive energy of flowing water and its sediment transport capability. Regions with high SPI values are generally subjected

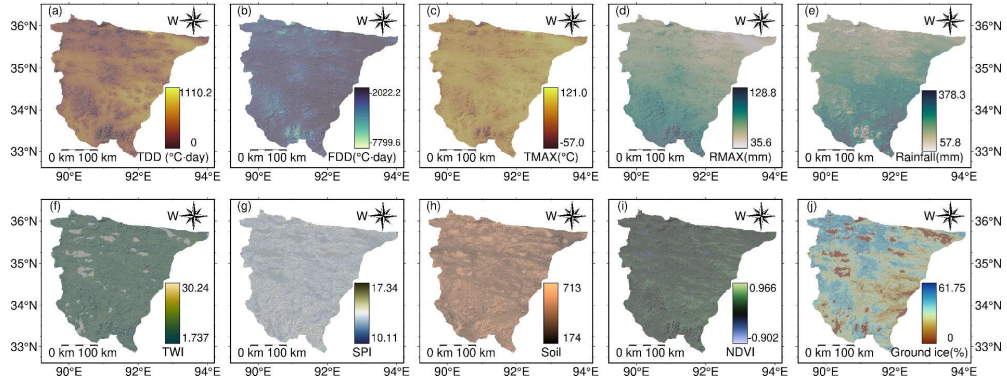


Fig. 2 Spatial distributions of important potential environmental factors. (a) TDD (Thawing Degree Days); (b) FDD (Freezing Degree Days); (c) Tmax (Maximum Summer Temperature); (d) Rmax (Maximum Summer Rainfall); (e) Rainfall (Annual Average Precipitation); (f) TWI (Topographic Wetness Index); (g) SPI (Stream Power Index); (h) Soil (Clay and Silt content); (i) NDVI (Normalized Difference Vegetation Index); (j) Ground ice content.

Table 1 Detailed information of potential environmental factors.

Category	Factors	Resolution	Data sources	References
Terrain	DEM	30 m	https://geoservice.dlr.de/web/	[31]
	Slope	30 m	Calculated using the Slope tool in ArcMap	[32]
	TWI	30 m	Calculated using the TWI tool in ArcMap	[33]
	SPI	30 m	Calculated using the SPI tool in ArcMap	[33]
Climate	TDD	1000 m	https://data.tpdc.ac.cn/home	[34]
	FDD	1000 m	https://data.tpdc.ac.cn/home	[34]
	Tmax	1000 m	https://data.tpdc.ac.cn/home	[34]
	Rmax	1000 m	https://data.tpdc.ac.cn/home	[35]
	Rainfall	1000 m	https://data.tpdc.ac.cn/home	[35]
Soil	Soil	250 m	https://soilgrids.org/	[36]
	NDVI	250 m	https://data.tpdc.ac.cn/home	[37]
	Permafrost	3000 m	https://data.tpdc.ac.cn/home	[38]
	Ground ice	1000 m	https://data.tpdc.ac.cn/home	[39]

to stronger hydrological erosion, which may destabilize soil structure and intensify moisture fluctuations. Such conditions can accelerate thaw-induced slope instability and enhance the development of RTSs.

All topographic variables were calculated from the TanDEM-X 30 m EDEM using spatial analysis tools in ArcGIS.

2.3.2 Climate factors

Climatic factors considered in this study included thawing degree days (TDD, $^{\circ}\text{C}\cdot\text{day}$), freezing degree days (FDD, $^{\circ}\text{C}\cdot\text{day}$), annual average precipitation (Rainfall, 0.1 mm), summer maximum temperature (T-max, $^{\circ}\text{C}$), and maximum precipitation

(R-max, 0.1 mm). TDD and FDD represent the cumulative annual temperatures above and below 0°C, respectively, and were used to characterize the thermal conditions during thawing and freezing periods [41].

Precipitation directly influences soil moisture conditions and acts as an important driver of ground-ice thaw. In addition, extreme climatic conditions represented by summer maximum temperature and maximum precipitation are closely associated with RTS initiation and expansion. Previous studies have shown that RTS activity is particularly active following warm and humid summers [42].

The climatic variables used in this study were derived from monthly precipitation and near-surface air temperature datasets covering the period from 2000 to 2021, which were obtained from the National Tibetan Plateau Data Center. To maintain spatial consistency with topographic variables, all climatic datasets were downscaled to 30 m resolution using a monthly lapse-rate correction method based on the TanDEM-X 30 m EDEM:

$$Y_{30m} = Y_{1km} + (DEM_{1km} - DEM_{30m}) \times LR \quad (1)$$

where DEM_{1km} and DEM_{30m} represent the elevations of the 1 km and 30 m DEMs, respectively; Y_{1km} and Y_{30m} denote the monthly average temperature or precipitation before and after downscaling; and LR is the monthly lapse rate.

2.3.3 Soil factors

Soil-related factors included permafrost type, ground-ice content, fine sediment content (clay + silt), and normalized difference vegetation index (NDVI). Permafrost type determines the thermal stability and ground-ice conditions of frozen ground. Ice-rich permafrost is particularly vulnerable to thaw-induced subsidence and slope failure under climatic warming. Regions with high ground-ice content generally exhibit higher RTS susceptibility because ground-ice melting produces water-saturated sediments that weaken soil structure and destabilize slopes. The resulting thermokarst depressions and localized ponding can further accelerate permafrost degradation and promote RTS expansion.

Fine-grained sediments (clay + silt) also play an important role in RTS development. Due to their strong water-retention capacity and swelling characteristics after saturation, these materials can significantly reduce soil shear strength during thawing processes. Therefore, areas containing abundant fine sediments are generally more susceptible to thaw-induced landslides and mass wasting [43].

NDVI was used to characterize vegetation conditions associated with permafrost stability. Vegetation cover can influence surface energy balance, soil moisture, and active-layer thickness. Previous studies have indicated that regions with relatively dense vegetation may still exhibit high RTS activity [44]. This phenomenon may be related to favorable moisture conditions for ground-ice accumulation beneath vegetated surfaces, while vegetation can also partially conceal underlying thermokarst instability.

3 Methods

This study proposes an integrated framework for retrogressive thaw slump (RTS) susceptibility assessment in permafrost regions, consisting of three major modules (Fig. 3). The first module focuses on the selection of dominant environmental factors using exploratory factor analysis (EFA), which was employed to reduce feature redundancy and identify key controlling variables associated with RTS occurrence.

The second module involves the generation of RTS and non-RTS samples. RTS samples were derived from the established inventory dataset, while non-RTS samples were generated by integrating Hierarchical Density-Based Spatial Clustering of Applications with Noise (HDBSCAN) and environmental similarity constraints. This strategy was designed to improve sample representativeness, reduce spatial sampling bias, and enhance model generalization ability.

The third module develops a heterogeneous ensemble machine learning framework for RTS susceptibility assessment. In this framework, convolutional neural networks (CNNs) were employed to automatically extract latent spatial features from environmental variables. The extracted deep features were subsequently integrated with multiple machine learning algorithms to improve predictive performance and spatial representation capability.

Based on the trained models and environmental variables, susceptibility maps of RTS occurrence were generated for the Hoh Xil region. The methodological principles and implementation procedures of each module are described in the following sections.

3.1 Factors selection

Exploratory factor analysis (EFA) was employed to identify the internal relationships among environmental variables by constructing latent factor models [45]. Before conducting EFA, the Kaiser–Meyer–Olkin (KMO) test [46] and Bartlett’s test of sphericity [?] were performed to evaluate the suitability of the dataset for factor analysis. According to the criteria proposed by Kaiser and Rice [47], the dataset was considered appropriate for EFA when the KMO value exceeded 0.60 and the significance level of Bartlett’s test satisfied $p < 0.05$.

Let X denote the original data matrix consisting of n observations and p variables. The factor model can be expressed as:

$$X = \Lambda F + \varepsilon \quad (2)$$

where Λ is the factor loading matrix, F represents the latent factor matrix, and ε denotes the residual error matrix. Prior to factor extraction, all variables were standardized using Z-score normalization.

The covariance matrix R of the standardized data can be expressed as:

$$R = \Lambda \Lambda^T + \Psi \quad (3)$$

where Ψ represents the unique variance matrix, and $\Lambda \Lambda^T$ is the reproduced correlation matrix explained by the retained factors.

The covariance matrix R was subsequently decomposed into the eigenvector matrix V and diagonal eigenvalue matrix D :

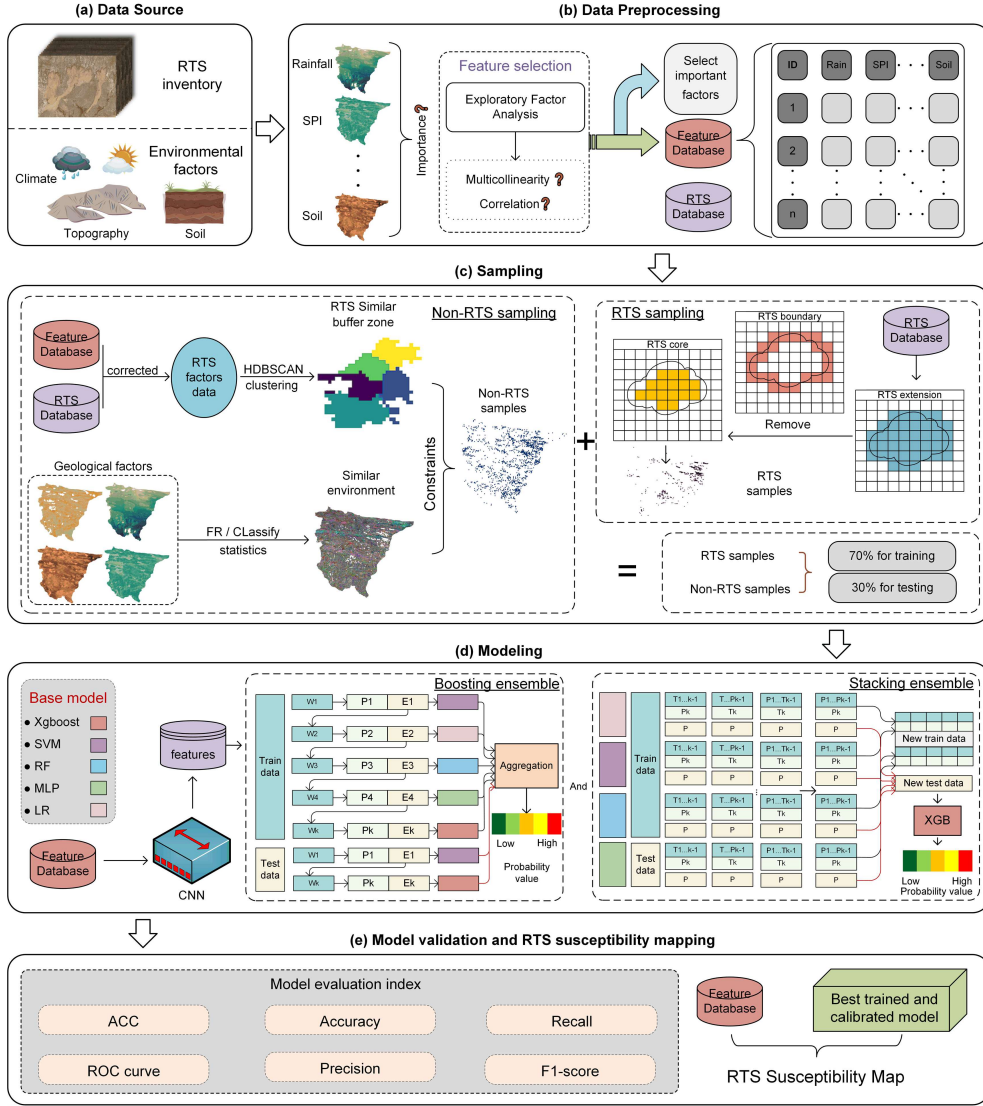


Fig. 3 Fig. 3. The framework of modeling and mapping for retrogressive thaw slumps.

$$R = VDV^T \quad (4)$$

257 The eigenvectors corresponding to the largest eigenvalues were retained to
 258 reconstruct the initial factor loading matrix:

$$\Lambda = VD^{1/2} \quad (5)$$

259 To improve interpretability, Promax oblique rotation was applied following Vari-
 260 max orthogonal rotation [48], allowing correlations among extracted latent factors.

261 The Varimax rotation maximizes the variance of squared factor loadings and can be
 262 expressed as:

$$V = \sum_{j=1}^m \left[\frac{1}{p} \sum_{i=1}^p l_{ij}^4 - \left(\frac{1}{p} \sum_{i=1}^p l_{ij}^2 \right)^2 \right] \quad (6)$$

263 where l_{ij} denotes the loading of variable i on factor j , m is the number of retained
 264 factors, and p is the number of observed variables.

265 After nonlinear transformation and regression adjustment, the final factor loading
 266 matrix was obtained. The importance of each environmental variable was determined
 267 according to its highest factor loading value. Variables with factor loadings greater
 268 than 0.7 were considered to have strong contributions to the corresponding latent
 269 factor. In addition, for reliable factor extraction in large datasets ($n > 100$), the
 270 communality of each variable was required to exceed 0.5, indicating that most of its
 271 variance could be explained by the retained factors.

272 3.2 Sampling strategy

273 3.2.1 RTS sampling

274 The conversion of RTS samples from vector polygons to raster data is an impor-
 275 tant step in susceptibility modeling. Different rasterization strategies may significantly
 276 influence susceptibility prediction results in permafrost regions [49]. In particular, the
 277 geometric accuracy of RTS boundaries and the spatial resolution of rasterized data can
 278 directly affect model reliability and spatial representation capability. Accurate preser-
 279 vation of RTS spatial morphology can improve the performance of polygon-based
 280 machine learning models and reduce classification uncertainty [15].

281 In this study, a core-area sampling strategy was adopted for RTS rasteriza-
 282 tion. Under this strategy, only raster cells completely enclosed within RTS polygon
 283 boundaries were retained as valid RTS samples (Fig. 3c). This approach minimizes
 284 boundary ambiguity and preserves the geometric integrity of RTS features during
 285 raster conversion.

286 The refined RTS inventory consisted of 1825 polygons, which were subsequently
 287 converted into raster data with a spatial resolution of 30 m. The extracted core
 288 areas contained a total of 14,944 pixels, which were used as positive samples for RTS
 289 susceptibility modeling.

290 3.2.2 Non-RTS sampling

291 (1) *HDBSCAN clustering*

292 Considering the spatial clustering characteristics of RTS samples, Hierarchi-
 293 cal Density-Based Spatial Clustering of Applications with Noise (HDBSCAN) was
 294 employed to identify regions with environmental conditions similar to RTS-positive
 295 areas (Fig. 3c) [50]. As an extension of the Density-Based Spatial Clustering of Appli-
 296 cations with Noise (DBSCAN) algorithm [51], HDBSCAN constructs hierarchical
 297 density relationships by iteratively performing DBSCAN over varying neighborhood
 298 radii and subsequently determines the optimal clusters based on cluster stability.

299 One major advantage of HDBSCAN is its capability to automatically identify
 300 noise points and isolate outliers into independent noise clusters [52]. This property is
 301 particularly suitable for geospatial datasets with uneven density distributions.

302 First, the mutual reachability distance (MRD) between two sample points was
 303 calculated as follows:

$$MRD_k(a, b) = \max \{core_k(a), core_k(b), d(a, b)\} \quad (7)$$

304 where $core_k(x)$ represents the core distance of point x , and $d(a, b)$ denotes the
 305 Euclidean distance between points a and b . The core distance is defined as the distance
 306 from point x to its k -th nearest neighbor:

$$core_k(x) = d(x, NN_k(x)) \quad (8)$$

307 where $NN_k(x)$ denotes the k -th nearest neighbor of point x , and k represents the
 308 minimum number of neighboring samples.

309 Subsequently, a minimum spanning tree (MST) was constructed based on MRD
 310 values to characterize the hierarchical connectivity among sampling points. The
 311 dataset was transformed into a weighted graph with MRD as edge weights, and the
 312 MST was generated using Prim's algorithm [53]. Hierarchical clustering was then
 313 established by progressively removing edges with the largest weights from the MST.

314 Assuming that the edge set of the MST is represented as:

$$e_1 \leq e_2 \leq \dots \leq e_{n-1}$$

315 clusters were partitioned by removing all connected components with edge weights
 316 exceeding a threshold λ . This process effectively pruned the cluster hierarchy under
 317 different density levels.

318 To determine the optimal cluster structure, cluster stability was quantitatively
 319 evaluated. For a cluster C , the birth threshold λ_{birth} and death threshold λ_{death}
 320 represent the minimum and maximum density levels at which the cluster exists,
 321 respectively. The stability of cluster C is defined as:

$$Stability(C) = \sum_{x_i \in C} (\lambda_i - \lambda_{birth}) \quad (9)$$

322 where λ_i denotes the density threshold when sample point x_i is assigned to cluster
 323 C . The density parameter λ is expressed as the reciprocal of the core distance:

$$\lambda = \frac{1}{core_k(x)} \quad (10)$$

324 The hierarchical clustering tree was traversed in a bottom-up manner to maximize
 325 the global stability metric and determine the optimal cluster partition. Samples that
 326 did not satisfy the stability criterion were classified as noise points.

327 (2) Classification of similar geological environments

328 After identifying environmentally similar regions using HDBSCAN clustering,
 329 candidate non-RTS areas were obtained. However, strong environmental heterogeneity

still existed among these regions. Random selection of non-RTS samples could therefore reduce environmental comparability with RTS samples and introduce sampling bias.

To improve the representativeness and validity of non-RTS samples, geological environmental consistency constraints were further incorporated [54]. Four environmental factors closely related to RTS formation, including TWI, slope, soil properties, and permafrost type, were selected to characterize geological similarity. The frequency ratio (FR) of each factor class was calculated as follows:

$$FR_i = \frac{F_i/F}{C_i/C} \quad (11)$$

where F_i represents the number of RTS pixels within the i -th factor interval, F is the total number of RTS pixels, C_i denotes the number of pixels within the corresponding factor interval, and C is the total number of pixels in the study area.

The FR values of different environmental classes were subsequently normalized using the min-max normalization method to determine their relative weights. Based on these weighted environmental factors, the entire study area was finally divided into 2800 similar geological environments.

(3) Optimized sampling strategy

A balanced sampling strategy was adopted to optimize the selection of RTS and non-RTS samples within each similar geological environment (SE_i). Let N_{RTS} and $N_{non-RTS}$ denote the numbers of RTS and non-RTS samples, respectively.

For RTS samples:

- If $N_{RTS}(SE_i) = 0$, the corresponding environment was excluded;
- If $N_{RTS}(SE_i) = 1$, the sample was assigned to the training set;
- If $N_{RTS}(SE_i) \geq 2$, the samples were randomly divided into 70% training data and 30% testing data.

For non-RTS samples:

- If the number of candidate non-RTS samples exceeded the RTS samples within the same environment, an equivalent number of non-RTS samples was randomly selected;
- If $N_{non-RTS}(SE_i) = 1$, the sample was assigned to the training set;
- If $N_{non-RTS}(SE_i) \geq 2$, the samples were divided into 70% training data and 30% testing data.

This strategy ensured balanced representation between RTS and non-RTS samples under comparable geological environmental conditions, thereby improving model robustness and reducing spatial sampling bias.

3.3 Susceptibility model

3.3.1 CNN for feature extraction

Convolutional neural networks (CNNs) are a class of deep neural networks derived from multilayer perceptrons and are particularly suitable for image processing

and spatial feature extraction [55]. CNNs employ local connectivity and weight-sharing mechanisms to reduce the number of trainable parameters, thereby improving computational efficiency and alleviating overfitting problems.

In this study, the input data for the CNN model consisted of a 10×10 matrix representing encoded environmental features for each sample. The transformation from original one-dimensional environmental feature vectors into two-dimensional spatial representations was achieved using an enhanced one-hot encoding strategy.

The convolutional layers extracted latent spatial features through multiple convolution kernels. The output of the j -th convolution kernel can be expressed as:

$$y_j = f \left(\sum_{i=1}^n x_i w_{ij} + b_j \right) \quad (12)$$

where x_i represents the input feature at position i , w_{ij} denotes the convolution kernel weight, b_j is the bias term, and $f(\cdot)$ represents the nonlinear activation function ReLU. Multiple convolution kernels were used to capture different spatial patterns associated with RTS occurrence.

To reduce the dimensionality of feature maps while preserving important spatial information, max-pooling operations were applied after convolution. The pooling process is defined as:

$$p_i = \max(x_i) \quad (13)$$

where x_i denotes the feature values within the pooling window centered at position i , and p_i is the pooled output feature.

The high-level latent features extracted by the CNN model were subsequently input into downstream machine learning classifiers for RTS susceptibility prediction. The hyperparameter settings of the CNN model are summarized in Table 2.

Table 2 Hyperparameter settings of CNNs.

Hyperparameter	Value	Hyperparameter	Value
Convolution kernel	3×3	Loss function	Cross-entropy
Pooling size	2×2	Learning rate	0.005
Activation function	ReLU	Dropout rate	0.3
Optimizer	Adagrad	Epoch	100

3.3.2 Machine learning based model

The machine learning-based models considered in this study include five representative algorithms: Extreme Gradient Boosting (XGBoost), Support Vector Machines (SVM), Random Forests (RF), Logistic Regression (LR), and Multi-layer Perceptrons (MLP).

XGBoost is an optimized gradient boosting framework that improves predictive performance through additive tree ensemble learning, incorporating regularization,

feature subsampling, and parallel computation, and is well suited for large-scale datasets [55].

SVM is a kernel-based classifier that constructs an optimal separating hyper-plane by maximizing the margin between classes, and is particularly effective for high-dimensional data with limited samples [55].

RF is an ensemble learning method that constructs multiple decision trees using bootstrap sampling and random feature selection, and enhances generalization ability through majority voting [55].

LR is a probabilistic linear model for binary classification, which estimates class probabilities using the logistic function and is widely used due to its simplicity, efficiency, and interpretability [55].

MLP is a classical feedforward neural network trained via backpropagation, capable of learning complex nonlinear mappings through multiple hidden layers and demonstrating universal function approximation ability [55].

Overall, these algorithms represent fundamental paradigms in machine learning, each offering distinct advantages in classification tasks while maintaining computational efficiency and practical applicability.

3.3.3 Heterogeneous ensemble approach

Although individual base models demonstrate computational efficiency due to their relatively simple structures, they may still struggle to maintain stable predictive performance when applied to large-scale geospatial datasets and complex nonlinear classification problems [55]. To address these limitations while preserving computational tractability, heterogeneous ensemble learning frameworks are adopted to integrate multiple base learners through complementary error-correction mechanisms. In this study, two ensemble strategies are employed (Fig. 3(d)): Stacking and Boosting.

(1) *Stacking*

Stacking (Stacked Generalization), introduced by Wolpert (1992), is a two-stage heterogeneous ensemble learning framework. In the first stage, multiple base learners (e.g., SVM, RF, and MLP) generate out-of-fold predictions through k -fold cross-validation:

$$z_i^{(m)} = h_m(D_i^{\text{train}}), \quad i = 1, 2, \dots, K \quad (16)$$

where D_i^{train} denotes the training subset in the i -th fold, and h_m represents the m -th base learner, with a total of M models. This strategy reduces overfitting by ensuring that predictions are generated on unseen validation folds. The resulting meta-features $Z = \{z_i^{(m)}\}$ are then used to train a meta-learner g (typically logistic regression):

$$\hat{y} = g(Z) \quad (17)$$

where $\hat{y}$ denotes the final prediction. This architecture leverages model heterogeneity and error diversity, thereby improving generalization performance through variance reduction.

433 (2) *Boosting*

434 Boosting is a sequential ensemble learning strategy that constructs a strong clas-
 435 sifier by iteratively focusing on misclassified samples through adaptive reweighting
 436 [55]. Given a dataset $\{(x_i, y_i)\}_{i=1}^N$, sample weights are initialized uniformly. At each
 437 iteration t , a base learner is trained to minimize the weighted classification error:

$$\varepsilon_t = \sum_{i=1}^N w_i^{(t)} \cdot \mathbb{I}(h_t(x_i) \neq y_i) \quad (18)$$

438 where $w_i^{(t)}$ denotes the sample weight at iteration t , $\mathbb{I}(\cdot)$ is the indicator func-
 439 tion, and $h_t(x_i)$ and y_i represent the predicted and true labels, respectively. The
 440 contribution of each learner is then computed as:

$$\alpha_t = \frac{1}{2} \ln \left(\frac{1 - \varepsilon_t}{\varepsilon_t} \right) \quad (19)$$

441 Sample weights are updated to emphasize misclassified instances:

$$w_i^{(t+1)} = \frac{w_i^{(t)} \exp(-\alpha_t y_i h_t(x_i))}{Z_t} \quad (20)$$

442 where Z_t is a normalization constant ensuring $\sum_i w_i^{(t+1)} = 1$. The final strong
 443 classifier is obtained through weighted voting of all weak learners:

$$\hat{y} = \text{sign} \left(\sum_{t=1}^T \alpha_t h_t(x) \right) \quad (21)$$

444 3.3.4 Performance metrics

445 In this study, to comprehensively evaluate model performance, we employ
 446 six complementary metrics, including Accuracy (ACC), Precision, Recall, F1-score,
 447 Receiver Operating Characteristic (ROC) curve, and Area Under the Curve (AUC).
 448 The definitions and formulations of these metrics are summarized in Table 3.

Table 3 Model performance metrics.

Performance metrics	Formula	Explanation
Accuracy (ACC)	$\frac{TP+TN}{TP+TN+FP+FN}$	TP: True Positive; FP: False Positive; TN: True Negative; FN: False Negative
Precision	$\frac{TP}{TP+FP}$	TP: True Positive; FP: False Positive
Recall	$\frac{TP}{TP+FN}$	TP: True Positive; FN: False Negative
F1-score	$\frac{2 \cdot \text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}$	Harmonic mean of Precision and Recall
ROC	TPR vs. FPR curve	TPR: True Positive Rate; FPR: False Positive Rate
AUC	Area under ROC curve	Measures overall classification performance

449 **Note:** TP represents the number of positive samples correctly predicted as positive,
 450 TN denotes the number of negative samples correctly predicted as negative, FP indi-
 451 cates the number of negative samples incorrectly predicted as positive, and FN refers
 452 to the number of positive samples incorrectly predicted as negative. The True Posi-
 453 tive Rate (TPR) represents the proportion of positive samples correctly predicted as
 454 positive, while the False Positive Rate (FPR) represents the proportion of negative
 455 samples incorrectly predicted as positive.

456 4 Results

457 4.1 Dominant environmental controls

458 The initiation of thaw slumps emerges from nonlinear interactions among coupled
 459 environmental drivers, including thermal regimes, topographical gradients, hydrolog-
 460 ical dynamics, and pedological properties. Traditional approaches that model these
 461 factors in isolation often suffer from overparameterization and poor out-of-distribution
 462 generalization. To address this, we first quantified multivariate associations between
 463 these environmental variables and thaw slump distribution, aiming to uncover their
 464 underlying structure and identify the most informative predictors, thereby improving
 465 the robustness and accuracy of susceptibility models.

466 The statistical tests applied to 12 candidate environmental variables indicate that
 467 Bartlett's test is highly significant ($\chi^2 = 210,577.414$, $p < 0.001$), and the KMO value
 468 (0.71) confirms sampling adequacy for factor analysis.

Table 4 Factor loadings of the observed variables

Factors	ML ₁	ML ₂	ML ₃	ML ₄	h^2
TDD	0.9592	-	-	-0.1001	0.9249
T-max	0.9090	-	-	-	0.9059
FDD	0.9081	-0.1502	-	-	0.8952
Ground ice	-0.7947	-0.3878	-	-	0.7130
R-max	-0.2946	0.8745	-	-	0.9565
Rainfall	0.4461	0.7205	-	-	0.9195
TWI	-	-	0.7861	-	0.8113
SPI	-	-	0.7803	-	0.8036
NDVI	0.1592	-	-	0.8775	0.8005
Soil	-0.3626	-	-	0.7005	0.7144
Slope	-	0.1106	-0.2417	0.1260	0.0923
Permafrost type	-	-0.6877	-	-0.1077	0.4898

a: Latent factors. b: Communality.

469 The results indicate that slope and permafrost type contribute minimally to
 470 the dominant factors and were therefore excluded from subsequent analysis. The
 471 remaining variables were grouped into four interpretable latent factors based on fac-
 472 tor loadings. (1) The temperature factor (ML₁) is strongly associated with total

473 and frozen-degree days, summer maximum temperature, and ground ice content; (2)
 474 the precipitation factor (ML₂) is dominated by annual and peak rainfall; (3) the
 475 terrain–hydrology factor (ML₃) reflects topographic wetness and flow accumulation
 476 characteristics (TWI and SPI); and (4) the soil–vegetation factor (ML₄) captures
 477 NDVI and soil properties, representing vegetation cover and substrate conditions.
 478 These four factors explain 84.45% of the total variance, and all variables exhibit strong
 479 communalities ($h^2 > 0.7$).

480 A random forest model was further used to quantify the relative importance of
 481 these grouped factors for thaw slump susceptibility (Fig. 4). The results indicate that
 482 precipitation is the most influential predictor, followed by temperature and soil-related
 483 factors, while terrain–hydrology exerts a comparatively weaker influence.

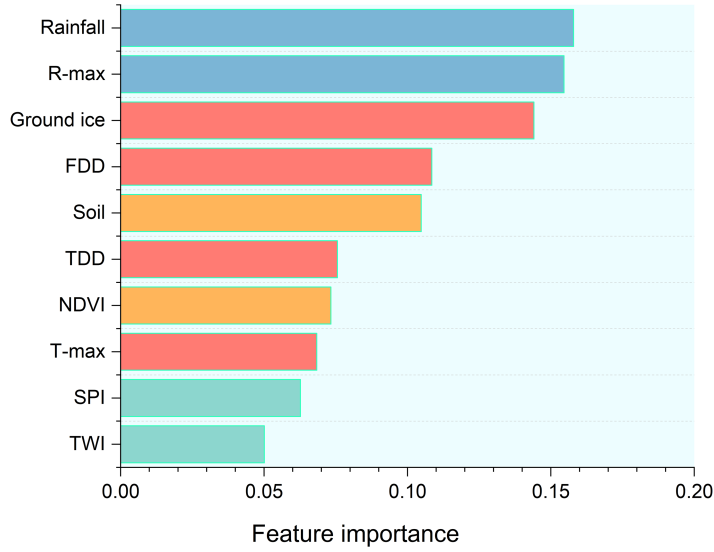


Fig. 4 Feature importance based on random forest algorithm

484 Subsequently, each environmental variable was categorized, and frequency ratio
 485 (FR) analysis was applied to quantify thaw slump susceptibility across different
 486 ranges. Higher FR values indicate stronger associations with thaw slump occurrence.
 487 The results show that the highest probabilities occur under annual precipitation of
 488 267.3–342.9 mm and summer peak rainfall of 816–966 mm, as well as in permafrost
 489 zones with 21–42% ground ice content and summer air temperatures of 4.8–6.5°C (see
 490 Table 8 in the Supplementary Materials). These ranges are consistent with metastable
 491 permafrost conditions reported in circumpolar regions (Obu et al., 2019).

492 Lower TWI and SPI values are more prevalent in thaw slump-prone areas,
 493 suggesting that limited subsurface drainage may enhance thaw slump initiation. Inter-
 494 estingly, higher NDVI values are also associated with increased susceptibility, possibly

reflecting vegetation-mediated thermal effects that accelerate ground ice degradation. Soils with higher proportions of sand and clay exhibit greater thaw slump density, likely due to reduced cohesion during thaw consolidation. Overall, this multi-step process reduces redundancy among variables and enhances the interpretability of the susceptibility modeling framework.

4.2 Model performance

Here, we implemented a hierarchical density-based clustering approach (HDBSCAN) to address potential sampling bias in negative samples and enhance the robustness of RTS susceptibility modeling. The silhouette coefficient of 0.4051 indicates that the clustering effectively captures structured patterns within the high-dimensional feature space, enabling reliable differentiation of RTS-prone regions (Fig. 5a).

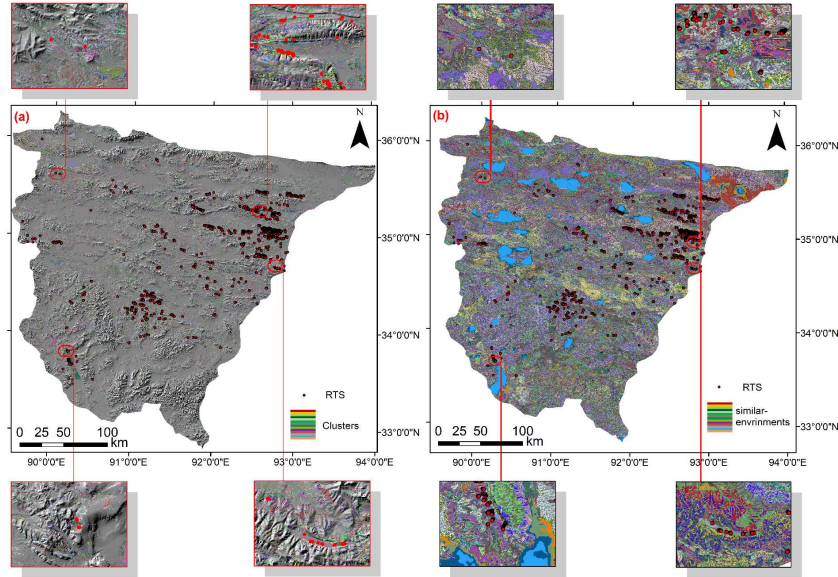


Fig. 5 HDBSCAN clustering results and similar environmental zones. (a) HDBSCAN clustering results of thawing slumps. (b) Similar environmental zones of thawing slumps.

After clustering, the study area was segmented into 2401 environmentally homogeneous units based on the FR-normalized weights of four geological factors (Fig. 6).

The similar environmental zones and the distribution of RTS sample quantities in various environments are presented in Fig. 5b and Fig. 7, respectively.

Notably, the aggregation of RTS occurrences within specific clusters highlights the influence of distinct environmental regimes on slump initiation. To construct a representative negative sample set, samples within HDBSCAN-identified RTS clusters were excluded, and non-RTS samples were proportionally selected from each

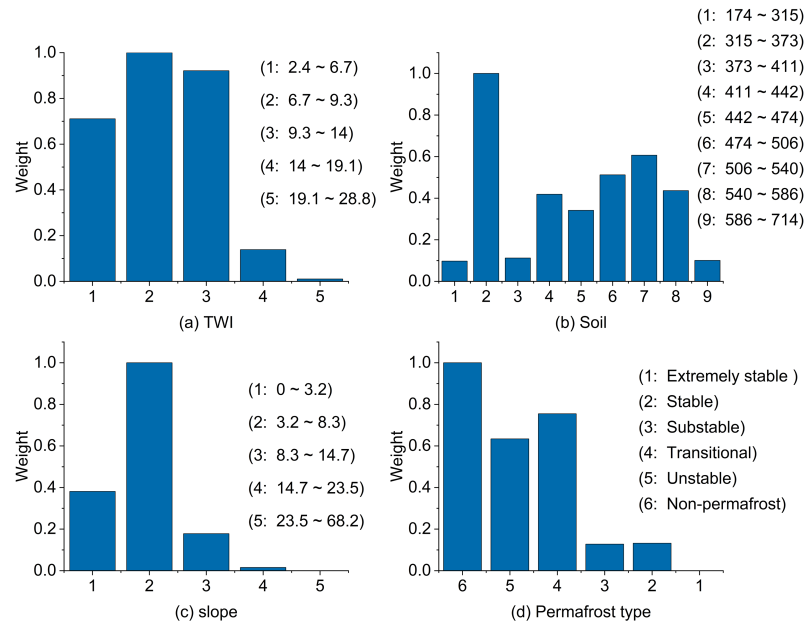


Fig. 6 Classes and FR weights of disaster-pregnant environmental factors.

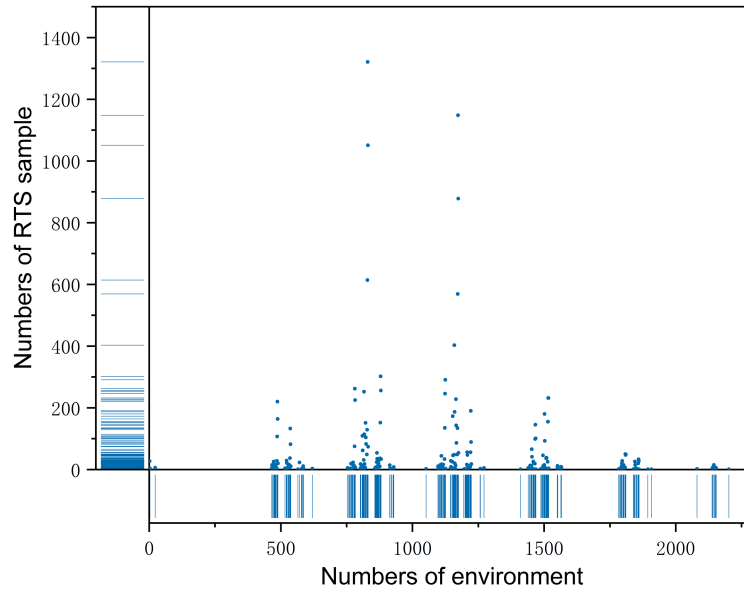


Fig. 7 Number of RTS samples distributed in each similar environment.

homogeneous unit, maintaining a 1:1 balance. This geostatistically informed sampling strategy reduces spatial and environmental bias, improving dataset balance and model generalization capability.

Table 5 Base-model performance

Model	Accuracy	Precision	Recall	F1-score	AUC
SVM	0.8285	0.8009	0.8775	0.8375	0.8920
XGB	0.8723	0.8627	0.8877	0.8750	0.9393
LR	0.8289	0.8090	0.8645	0.8358	0.8938
RF	0.8852	0.8706	0.9070	0.8884	0.9468
MLP	0.8316	0.8199	0.8530	0.8361	0.8946

As shown in Table 5, the RF algorithm demonstrates superior discriminatory performance among baseline classifiers, achieving the highest accuracy (0.8852). It outperforms XGBoost (0.8723), MLP (0.8316), Logistic Regression (0.8289), and SVM (0.8285), supporting its selection as the meta-learner in the stacking ensemble framework.

Table 6 Performance of different heterogeneous ensemble models

Model	Non-RTSs	Accuracy	Precision	Recall	F1-score	AUC
Stacking	RM	0.9371	0.9301	0.9462	0.9381	0.9770
Stacking	HSM	0.9412	0.9341	0.9504	0.9422	0.9777
Boosting	RM	0.9380	0.9169	0.9643	0.9400	0.9800
Boosting	HSM	0.9482	0.9400	0.9577	0.9485	0.9853

The ensemble learning framework achieves improved performance through synergistic integration of feature learning and sampling optimization (Table 6). The ROC curves further demonstrate strong model fitting ability (Fig. 8a).

Incorporating CNN significantly improves predictive performance, particularly when combined with the HDBSCAN-based Similar Environment Method (HSM) for negative sampling. The Boosting CNN-HSM model achieves the best performance, with an accuracy of 99.01%, recall of 99.72%, and AUC of 0.99, outperforming all baseline models. The Stacking ensemble also maintains strong generalization ability, with an F1-score of 0.92 and AUC of 0.96. Comparative results indicate that the HSM strategy consistently outperforms the Random Method (RM), confirming the advantage of density-aware spatial sampling in class-imbalanced geohazard modeling. The final Boosting-CNN-HSM framework is therefore adopted to generate high-resolution RTS susceptibility maps.

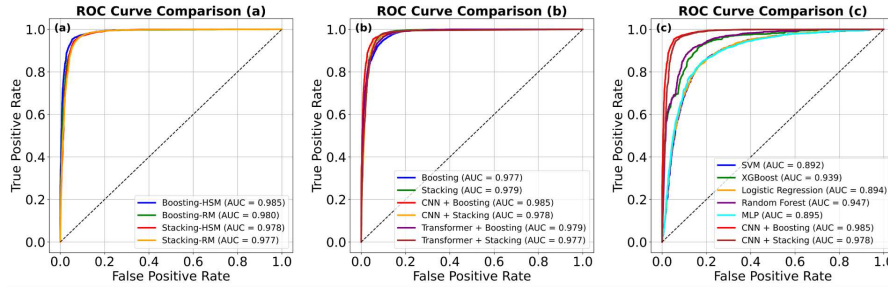


Fig. 8 ROC curves of different models using testing dataset: (a) Stacking-RM, Boosting-RM, Stacking-HSM, Boosting-HSM; (b) Stacking, Boosting, CNN+Stacking, CNN+Boosting, Transformer+Stacking, Transformer+Boosting; (c) SVM, XGBoost, LR, RF, MLP, CNN+Stacking, CNN+Boosting.

4.3 RTS susceptibility distribution

The RTS susceptibility maps were generated to validate the spatial predictive capability of different model-sampling integrations. Continuous susceptibility values were classified into five hazard probability levels: Very Low, Low, Moderate, High, and Very High, using the Jenks Natural Breaks Optimization algorithm, which maximizes inter-class variance while minimizing intra-class variance (Wang et al., 2019). Significant spatial heterogeneity is observed across different modeling and sampling configurations, as illustrated in Fig. 9.

The Stacking-RM model produces overly generalized susceptibility patterns, characterized by broad and poorly localized high-risk zones (Fig. 9a), indicating limited spatial discrimination and overprediction. Although the Boosting-RM model shows slight improvement, identifying several high-susceptibility areas in mountainous regions, the resulting map remains noisy with blurred boundaries and scattered false positives (Fig. 9b).

In contrast, incorporating the HSM strategy into the Stacking framework significantly improves spatial coherence. High-risk zones become more compact and better aligned with geomorphological transitions, particularly along slope-valley interfaces. However, some misclassification persists along boundary areas, where moderate-to-high susceptibility is assigned to low-risk regions (Fig. 9c), suggesting partial overfitting to complex terrain transitions.

Among all configurations, the Boosting-HSM model exhibits the highest spatial fidelity. As shown in Fig. 9d, its susceptibility map closely matches observed RTS distributions, particularly in the mountainous corridor west of the Qinghai-Tibet Highway. It effectively reduces false alarms in environmentally stable regions and produces smooth spatial transitions with well-defined class boundaries, demonstrating strong operational applicability.

Accordingly, the Boosting-HSM model is selected as the optimal configuration for final susceptibility mapping. The spatial extents of each susceptibility class are as follows: very low (6,114 km²), low (6,902 km²), moderate (3,400 km²), high (7,587 km²), and very high (7,734 km²). Notably, most observed RTS events fall within the high and very high susceptibility zones, confirming the model's predictive reliability.

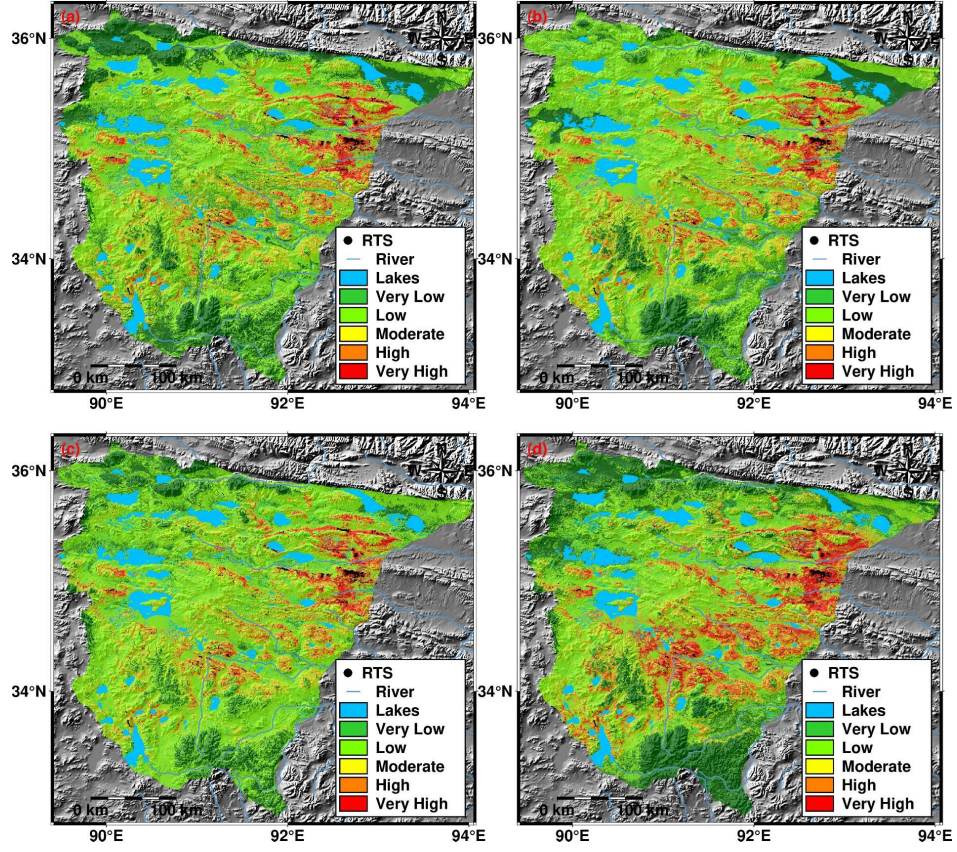


Fig. 9 RTS susceptibility mapping in the Hoh Xil. (a) Stacking-RM, (b) Boosting-RM, (c) Stacking-HSM, (d) Boosting-HSM.

The spatial patterns are highly consistent with the geomorphological characteristics of thaw slump development, further demonstrating the robustness and transferability of the proposed integrated modeling framework.

5 Discussion

5.1 Advantages of feature extraction

The predictive accuracy of RTS susceptibility models is intrinsically dependent on the quality and representation capacity of environmental features. In this study, six configurations were designed to evaluate the effectiveness of different feature extraction strategies, including (a, b) raw (non-extracted) features, (c, d) CNN-extracted features, and (e, f) Transformer-extracted features, each combined with either Stacking or Boosting classifiers. All experiments adopt the HSM sampling strategy to ensure data consistency.

Table 7 Performance of models based on different feature extraction methods

Model	Accuracy	Precision	Recall	F1-score	AUC
Boosting	0.9238	0.9151	0.9356	0.9252	0.9770
Stacking	0.9388	0.9298	0.9502	0.9399	0.9794
CNN + Boosting	0.9482	0.9400	0.9577	0.9485	0.9853
CNN + Stacking	0.9412	0.9341	0.9504	0.9422	0.9777
Transformer + Boosting	0.9315	0.9151	0.9522	0.9334	0.9788
Transformer + Stacking	0.9431	0.9269	0.9630	0.9446	0.9769

Quantitative evaluation indicates that CNN-extracted features consistently outperform other configurations in terms of both accuracy and stability (Table 7). The ROC curves on the validation set further confirm the superior discriminative ability and spatial coherence of CNN-based models (Fig. 8b). To further assess spatial prediction quality, susceptibility maps were generated for each configuration (Fig. 10).

The Stacking model using raw features produces overly smoothed susceptibility patterns, with diffuse high-risk zones that fail to capture terrain-specific hazard structures (Fig. 10a). The Boosting model slightly improves spatial granularity; however, fragmented patches and false positives remain in stable regions, highlighting the limitations of raw feature inputs (Fig. 10b).

In contrast, CNN-extracted features significantly enhance spatial resolution and hazard delineation. The Stacking–CNN model yields more coherent susceptibility zones that align well with topographic structures, particularly along valley slopes (Fig. 10c). The Boosting–CNN model achieves the best performance, producing highly refined spatial patterns that closely match observed RTS locations, with reduced false positives and smoother risk gradients (Fig. 10d). These results demonstrate the strong ability of CNNs to capture local geomorphological variability critical for thaw slump initiation.

Although Transformer architectures are effective in modeling long-range dependencies and enabling parallel computation (Vaswani et al., 2017), their performance in RTS feature extraction is less satisfactory. The Stacking–Transformer model exhibits overly diffused spatial patterns with weak localization ability (Fig. 10e). The Boosting–Transformer model shows slight improvement in matching observed RTS distributions but still lacks fine-scale spatial sharpness compared to CNN-based models (Fig. 10f). This limitation may be attributed to the difficulty of Transformers in capturing fine-grained 2D spatial dependencies without dense positional encoding.

Overall, these results highlight the critical importance of spatially aware feature extraction in improving model interpretability and predictive performance. CNNs are particularly effective in preserving local spatial continuity and capturing terrain-driven variability, thereby enabling more accurate and operationally meaningful susceptibility mapping. Among all configurations, the Boosting–CNN model demonstrates the best overall performance and is therefore identified as the most robust framework for RTS hazard assessment.

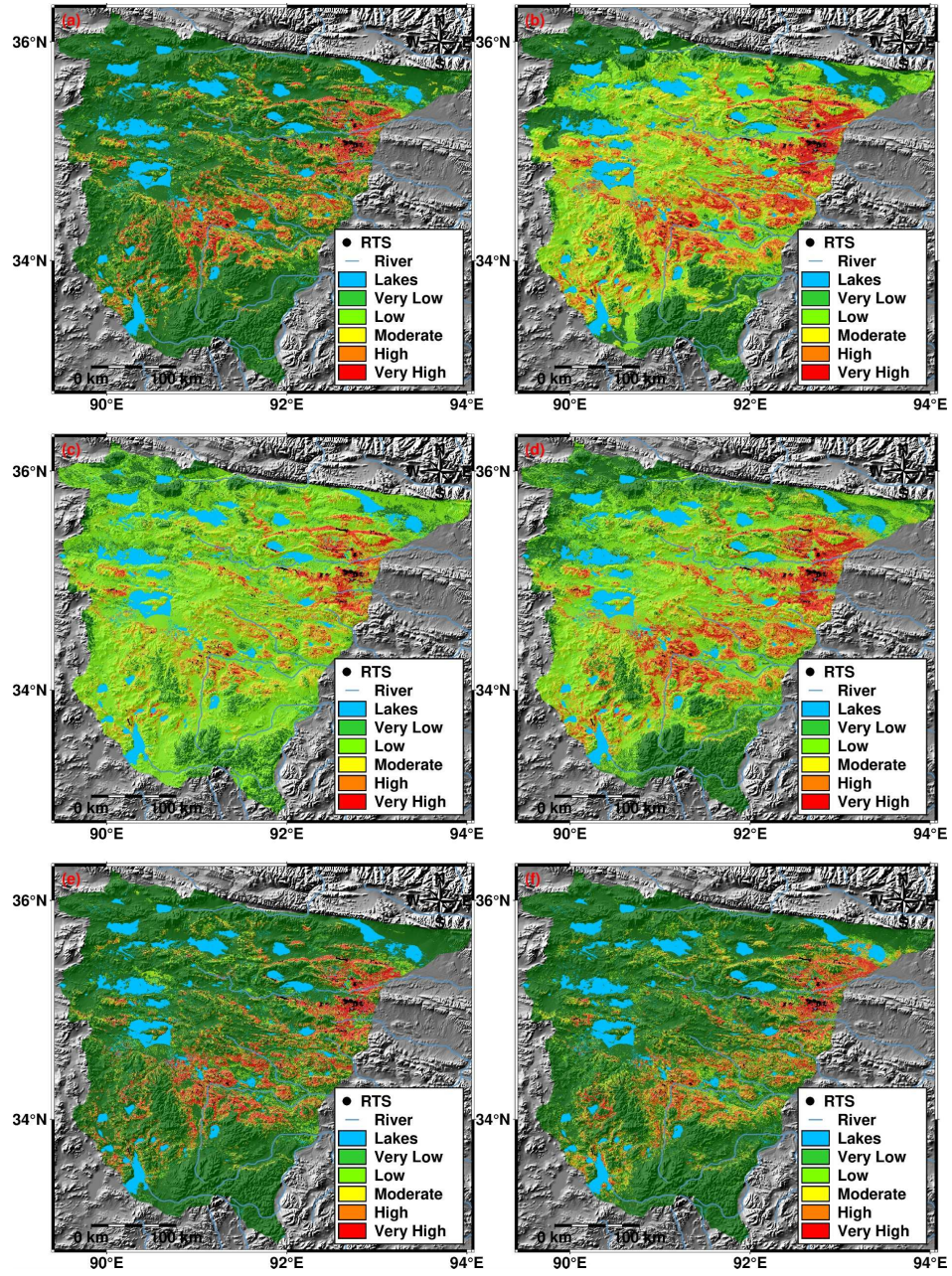


Fig. 10 RTS susceptibility mapping using different feature extraction methods in the Hoh Xil. (a) Stacking, (b) Boosting, (c) Stacking-CNN, (d) Boosting-CNN, (e) Stacking-Transformer, (f) Boosting-Transformer.

5.2 Powerful ensemble learning

Ensemble learning forms the methodological cornerstone of the RTS susceptibility modeling framework, serving as the primary classification strategy following feature extraction. It is designed to address the challenges associated with high-dimensional, nonlinear, and imbalanced geospatial data. By integrating multiple base learners, ensemble models achieve improved predictive accuracy and spatial generalization capability. A hierarchical evaluation framework was employed to systematically compare individual classifiers and ensemble models, with particular emphasis on their predictive performance and spatial discrimination ability in identifying RTS-prone regions. All experiments adopt the HSM sampling strategy and CNN-based feature representations to ensure data consistency.

As summarized in Tables 5 and 6, the RF model achieves the highest predictive accuracy (0.875) and AUC value (0.930) among all standalone classifiers (Fig. 8c), outperforming other individual models across most evaluation metrics. In contrast, SVM and MLP exhibit relatively weaker performance, likely due to limited sensitivity to feature scale and nonlinear interactions. The Stacking ensemble, which integrates predictions from multiple base learners via a meta-learner, improves overall performance compared to most individual classifiers, particularly in balancing precision and recall. However, Boosting-based ensemble models consistently outperform both individual classifiers and Stacking-based ensembles in terms of AUC and F1-score.

The Boosting model achieves the best overall performance across all evaluation metrics. This improvement is attributed to its iterative optimization mechanism and adaptive weighting of misclassified samples, which enhance its ability to capture complex nonlinear decision boundaries. Moreover, the Boosting model achieves the highest AUC, with its ROC curve consistently dominating other models (Fig. 8c), demonstrating superior classification capability.

The spatial distribution of RTS susceptibility for each model is presented in Fig. 11. The SVM model produces highly fragmented susceptibility patterns, with scattered high-risk zones and poor alignment with known geomorphological features (Fig. 11a). The XGBoost model improves spatial coherence and better delineates high-risk clusters in mountainous regions; however, overestimation persists in stable terrain (Fig. 11b). The Logistic Regression model generates overly simplified susceptibility maps, failing to capture fine-scale spatial variability (Fig. 11c). The RF model produces more structured spatial patterns, although some high-risk regions remain weakly defined (Fig. 11d). The MLP model shows inconsistent gradients and blurred class boundaries, indicating limited spatial sensitivity (Fig. 11e).

In contrast, ensemble frameworks significantly improve spatial prediction performance. The Stacking model better identifies high-risk corridors and produces smoother susceptibility transitions (Fig. 11f), although some overprediction remains along class boundaries, reflecting sensitivity to local noise. The Boosting model achieves the most refined spatial output (Fig. 11g), with sharply delineated high-susceptibility zones that closely align with geomorphological indicators of RTS occurrence. False positives in stable regions are significantly reduced, while low-risk areas are consistently identified, indicating a balanced performance in terms of sensitivity and specificity.

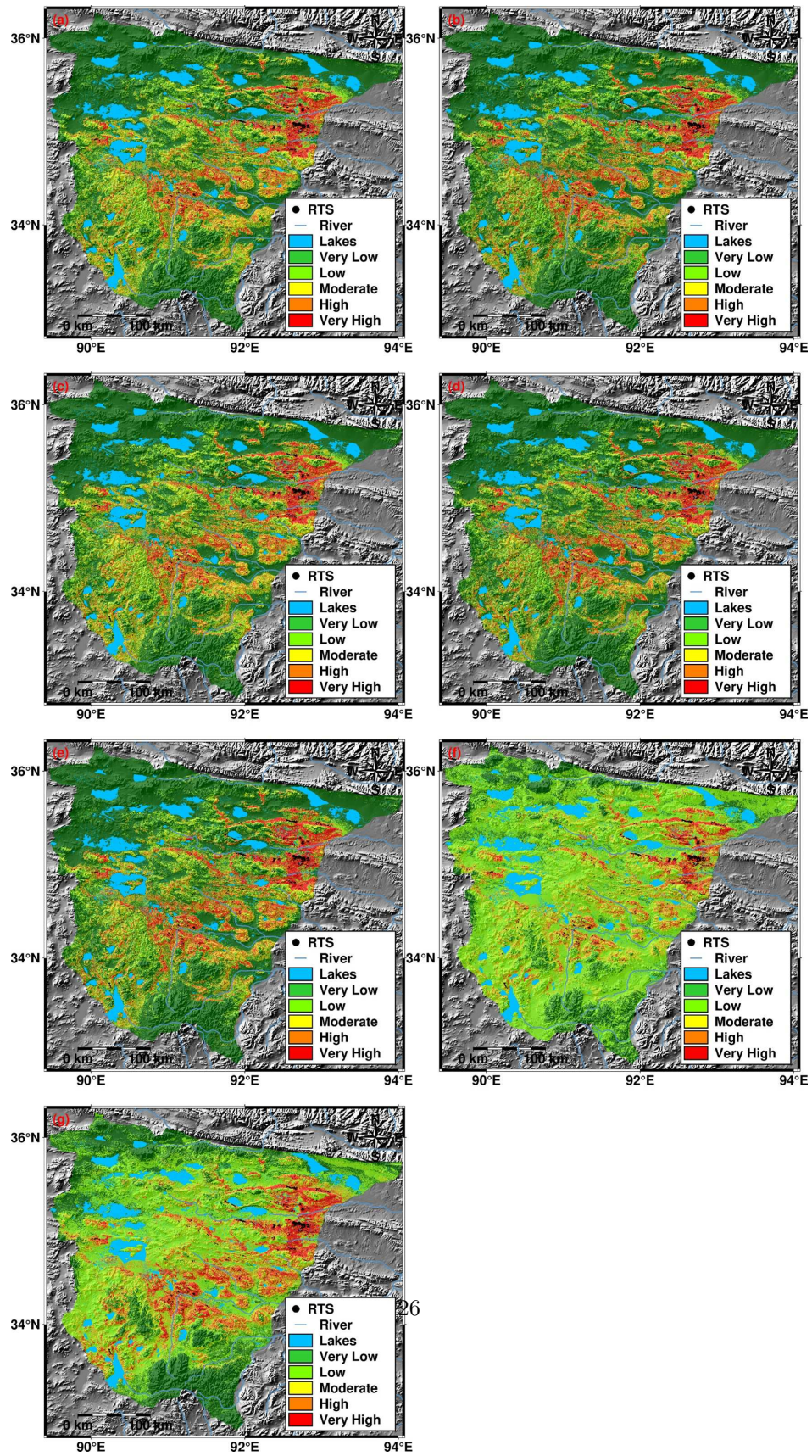


Fig. 11 Visualization of RTS susceptibility mapping using different basic models and ensemble models in the Hoh Xil. (a) SVM, (b) XGBoost, (c) LR, (d) RF, (e) MLP, (f) Stacking-CNN, (g) Boosting-CNN.

Overall, these results highlight the critical role of ensemble learning in enhancing model robustness and generalization ability. Boosting-based methods demonstrate strong capability in integrating heterogeneous predictive signals and mitigating overfitting under complex terrain conditions. When combined with CNN-based feature extraction, ensemble learning becomes particularly effective for RTS susceptibility mapping in highly heterogeneous landscapes. The Boosting–CNN framework thus emerges as the most robust and operationally reliable approach for high-resolution RTS hazard assessment.

6 Conclusion

The increasing occurrence of retrogressive thaw slumps (RTSs) under accelerated climate warming poses significant risks to infrastructure and environmental stability in permafrost regions. To address this challenge, this study proposes a robust and scalable framework that integrates multivariate factor analysis, deep learning-based spatial feature extraction, and ensemble learning to improve the accuracy and reliability of RTS susceptibility assessment.

Applied to the Hoh Xil region on the Qinghai–Tibet Plateau, where complex periglacial environments support widespread RTS development, the proposed framework effectively captures the nonlinear interactions among key environmental drivers controlling slope instability. The results indicate that hydrothermal conditions are the dominant triggering mechanism, particularly under the combined influence of extreme precipitation events and abnormal freeze–thaw cycles. In addition, predisposing factors such as high ground ice content and fine-grained soil textures further enhance slope failure susceptibility.

The incorporation of clustering-based negative sample selection constrained by environmental similarity effectively mitigates sampling bias and improves model generalization. Furthermore, convolutional neural networks (CNNs) enable the extraction of spatially sensitive terrain features associated with RTS initiation, thereby supporting the generation of high-resolution susceptibility maps that are highly consistent with observed RTS distributions.

Comparative results from ensemble modeling demonstrate that the Boosting-based framework achieves the best predictive performance, with an AUC of 0.9853, significantly outperforming all baseline models. Overall, the proposed integrated approach not only improves predictive accuracy but also enhances interpretability of RTS controlling mechanisms. This framework provides a transferable and scalable solution for climate-adaptive hazard assessment and infrastructure planning in cold-region environments.

Acknowledgements. The authors would like to thank the Tibetan Plateau Data Center for providing the RTS inventory and environmental datasets used in this study.

Declarations

- Funding

697 This research was funded by the Sichuan Science and Technology Program (Grant
698 Nos.2019YJ0504); the National Natural Science Foundation of China (Grant Nos.
699 42271461); 2021 Chengdu University of Technology Young and Middle-aged Key
700 Teachers Program.

- 701 • Conflict of interest

702 The authors declare that they have no conflict of interest.

- 703 • Ethics approval

704 Not applicable.

- 705 • Consent for publication

706 Not applicable.

- 707 • Data availability

708 The datasets generated and analyzed during the current study are available from
709 the corresponding author upon reasonable request.

- 710 • Code availability

711 The code used in this study is available from the corresponding author upon
712 reasonable request.

- 713 • Author contribution

714 Tanshuo Li: Conceptualization, methodology, data curation, formal analysis, soft-
715 ware, visualization, writing—original draft preparation.

716 Jiang Ye: Supervision, conceptualization, funding acquisition, methodology,
717 writing—review and editing.

718 Huiyuan Luo: Data processing, validation, investigation, writing—review and edit-
719 ing.

720 Qiang Xu: Resources, supervision, project administration, writing—review and
721 editing.

722 References

- 723 [1] Romanovsky, V.E., Smith, S.L., Christiansen, H.H.: Permafrost thermal state in
724 the polar northern hemisphere during the international polar year 2007–2009:
725 A synthesis. *Permafrost and Periglacial Processes* **21**(2), 106–116 (2010) <https://doi.org/10.1002/ppp.689>
726
- 727 [2] Zhao, L., Wu, Q., Marchenko, S., Sharkhuu, N.: Thermal state of permafrost and
728 active layer in central asia during the international polar year. *Permafrost and*
729 *Periglacial Processes* **21**(2), 198–207 (2010) <https://doi.org/10.1002/ppp.688>
- 730 [3] Chadburn, S., Burke, E., Cox, P., Friedlingstein, P., Hugelius, G., Westermann,
731 S.: An observation-based constraint on permafrost loss as a function of global
732 warming. *Nature Climate Change* **7**(5), 340–344 (2017)
- 733 [4] Runge, A., Nitze, I., Grosse, G.: Remote sensing annual dynamics of rapid per-
734 mafrost thaw disturbances with landtrendr. *Remote Sensing of Environment* **268**,
735 112752 (2022) <https://doi.org/10.1016/j.rse.2021.112752>

- 736 [5] Yang, D., Qiu, H., Ye, B., Liu, Y., Zhang, J., Zhu, Y.: Distribution and recur-
737 rence of warming-induced retrogressive thaw slumps on the central qinghai-tibet
738 plateau. *Journal of Geophysical Research: Earth Surface* **128**(8), 2022–007047
739 (2023) <https://doi.org/10.1029/2022JF007047>
- 740 [6] Lin, Z., Niu, F., Xu, Z., Xu, J., Wang, P.: Thermal regime of a thermokarst lake
741 and its influence on permafrost, beiluhe basin, qinghai–tibet plateau. *Permafrost
742 and Periglacial Processes* **21**(4), 315–324 (2010) [https://doi.org/10.1002/ppp.
743 692](https://doi.org/10.1002/ppp.692)
- 744 [7] Niu, F., Luo, J., Lin, Z., Fang, J., Liu, M.: Thaw-induced slope failures and stabil-
745 ity analyses in permafrost regions of the qinghai-tibet plateau, china. *Landslides*
746 **13**, 55–65 (2016)
- 747 [8] Costard, F., Dupeyrat, L., Séjourné, A., Bouchard, F., Fedorov, A., Saint-Bézar,
748 B.: Retrogressive thaw slumps on ice-rich permafrost under degradation: Results
749 from a large-scale laboratory simulation. *Geophysical Research Letters* **48**(1),
750 2020–091070 (2021)
- 751 [9] Liu, Y., Qiu, H., Kamp, U., Wang, N., Wang, J., Huang, C., Tang, B.: Higher tem-
752 perature sensitivity of retrogressive thaw slump activity in the arctic compared
753 to the third pole. *Science of the Total Environment* **914**, 170007 (2024)
- 754 [10] Jin, H., Huang, Y., Bense, V.F., Ma, Q., Marchenko, S.S., Shepelev, V.V., Hu, Y.,
755 Liang, S., Spektor, V.V., Jin, X.: Permafrost degradation and its hydrogeological
756 impacts. *Water* **14**(3), 372 (2022) <https://doi.org/10.3390/w14030372>
- 757 [11] Fan, X., Lin, Z., Yao, M., Wang, Y., Gu, Q., Luo, J., Wu, X., Gao, Z.: Spatio-
758 temporal evolution and susceptibility assessment of thaw slumps associated with
759 climate change in the hoh xil region, in the hinterland of the qinghai–tibet
760 plateau. *Remote Sensing* **17**(9), 1614 (2025)
- 761 [12] Yang, Z., Ni, W., Niu, F., Li, L., Ren, S.: Spatiotemporal distribution character-
762 istics and influencing factors of freeze–thaw erosion in the qinghai–tibet plateau.
763 *Remote Sensing* **16**(9), 1629 (2024)
- 764 [13] Lewkowicz, A.G.: Rate of short-term ablation of exposed ground ice, banks island,
765 northwest territories, canada. *Journal of Glaciology* **32**(112), 511–519 (1986)
- 766 [14] Lantz, T.C., Kokelj, S.V.: Increasing rates of retrogressive thaw slump activity
767 in the mackenzie delta region, nwt, canada. *Geophysical Research Letters* **35**(6)
768 (2008)
- 769 [15] Huang, F., Yan, J., Fan, X., Yao, C., Huang, J., Chen, W., Hong, H.: Uncer-
770 tainty pattern in landslide susceptibility prediction modelling: Effects of different
771 landslide boundaries and spatial shape expressions. *Geoscience Frontiers* **13**(2)
772 (2022) <https://doi.org/10.1016/j.gsf.2021.101317>

- 773 [16] Lacelle, D., Brooker, A., Fraser, R.H., Kokelj, S.V.: Distribution and growth
774 of thaw slumps in the richardson mountains–peel plateau region, northwestern
775 canada. *Geomorphology* **235**, 40–51 (2015) [https://doi.org/10.1016/j.geomorph.](https://doi.org/10.1016/j.geomorph.2015.01.024)
776 [2015.01.024](https://doi.org/10.1016/j.geomorph.2015.01.024)
- 777 [17] Luo, J., Niu, F., Lin, Z., Liu, M., Yin, G., Gao, Z.: Inventory and frequency of
778 retrogressive thaw slumps in permafrost region of the qinghai–tibet plateau. *Geo-*
779 *physical Research Letters* **49**(23), 2022–099829 (2022) [https://doi.org/10.1029/](https://doi.org/10.1029/2022GL099829)
780 [2022GL099829](https://doi.org/10.1029/2022GL099829)
- 781 [18] Luo, J., Niu, F., Lin, Z., Liu, M., Yin, G.: Recent acceleration of thaw slumping in
782 permafrost terrain of qinghai–tibet plateau: An example from the beiluhe region.
783 *Geomorphology* **341**, 79–85 (2019)
- 784 [19] Xia, Z., Huang, L., Fan, C., Jia, S., Lin, Z., Liu, L., Luo, J., Niu, F., Zhang, T.:
785 Retrogressive thaw slumps along the qinghai–tibet engineering corridor: A com-
786 prehensive inventory and their distribution characteristics. *Earth System Science*
787 *Data Discussions* **2022**, 1–19 (2022)
- 788 [20] Lantuit, H., Pollard, W.: Temporal stereophotogrammetric analysis of retrogres-
789 sive thaw slumps on herschel island, yukon territory. *Natural Hazards and Earth*
790 *System Sciences* **5**(3), 413–423 (2005) <https://doi.org/10.5194/nhess-5-413-2005>
- 791 [21] Lantuit, H., Pollard, W.: Fifty years of coastal erosion and retrogressive thaw
792 slump activity on herschel island, southern beaufort sea, yukon territory, canada.
793 *Geomorphology* **95**(1–2), 84–102 (2008) [https://doi.org/10.1016/j.geomorph.](https://doi.org/10.1016/j.geomorph.2006.07.040)
794 [2006.07.040](https://doi.org/10.1016/j.geomorph.2006.07.040)
- 795 [22] Eini, M., Kaboli, H.S., Rashidian, M., Hedayat, H.: Hazard and vulnerability in
796 urban flood risk mapping: Machine learning techniques and considering the role
797 of urban districts. *International Journal of Disaster Risk Reduction* **50**, 101687
798 (2020) <https://doi.org/10.1016/j.ijdrr.2020.101687>
- 799 [23] Rafiei-Sardooi, E., Azareh, A., Choubin, B., Mosavi, A.H., Clague, J.J.: Eval-
800 uating urban flood risk using hybrid method of topsis and machine learning.
801 *International Journal of Disaster Risk Reduction* **66**, 102614 (2021) [https://doi.](https://doi.org/10.1016/j.ijdrr.2021.102614)
802 [org/10.1016/j.ijdrr.2021.102614](https://doi.org/10.1016/j.ijdrr.2021.102614)
- 803 [24] Li, Z.-J., Zong-Xing, L., Ling-Ling, S., Jin-Zhu, M.: Characteristic and factors
804 of stable isotope in precipitation in the source region of the yangtze river. *Agri-*
805 *cultural and Forest Meteorology* **281**, 107825 (2020) [https://doi.org/10.1016/j.](https://doi.org/10.1016/j.agrformet.2019.107825)
806 [agrformet.2019.107825](https://doi.org/10.1016/j.agrformet.2019.107825)
- 807 [25] Zhang, Y., Zhao, L., Xie, C., Wu, T., Wu, X., Yang, G., Yang, S., Wang, W.,
808 Pang, Q., Liu, G.: Permafrost characteristics and potential influencing factors in
809 the lake regions of hoh xil, qinghai–tibet plateau. *Geoderma* **437**, 116572 (2023)
810 <https://doi.org/10.1016/j.geoderma.2023.116572>

- [26] Yin, G., Niu, F., Luo, J., Wang, J., Liu, M., Dong, T., Cao, Y., Li, A.: High potential for pile-bearing capacity loss and ground subsidence over permafrost regions across the northern hemisphere. *Global and Planetary Change* **226**, 104156 (2023) <https://doi.org/10.1016/j.gloplacha.2023.104156>
- [27] Fan, Y., Du, H., Han, Z., Liu, X., Liu, X., Li, Z.: Assessment of dust emissions and their controlling factors on the hoh xil, north-central part of the qinghai-tibetan plateau. *Environmental and Sustainability Indicators* **24**, 100487 (2024) <https://doi.org/10.1016/j.indic.2024.100487>
- [28] Lin, Z., Fan, X., Yao, M.: Distribution Data of Thaw Slump/Landslide in Hoh Xil Area (2018–2019)
- [29] Jones, M.K.W., Pollard, W.H., Jones, B.M.: Rapid initialization of retrogressive thaw slumps in the canadian high arctic and their response to climate and terrain factors. *Environmental Research Letters* **14**(5), 055006 (2019)
- [30] Yin, G., Luo, J., Niu, F., Lin, Z., Liu, M.: Machine learning-based thermokarst landslide susceptibility modeling across the permafrost region on the qinghai-tibet plateau. *Landslides* **18**(7), 2639–2649 (2021)
- [31] González, C., Bachmann, M., Bueso-Bello, J.-L., Rizzoli, P., Zink, M.: A fully automatic algorithm for editing the tandem-x global dem. *Remote Sensing* **12**(23) (2020) <https://doi.org/10.3390/rs12233961>
- [32] Wilson, J.P., Gallant, J.C.: *Terrain Analysis: Principles and Applications*. John Wiley & Sons, ??? (2000)
- [33] Beven, K.J., Kirkby, M.J.: A physically based, variable contributing area model of basin hydrology/un modèle à base physique de zone d’appel variable de l’hydrologie du bassin versant. *Hydrological Sciences Journal* **24**(1), 43–69 (1979)
- [34] Yang, Y., You, Q.: Dataset of 0.1° Monthly Surface Air Temperature over the Tibetan Plateau (1901–2020)
- [35] Peng, S.: 1-km Monthly Precipitation Dataset for China (1901–2023)
- [36] Poggio, L., De Sousa, L.M., Batjes, N.H., Heuvelink, G.B., Kempen, B., Ribeiro, E., Rossiter, D.: Soilgrids 2.0: Producing soil information for the globe with quantified spatial uncertainty. *SOIL* **7**(1), 217–240 (2021)
- [37] Gao, J., Shi, Y., Zhang, H., Chen, X., Zhang, W., Shen, W., Xiao, T., Zhang, Y.: China Regional 250m Normalized Difference Vegetation Index Data Set (2000–2023)
- [38] Cheng, G., Li, S., Nan, Z., Tong, B.: Map of Permafrost on the Qinghai-Tibet Plateau (1:3,000,000) (1983–1996)

- [39] Zou, D., Pang, Q., Zhao, L., Wang, L., Hu, G., Du, E., Liu, G., Liu, S., Liu, Y.: Permafrost Ground-Ice Content to 10 M Depth on the Qinghai-Tibet Plateau
- [40] Niu, F., Luo, J., Lin, Z., Liu, M., Yin, G.: Thaw-induced slope failures and susceptibility mapping in permafrost regions of the qinghai-tibet engineering corridor, china. *Natural Hazards* **74**, 1667–1682 (2014)
- [41] Yin, G., Luo, J., Niu, F., Liu, M., Gao, Z., Dong, T., Ni, W.: High-resolution assessment of retrogressive thaw slump susceptibility in the qinghai-tibet engineering corridor. *Research in Cold and Arid Regions* **15**(6), 288–294 (2023) <https://doi.org/10.1016/j.rcar.2023.11.006>
- [42] Lewkowicz, A., Way, R.: Extremes of summer climate trigger thousands of thermokarst landslides in a high arctic environment. *Nature Communications* **10**, 1329 (2019)
- [43] Kokelj, S.V., Kokoszka, J., Der Sluijs, J., Rudy, A.C., Tunnicliffe, J., Shakil, S., Tank, S.E., Zolkos, S.: Thaw-driven mass wasting couples slopes with downstream systems, and effects propagate through arctic drainage networks. *The Cryosphere* **15**(7), 3059–3081 (2021)
- [44] Lou, P., Wu, T., Chen, J., Fu, B., Zhu, X., Chen, J., Wu, X., Yang, S., Li, R., Lin, X.: Recognition of thaw slumps based on machine learning and uavs: A case study in the qilian mountains, northeastern qinghai-tibet plateau. *International Journal of Applied Earth Observation and Geoinformation* **116**, 103163 (2023) <https://doi.org/10.1016/j.jag.2022.103163>
- [45] Beyle, H.C.: The vectors of mind-multiple-factor analysis for the isolation of primary traits. *Public Opinion Quarterly* **1**(4), 146–146 (1937) <https://doi.org/10.1037/10018-000>
- [46] Kaiser, H.F.: An index of factorial simplicity. *Psychometrika* **39**(1), 31–36 (1974)
- [47] Kaiser, H.F., Rice, J.: Little jiffy, mark iv. *Educational and Psychological Measurement* **34**(1), 111–117 (1974)
- [48] Goretzko, D., Pham, T.T.H., Bühner, M.: Exploratory factor analysis: Current use, methodological developments and recommendations for good practice. *Current Psychology* **40**, 3510–3521 (2021)
- [49] Guo, Z., Tian, B., Zhu, Y., He, J., Zhang, T.: How do the landslide and non-landslide sampling strategies impact landslide susceptibility assessment? — a catchment-scale case study from china. *Journal of Rock Mechanics and Geotechnical Engineering* **16**(3), 877–894 (2024) <https://doi.org/10.1016/j.jrmge.2023.07.026>
- [50] He, G., Gao, X., Li, L., Gao, P.: Oct monitoring data processing method of laser

- 882 deep penetration welding based on hdbscan. *Optics & Laser Technology* **179**
883 (2024) <https://doi.org/10.1016/j.optlastec.2024.111303>
- 884 [51] Fan, T., Guo, N., Ren, Y.: Consumer clusters detection with geo-tagged social
885 network data using dbscan algorithm: A case study of the pearl river delta in
886 china. *GeoJournal* **86**(1), 317–337 (2021)
- 887 [52] Li, M., Bi, X., Wang, L., Han, X.: A method of two-stage clustering learning based
888 on improved dbscan and density peak algorithm. *Computer Communications*
889 **167**, 75–84 (2021) <https://doi.org/10.1016/j.comcom.2020.12.019>
- 890 [53] Wang, G.-W., Zhang, C.-X., Zhuang, J.: Clustering with prim’s sequential rep-
891 resentation of minimum spanning tree. *Applied Mathematics and Computation*
892 **247**, 521–534 (2014) <https://doi.org/10.1016/j.amc.2014.09.026>
- 893 [54] Lu, J., He, Y., Zhang, L., Zhang, Q., Gao, B., Chen, H., Fang, Y.: Ensemble
894 learning landslide susceptibility assessment with optimized non-landslide samples
895 selection. *Geomatics, Natural Hazards and Risk* **15**(1), 2378176 (2024) <https://doi.org/10.1080/19475705.2024.2378176>
896
- 897 [55] LeCun, Y., Boser, B., Denker, J.S., Henderson, D., Howard, R.E., Hubbard,
898 W., Jackel, L.D.: Backpropagation applied to handwritten zip code recognition.
899 *Neural Computation* **1**(4), 541–551 (1989) [https://doi.org/10.1162/neco.1989.1.](https://doi.org/10.1162/neco.1989.1.4.541)
900 [4.541](https://doi.org/10.1162/neco.1989.1.4.541)