Earth Syst. Sci. Data, 18, 3779–3801, 2026
https://doi.org/10.5194/essd-18-3779-2026

Open Access
Earth System
Science
Data

A physically consistent soil thickness map of the Qinghai–Tibet Plateau derived from coupled erosion mechanisms

Lihua Chen[1,2,★], **Xingyu Ding**[3,★], **Shuping Zhao**[1,2], **Fujun Niu**[1], **Keting Feng**[4], **and Zhuotong Nan**[1,2,5]

[1]Yangtze River Delta Urban Wetland Ecosystem National Field Scientific Observation and Research Station,
School of Environment and Geographic Sciences, Shanghai Normal University, Shanghai 200234, China
[2]Key Laboratory of Ministry of Education on Virtual Geographic Environment,
Nanjing Normal University, Nanjing 210023, China
[3]Hangzhou No. 14 High School Qingshanhu, Hangzhou 311300, China
[4]Institute of Earth Environment, Chinese Academy of Sciences, Xi'an 710061, China
[5]Jiangsu Center for Collaborative Innovation in Geographical Information Resource Development
and Application, Nanjing 210023, China
★These authors contributed equally to this work.

Correspondence: Shuping Zhao (shuping@njnu.edu.cn) and Zhuotong Nan (nanzt@njnu.edu.cn)

Received: 22 December 2025 – Discussion started: 12 March 2026
Revised: 24 May 2026 – Accepted: 27 May 2026 – Published: 3 June 2026

Abstract. Accurate, spatially explicit data on solum thickness (the vertical distance from the ground surface to the C horizon) is a critical missing link for Earth system modeling on the Qinghai–Tibet Plateau (QTP), where complex terrain and sparse subsurface observations limit the performance of traditional empirical mapping approaches. This study presents a new high-resolution (1 km) solum thickness dataset for the QTP developed using a revised geomorphic mass balance model. The model mechanistically couples climate-driven soil production with a multi-process erosion scheme that partitions sediment transport into hydraulic, aeolian, and gravitational components, while employing cluster-specific calibration and adjustable weighting coefficients to account for regional environmental heterogeneity. The resulting dataset reveals that solum thickness ranges from 0.39 to 2.04 m, with a plateau-wide mean of 0.89 m, exhibiting a pronounced decreasing gradient from the warm, humid southeast to the cold, arid northwest. Validation against 552 soil profile observations using 4-fold cross validation yields a Root Mean Square Error (RMSE) of 0.34 m and a Mean Relative Error of 0.78. This physically based approach achieves relative improvements of approximately 10 %–17 % in RMSE relative to existing national-scale products under shared observational constraints. Importantly, the model effectively reproduces realistic geomorphological patterns, such as sharp ridge-valley differentiation and elevation-dependent gradients that purely data-driven models often fail to capture in undersampled high-altitude regions. This dataset provides essential boundary conditions for hydrological, ecological, and cryospheric models, supporting more reliable assessments of the "Third Pole" under a changing climate. The dataset is available at figshare via https://doi.org/10.6084/m9.figshare.30925358 (Chen et al., 2025).

1 Introduction

As the central medium for material and energy exchange within terrestrial ecosystems, soil acts as a fundamental regulator of surface hydrological processes (Zhang et al., 2024a; Fu et al., 2011) and global biogeochemical cycles (Smith et al., 2015). Soil thickness, specifically defined in this study as the solum thickness (the vertical distance from the ground surface to the C horizon or parent material interface) (Xiao et al., 2023), governs hydraulic storage capacity, nutrient availability, and rooting depth (Berhe et al., 2018). While total regolith depth (depth to bedrock) includes the underlying saprolite, the solum represents the zone of active pedogenesis and high-frequency land-atmosphere interaction (Liu et al., 2013); thus its accurate spatial characterization is essential for terrestrial and Earth system models.

The Qinghai–Tibet Plateau (QTP) is a globally significant, ecologically sensitive region highly vulnerable to climate change (Zhang et al., 2022; Piao et al., 2019). The plateau's unique environment drives distinct soil formation mechanisms, including cryogenic constraints on organic matter decomposition (Chen et al., 2019), intense physical disintegration via frost action (Hall et al., 2002), and high rates of mass transfer from glacial and fluvial incision (Montgomery, 2002; Korup and Montgomery, 2008). These processes create pronounced spatial heterogeneity in soil development and preservation across the plateau. Consequently, a physically consistent solum thickness dataset is a fundamental prerequisite for modeling the region's hydrological regime (Fu et al., 2011), quantifying carbon stocks (Ke et al., 2025), and assessing the stability of its fragile alpine ecosystems (Berhe et al., 2018).

Despite this need, existing solum thickness or soil property datasets for the QTP, such as the China Dataset of Soil Properties for Land Surface Modelling, version 1 (hereafter the Shangguan map) (Shangguan et al., 2013) and the Chinese High-Resolution National Soil Information Grid Basic Attribute Dataset (hereafter the Liu map) (Liu et al., 2021), are primarily produced using empirical and machine learning models. These methods are fundamentally constrained by their reliance on extensive, representative field measurements (Gupta et al., 2024; Rosin et al., 2025). The extreme altitude and logistical challenges of the QTP result in a sparse and unevenly distributed network of soil profiles, which are largely clustered in accessible, flat areas, while hillslopes remain significantly undersampled. In such data-scarce and topographically complex environments, empirical models often suffer from limited generalization capability and fail to reproduce realistic geomorphic gradients.

To address this critical data gap, process-based models offer a robust alternative by mechanistically linking solum thickness to the physical interplay between weathering and surface erosion. By grounding simulations in geomorphic mass balance principles (Dietrich et al., 1995), this approach reduces the structural dependence on dense field measurements and is especially well-suited for estimating solum thickness on complex terrain where observations are absent. However, the application of such models on the tectonically active and permafrost-influenced QTP requires careful treatment of key assumptions (e.g., steady-state conditions) and propagation of uncertainties, which we address through quantitative sensitivity and ensemble analyses.

In this study, we present a revised mass balance model to simulate solum thickness across the QTP at a 1 km resolution. We enhance the traditional modeling framework by explicitly partitioning total erosion into hydraulic (water), aeolian (wind), and gravitational components, and by introducing adjustable weighting coefficients to handle regional environmental heterogeneity. We further conduct quantitative uncertainty and sensitivity assessments to evaluate physical plausibility and the impacts of model assumptions and input data uncertainties. We then validate the resulting high-resolution solum thickness dataset against field observations and compare it with existing national products. This work delivers a foundational, physically consistent solum thickness dataset essential for reducing uncertainty in regional hydrological and ecological simulations.

2 Materials and methods

2.1 Study area

The QTP, extending across 26°00′–39°47′ N and 73°19′–104°47′ E, represents the world's largest and highest physiographic unit, covering an area of approximately 2.5 million km^2 with an average elevation exceeding 4000 m a.s.l. (Fig. 1a). The region is characterized by a pronounced continental climate marked by significant spatial gradients; mean annual air temperature (MAT) decreases from approximately 10 °C in the humid southeast to −8 °C in the arid northwest (Yu et al., 2024). Mean annual precipitation (MAP) follows a similar trend, declining from over 1000 mm to less than 100 mm (Yang et al., 2010). Alpine grassland is the dominant land cover type (Wang et al., 2022), reflecting these severe climatic constraints. Permafrost is widespread, underlying approximately 67 % of the plateau's total area and acting as a primary regulator of cryogenic soil processes (Cao et al., 2023). In the core permafrost zones, intense freeze-thaw cycles, limited vegetation cover, and sparse observational data create particularly challenging conditions for both soil formation and large-scale modeling.

Soils on the QTP are typically pedogenically young and fragile, shaped by the combined effects of intense glacial erosion, pervasive freeze-thaw cycles, and wind erosion (Wei et al., 2025; Montgomery, 2002; Li et al., 2012). These active geomorphic processes drive rapid surface renewal, resulting in a complex soil-forming environment where diverse soil types exhibit high spatial heterogeneity (Fig. 1b).

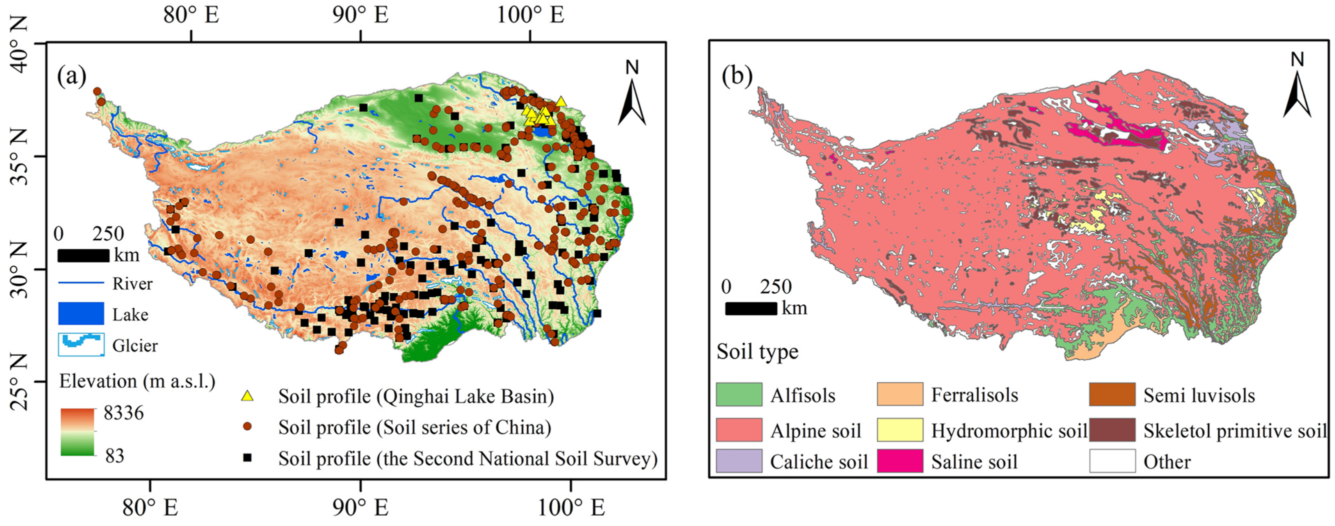


Figure 1. Overview of the study area. **(a)** Topography of the Qinghai–Tibet Plateau (QTP) showing the spatial distribution of the 552 soil profiles used for model calibration and validation. **(b)** Distribution of major genetic soil types across the QTP.

2.2 Data acquisition and preprocessing

Multiple datasets were compiled to drive the revised mass balance model, calibrate its parameters, and validate the outputs. A total of 552 soil profile observations were compiled across the QTP to serve as the ground truth for model calibration and performance assessment (Fig. 1a). These profiles were integrated from several primary sources, including the Second National Soil Survey (Pan and Shi, 2015), the Chinese Soil Series Chronicle (Central and Western Volumes) (Zhao and Li, 2020; Li et al., 2020; Yuan, 2020; Huang and Lu, 2020; Wu et al., 2020; Yang and Zhang, 2020), and a study of parent material age in the Qinghai Lake Basin (Zhang et al., 2024b). A critical methodological distinction is that these observations represent the vertical distance from the ground surface to the base of the pedogenic horizons, effectively measuring solum thickness. Because excavations in these national surveys typically terminate at the C horizon interface rather than the lithic contact (Shangguan et al., 2013), these data provide a consistent reference for the solum thickness simulated in this study.

Climate forcing was derived from high-resolution datasets for the period 2000–2015 to represent the modern steady-state climate. MAT data were sourced from the Dataset of 0.01° Surface Air Temperature over Tibetan Plateau (1979–2018) (Ding et al., 2022), while MAP data were extracted from the 1 km Monthly Precipitation Dataset for China (1901–2024) (Peng, 2025). Vegetation conditions, which regulate surface erosion, were parameterized using a 1 km MODIS(MOD13A2)-derived Normalized Difference Vegetation Index (NDVI) product (2001–2020) (Zhu, 2022) over the 2000–2015 interval. Topographic attributes, including slope, aspect, and curvature, were extracted from the 30 m NASADEM digital elevation model and subsequently resampled to a 1 km spatial grid using bilinear interpolation. Land use data for 2015 at a 300 m resolution (Xu, 2019) and soil parent material from the 1 : 2.5 million scale Quaternary Geological Map of the QTP (Li, 2020) were similarly standardized to a 1 km resolution grid.

The model's erosion framework is driven by spatially explicit estimates of hydraulic and aeolian mass transfer. The hydraulic erosion rate (E_{water}) was adopted from a 1 km resolution dataset (2002–2016) (Teng et al., 2018) generated using the Revised Universal Soil Loss Equation (RUSLE), while the aeolian erosion rate (E_{wind}) was sourced from a 1 km resolution map (1980–2015) (Teng et al., 2021) created via the Revised Wind Erosion Equation (RWEQ) model. To ensure geomorphic realism, we integrated the global depth-to-bedrock (DTB) dataset developed by Shangguan et al. (2017) as a critical structural constraint. This dataset, generated at a 1 km resolution using an ensemble of machine learning algorithms, defines the total available weathered mantle. However, it should be noted that uncertainties remain in the interior regions of the QTP due to highly sparse field observations and a heavy reliance on model-based extrapolation.

To evaluate the performance of our model, two existing national-scale datasets were utilized for comparison. The first dataset, the Shangguan map (Shangguan et al., 2013), was developed at a spatial resolution of 30 arcsec (∼ 1 km) using a polygon linkage method that integrated 8979 soil profiles from the Second National Soil Survey (1979–1985) with the 1 : 1 million Soil Map of China. Notably, the Shangguan map explicitly delineates non-soil map units (e.g., glaciers, rock debris, and water bodies), which appear as void spaces (null values) in the dataset. The second dataset, the Liu map (Liu et al., 2021), is a high-resolution (90 m) gridded dataset generated through a predictive mapping paradigm using a quantile regression forest algorithm. The model was trained on approximately 5000 representative soil profiles from the more recent National Soil Series Survey (2009–2019). While both datasets utilize extensive national survey data, their performance on the QTP is potentially limited by the lower density of profiles in high-altitude regions compared to eastern China. The observational dataset compiled for our study integrates the primary data sources used in both the Shangguan map and the Liu map, supplemented by additional regional survey data to maximize representativeness.

Finally, to constrain the simulation physically, non-soil map units were excluded by masking glaciers and water bodies using the Second Chinese Glacier Inventory (Guo et al., 2015b) and the China Lake Dataset (1960s–2020) (Zhang et al., 2019), with the aquatic mask specifically based on the 2020 lake distribution.

2.3 The revised mass balance model

To simulate the spatial heterogeneity of solum thickness across the QTP, we revised the classical mass balance framework to account for the region's specific geomorphic processes. Unlike traditional landscape evolution models that treat the subsurface as a single homogeneous layer, our framework employs a coupled two-interface system. This system distinguishes between the solum (h_s; m), defined as the genetically developed A and B horizons, and the underlying saprolite or C horizon (h_c; m). This distinction is critical for the QTP, where the physical and chemical coupling between surface erosion and the deep weathering front is mediated by the thickness of the solum.

2.3.1 Governing mass balance equations

The temporal evolution for the solum and saprolite thicknesses is governed by the conservation of mass. The change in thickness over time (t) for each layer is defined by the balance between material production from the underlying interface and material loss at the upper interface:

$$\frac{\partial h_\mathrm{s}}{\partial t} = \frac{\rho_\mathrm{c}}{\rho_\mathrm{s}} P_{\mathrm{c}\rightarrow\mathrm{s}} - E, \qquad (1)$$

$$\frac{\partial h_\mathrm{c}}{\partial t} = \frac{\rho_\mathrm{r}}{\rho_\mathrm{c}} P_{\mathrm{r}\rightarrow\mathrm{c}} - \frac{\rho_\mathrm{c}}{\rho_\mathrm{s}} P_{\mathrm{c}\rightarrow\mathrm{s}}, \qquad (2)$$

where $P_{\mathrm{c}\rightarrow\mathrm{s}}$ represents the production of solum from the C horizon ($\mathrm{m\,a^{-1}}$), $P_{\mathrm{r}\rightarrow\mathrm{c}}$ is the production of the C horizon from the unweathered bedrock ($\mathrm{m\,a^{-1}}$), and E represents the total physical erosion rate at the surface ($\mathrm{m\,a^{-1}}$). The bulk density terms for rock (ρ_r; $\mathrm{kg\,m^{-3}}$), saprolite (ρ_c; $\mathrm{kg\,m^{-3}}$), and solum (ρ_s; $\mathrm{kg\,m^{-3}}$) are integrated to account for the volumetric expansion occurring during pedogenesis. Chemical weathering terms are omitted because physical weathering and cryogenic processes are expected to dominate mass transfer across much of the QTP, particularly in high-elevation and permafrost regions.

2.3.2 Climate control on pedogenesis

The production of solum from the C horizon ($P_{\mathrm{c}\rightarrow\mathrm{s}}$) is modelled as an exponential decay function of the solum thickness (Pelletier and Rasmussen, 2009b), representing the insulating effect of the soil mantle on the weathering front. However, to account for the extreme climatic gradients of the QTP, we strictly define the potential production rate as a thermodynamic function rather than a static constant. The production function is formulated as:

$$P_{\mathrm{c}\rightarrow\mathrm{s}} = \Phi(T,P) P_{\mathrm{s}_0} e^{\frac{-h_\mathrm{s}}{h_0}}, \qquad (3)$$

where h_0 is the characteristic soil depth (m) and P_{s_0} is the potential production rate at the interface ($\mathrm{m\,a^{-1}}$). In this context, P_{s_0} represents the maximum rate at which the C horizon is converted to solum when $h_\mathrm{s} = 0$. The term $\Phi(T,P)$ represents the climate-dependency of the process, specifically integrating the effects of air temperature (T; °C) and precipitation (P; mm). Together, $\Phi(T,P)P_{\mathrm{s}_0}$ represents the maximum rate of solum formation under specific climatic conditions. We quantify this potential weathering rate using the Effective Energy and Mass Transfer (EEMT) model (Rasmussen et al., 2005; Rasmussen and Tabor, 2007; Pelletier et al., 2013), which calculates the energy available for mineral transformation based on local thermodynamics:

$$\Phi(T,P) P_{\mathrm{s}_0} = a e^{b\cdot \mathrm{EEMT}}, \qquad (4)$$

where a (lithology adjustment coefficient, $\mathrm{m\,a^{-1}}$) and b (climate response coefficient, $\mathrm{m^2\,a\,kJ^{-1}}$) are lithological dependent coefficients. EEMT ($\mathrm{kJ\,m^{-2}\,a^{-1}}$) is calculated from MAT (°C) and MAP (mm) (Rasmussen and Tabor, 2007):

$$\mathrm{EEMT} = 347134 \exp\left\{-\frac{1}{2}\left[\left(\frac{\mathrm{MAT}-21.5}{-10.1}\right)^2 + \left(\frac{\mathrm{MAP}-4412}{1704}\right)^2\right]\right\}. \qquad (5)$$

This formulation ensures that the model dynamically adjusts the weathering potential based on the varying thermal and moisture regimes across the plateau's vast area.

2.3.3 Multi-process erosion framework

Standard mass balance models typically approximate erosion solely as a diffusive process driven by gravity. However, the QTP is characterized by a multi-process erosional environment where wind and water play dominant roles in different regions (Wei et al., 2025). Other potential mass-loss factors, such as explicit freeze-thaw driven transport and human activities, were not modeled as separate erosion categories. Although freeze-thaw processes can also contribute to sediment transport in periglacial environments (Guo et al., 2015a), their primary role in the present framework is represented as a weathering mechanism that generate loose detrital material subsequently mobilized by hydraulic, aeolian, and gravitational transport. Furthermore, given the vast area and sparse population of the QTP, the impact of direct human activities on regional-scale erosion patterns is considered negligible compared to natural forcing (Sun et al., 2020) and was therefore excluded from the model.

To reflect the environmental heterogeneity of the region, we expanded the erosion term (E) into Eq. (1) into a weighted sum of three distinct components:

$$E=\alpha \frac{E_{\text{water}}}{10\rho_{\text{s}}}+\beta \frac{E_{\text{wind}}}{10\rho_{\text{s}}}+\gamma E_{\text{gravity}}, \tag{6}$$

where E_{water} and E_{wind} ($\mathrm{t\,ha^{-1}\,a^{-1}}$) are input erosion rates derived from the RUSLE method (Renard et al., 1994) and the RWEQ model (Fryrear et al., 2000), respectively. For the QTP, the two erosion inputs were sourced from regional datasets (Teng et al., 2018, 2021). Gravitational erosion (E_{gravity}; $\mathrm{m\,a^{-1}}$) representing downslope movement such as soil creep or freeze-thaw action, is modeled as a linear function of topographic curvature (Roering, 2008):

$$E_{\text{gravity}}=D\nabla^2 z, \tag{7}$$

where D is the soil diffusion coefficient ($\mathrm{m^2\,a^{-1}}$) and $\nabla^2 z$ is the topographic curvature over the elevation grid z.

The term $10\rho_{\text{s}}$ converts mass-based erosion rates into linear lowering rates ($\mathrm{m\,a^{-1}}$). The dimensionless weighting coefficients (α, β, γ) are introduced to account for the spatial heterogeneity of process dominance. These coefficients allow the model to be calibrated such that aeolian erosion determines thickness in the arid interior, while hydraulic erosion dominates in the humid monsoon zones.

2.3.4 Steady-state solution and boundary conditions

To solve for the solum thickness (h_{s}), we assume a geomorphic steady state where the production of soil is balanced by surface erosion over long timescales ($\partial h/\partial t=0$) (Phillips, 2010). While instantaneous fluxes may fluctuate (Gabet and Mudd, 2009), the steady-state framework remains a robust and widely applied approach for simulating large-scale spatial patterns over geomorphic timescales (Phillips, 2010). Substituting Eqs. (3) and (6) into Eq. (1) and rearranging the terms yields the equilibrium condition:

$$\frac{\rho_{\text{c}}}{\rho_{\text{s}}}\left(\Phi(T,P)P_{\text{s}0}e^{-h_{\text{s}}^*/h_0}\right)=\alpha \frac{E_{\text{water}}}{10\rho_{\text{s}}}+\beta \frac{E_{\text{wind}}}{10\rho_{\text{s}}}+\gamma E_{\text{gravity}}. \tag{8}$$

Solving this equation for the steady-state thickness (h_{s}^*), we obtain the final equation used for mapping:

$$h_{\text{s}}^*=-h_0\ln\left[\frac{\rho_{\text{s}}}{\rho_{\text{c}}}\frac{\left(\alpha E_{\text{water}}+\beta E_{\text{wind}}+\gamma E_{\text{gravity}}\right)}{\Phi(T,P)P_{\text{s}0}}\right]. \tag{9}$$

Within this derivation, the ratio of saprolite to solum bulk density ($\rho_{\text{c}}/\rho_{\text{s}}$) is assumed to be a constant value of 1.35, a standard parameter adopted from established geomorphological modeling studies (Pelletier and Rasmussen, 2009a) to represent volumetric expansion during pedogenesis. Equation (9) allows us to resolve the solum thickness. However, the second governing equation (Eq. 2) describing the saprolite evolution cannot be similarly solved due to a lack of constraints on the bedrock-to-saprolite production rate ($P_{\text{r}\rightarrow\text{c}}$). To resolve this, we utilize the spatially explicit DTB dataset from Shangguan et al. (2017) as a physical boundary condition. Since DTB represents the total regolith depth ($\text{DTB}=h_{\text{s}}+h_{\text{c}}$), the saprolite thickness is constrained by the relationship $h_{\text{c}}^*=\text{DTB}-h_{\text{s}}^*$. This ensures that the simulated solum thickness (h_{s}^*) is always physically bounded by the total available regolith, effectively closing the system of equations.

2.4 Environmental clustering for parameterization

Applying the mass balance model (Eq. 9) requires the specification of distinct kinetic and transport coefficients that govern soil evolution. Given the vast area and extreme environmental gradients of the QTP, assuming a single set of global parameters would oversimplify regional geomorphic distinctiveness and introduce systematic bias. To address this, we implemented a zonally adaptive parameterization strategy that partitions the plateau into homogeneous environmental clusters, allowing the model to explicitly account for the spatial non-stationarity of weathering and erosion processes. The process involved two main steps.

We first evaluated the predictive importance of 16 potential covariates derived from topographic, climatic, and ecological datasets (Table 1) (Yamashita et al., 2024; Wang et al., 2021). A Random Forest algorithm was employed to rank the predictive importance of these variables for solum thickness using the Mean Decrease Accuracy (MDA) metric. Variables with non-positive MDA values were deemed non-contributory and were excluded to reduce noise.

Second, based on the remaining key predictors, we partitioned the study area into eight environmentally distinct clusters using the k-prototypes (Huang, 1998) unsupervised clustering algorithm ($k=8$). This algorithm was specifically selected for its ability to handle mixed data types, effectively integrating both continuous and categorical variables. Consistent with the data processing in Sect. 2.2, non-soil areas (glaciers and water bodies) were excluded from the clustering process. This resulting zoning system provided the framework for calibrating the revised mass balance model independently for each cluster, thereby better capturing the region's complex environmental gradients.

2.5 Model calibration and performance assessment

Obtaining direct field measurement of the model parameters (Table 2) required by the mass balance model is prac-

Table 1. The initial pool of environmental variables.

Category	Environmental variable	Abbreviation
Terrain attributes	Elevation	Ele
	Slope	Slo
	Aspect	Asp
	Slope Length (Hickey et al., 1994)	SlL
	Slope Position	SlP
	General Curvature	GeC
	Plane Curvature	PlC
	Profile Curvature	PrC
	Terrain Ruggedness Index (Moreno et al., 2003)	TRI
Ecological Indicator	Normalized Difference Vegetation Index	NDVI
Hydrological Indicator	Topographic Wetness Index	TWI
Geological	Parent Material of Soil Formation	PMSF
Climatic Factors	Mean Annual Precipitation	MAP
	Mean Annual Temperature	MAT
Land Cover	Land Use Type	LUT
Erosion Potential	Sediment Transport Index	STI

tically infeasible across the vast and extreme environment of the QTP. We therefore implemented a parameter optimization approach using inverse modeling against the available soil profile datasets. The model parameters (Table 2) were calibrated independently for each of the eight environmental clusters using the Shuffled Complex Evolution (SCE-UA) algorithm (Duan et al., 1993). This method was selected for its robust performance in calibrating physical land surface models (Guo et al., 2013; Vrugt et al., 2003) as its complex shuffling mechanism is highly effective at avoiding convergence on local minima. To prevent equifinality and maintain physical consistency, the parameters were constrained within physically permissible ranges derived from literature (Table 2).

To rigorously evaluate the reliability of the simulated solum thickness, we implemented a two-tiered assessment strategy comprising internal cross-validation and external benchmarking against existing national datasets. Given the practical difficulty of collecting independent validation samples in the extreme environment of the QTP, we employed a k-fold cross-validation scheme (with $k = 4$) to maximize the utility of the 552 available soil profiles. Within each environmental cluster, the observational data were randomly partitioned into four equal subsets. The model was then iteratively trained using three of the subsets, with the remaining subset reserved for validation. This process was repeated four times to ensure every subset served as a validation set exactly once.

The accuracy of the solum thickness simulation was quantitatively characterized using three statistical metrics: Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and Mean Relative Error (MRE). These metrics characterize the deviation between simulated values ($\hat{m}(x_i)$) against measured values ($m(x_i)$):

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(\hat{m}(x_i) - m(x_i)\right)^2}, \tag{10}$$

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}\left|\hat{m}(x_i) - m(x_i)\right|, \tag{11}$$

$$\mathrm{MRE} = \frac{1}{n}\sum_{i=1}^{n}\frac{\left|\hat{m}(x_i) - m(x_i)\right|}{m(x_i)}. \tag{12}$$

Beyond internal consistency, we conducted a comparative performance analysis against two state-of-the-art regional products: the Shangguan map and the Liu map. Error metrics for all three datasets were calculated against the compiled 552 soil profiles. It is important to acknowledge that these 552 observations partially overlap with the source data used to generate the reference maps. Therefore, the performance comparison must be interpreted as a relative assessment under shared observational constraints, rather than a fully independent external validation. This approach still provides a rigorous benchmark of how well the process-based framework reproduces geomorphic reality compared to traditional polygon-linkage and machine-learning paradigms using the best available data for the region.

Table 2. Parameters of the revised mass balance model.

Parameter	Description	Value range	Unit	Basis/references
ρ_s	Soil bulk density	2400–2900	$\mathrm{kg\,m^{-3}}$	Rühlmann et al. (2006)
h_0	Characteristic soil depth	0–0.5	m	Pelletier and Rasmussen (2009b), Pelletier et al. (2013)
a	Lithology adjustment coefficient	0.5–100	$\mathrm{m\,a^{-1}}$	Pelletier and Rasmussen (2009b), Pelletier et al. (2013)
b	Climate response coefficient	0–0.1	$\mathrm{m^2\,a\,kJ^{-1}}$	Pelletier and Rasmussen (2009b), Pelletier et al. (2013)
α	Hydraulic erosion weight	0–1	–	Mathematically constrained
β	Aeolian erosion weight	0–1	–	Mathematically constrained
γ	Gravitational erosion weight	0–1	–	Mathematically constrained
D	Soil diffusion coefficient	0.00001–0.01	$\mathrm{m^2\,a^{-1}}$	Roering (2008)

3 Results

3.1 Environmental clusters and model calibration

The Random Forest-based importance analysis revealed significant disparities in the predictive power of the 16 potential environmental variables. As illustrated in Fig. 2a, elevation and NDVI emerged as the dominant factors controlling the spatial heterogeneity of the QTP environment. This reflects the fundamental thermodynamic constraints on pedogenesis across the QTP: elevation governs the thermal energy available for chemical weathering, while precipitation and vegetation regulate both the moisture necessary for mineral transformation and the surface stability against erosion. Conversely, local morphometric variables such as general curvature (GeC), the Sediment Transport Index (STI), land use type (LUT), the Terrain Ruggedness Index (TRI) and the Topographic Wetness Index (TWI) exhibited non-positive MDA values, indicating negligible contribution to the environmental zoning. Consequently, these five variables were excluded, retaining 11 key environmental variables for the clustering process.

Based on these variables, the QTP was partitioned into eight spatially distinct clusters (Fig. 2b). These clusters exhibit high internal homogeneity but significant inter-cluster heterogeneity, effectively capturing the plateau's diverse environmental regimes (Table 3). Cluster 1, located in the southeastern and eastern margins, represents the zone with the most favorable moisture and thermal regimes; it features the lowest mean elevation (2136 m a.s.l.) and the highest precipitation (832.41 mm) and air temperature (MAT: 13.09 °C), supporting the densest vegetation on the plateau (mean NDVI: 0.86) and rapid soil development. In sharp contrast, Cluster 2 is concentrated in the Qaidam Basin (northeastern QTP) and represents the plateau's arid interior, characterized by flat terrain, minimal precipitation (~ 222 mm), and desert vegetation. Clusters 3 and 4 serve as transitional zones in the eastern and southeastern QTP; they exhibit similar temperature and precipitation patterns and vegetation conditions to one another but are differentiated by elevation (3597 and 4354 m, respectively).

Table 3. Summary of the representative environmental characteristics for the eight clusters. MAT, MAP: mean annual air temperature and mean annual precipitation (2000–2015), respectively.

	Elevation (m a.s.l.)	MAT (°C)	MAP (mm)	NDVI (–)
Cluster 1	2136	13.09	832.41	0.86
Cluster 2	2938	4.58	222.49	0.21
Cluster 3	3597	3.05	610.26	0.64
Cluster 4	4354	2.23	809.56	0.66
Cluster 5	4230	−1.15	399.75	0.32
Cluster 6	4697	−1.27	449.22	0.34
Cluster 7	5036	−3.32	412.39	0.20
Cluster 8	5419	−4.01	505.06	0.16

The remaining groups, Clusters 5 through 8, constitute the high-altitude cold zones. Cluster 5 is predominantly located in the northeastern QTP, while Cluster 6 extends across the central and eastern-central portions of the plateau, including the Three Rivers Source Region. Both are defined by high elevations ($>$ 4200 m a.s.l.) and reduced climatic variability. Cluster 7 covers the central and western interior, the core permafrost zone, exhibiting low temperatures (MAT: −3.32 °C) and sparse vegetation. Finally, Cluster 8 occupies the most extreme high-elevation zone, such as the northern Himalayas where soil formation is kinetically limited by low temperatures and sparse biomass production.

The calibrated model parameters (Table 4) provide physical insights into these zonal differences. The lithology adjustment coefficient (a), which scales the maximum potential weathering rate, reveals a fundamental divergence between the plateau's interior core and its margins. In the central permafrost and high-mountain zones (Clusters 7 and 8), a values are extremely low ($< 1\,\mathrm{m\,a^{-1}}$), reflecting severe lithological constraints that inhibit soil production regardless of climatic inputs. In contrast, the surrounding regions (Clusters 1–4, 6) exhibit high potential weathering rates ($a > 80\,\mathrm{m\,a^{-1}}$). This suggests that in these active soil-forming zones, the calibrated model requires relatively high effective production rates to reproduce the observed solum thickness under these environmental conditions.

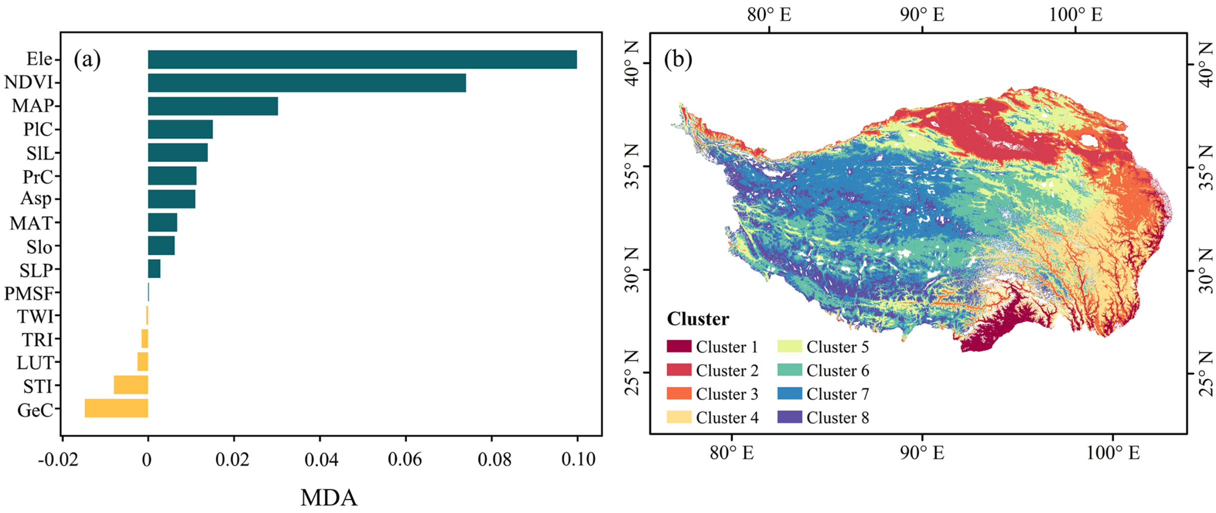


Figure 2. Environmental variable importance and environmental zoning of the QTP. **(a)** Importance ranking of the initial environmental variables based on the Random Forest algorithm. Higher Mean Decrease Accuracy (MDA) values indicate greater predictive importance. **(b)** Spatial distribution of the eight environmental clusters derived from the k-prototype clustering algorithm.

Table 4. Calibrated parameters of the revised mass balance model for the eight environmental clusters. n: the number of soil profile observations used for calibration in each cluster; RMSE: root mean square error resulting from the simulation. NA: cross-validation was not performed for Cluster 8 because sample size ($n = 4$) is insufficient for reliable results.

	ρ_s	h_0	a	b	α	β	γ	D	RMSE	n
Cluster 1	2764.99	0.038	86.03	0.00039	0.0012	0.00075	0.18	0.0016	0.38	35
Cluster 2	2895.88	0.048	98.34	1.42×10^{-06}	0.0024	0.0021	0.014	0.00041	0.34	131
Cluster 3	2855.11	0.043	99.7	4.02×10^{-05}	0.0014	0.0069	0.017	0.00011	0.38	136
Cluster 4	2895.74	0.026	99.69	0.0012	0.00043	0.0072	0.0035	6.23×10^{-05}	0.40	51
Cluster 5	2895.67	0.056	6.89	4.13×10^{-05}	0.15	0.0092	0.94	0.0096	0.34	71
Cluster 6	2899.59	0.042	99.18	3.84×10^{-05}	0.053	0.0012	0.052	6.95×10^{-05}	0.30	79
Cluster 7	2401.88	0.10	0.58	0.00014	0.99	0.96	0.99	0.0093	0.29	15
Cluster 8	2400.32	0.084	0.63	0.00069	0.99	0.0004	0.0056	0.00065	NA	4

The erosion weighting coefficients (α, β, γ) highlight spatial disparities in sediment transport. In the arid interior (Cluster 7), coefficients are near unity (≥ 0.96), indicating that empirical erosion estimates closely match the low weathering rates. Conversely, in the humid southeast (Cluster 1) and the source regions (Cluster 6), the hydraulic erosion weights are notably low ($\alpha < 0.06$), suggesting that vegetation cover significantly dampens actual sediment export relative to the potential erosion energy predicted by RUSLE.

3.2 Spatial distributions of simulated solum thickness

The simulation results from the revised mass balance model indicate that solum thickness across the QTP ranges from 0.39 to 2.04 m, with a spatially averaged mean of 0.89 m. Spatially, solum thickness generally decreases from the southeast to the northwest (Fig. 3). Regions with the greatest solum thickness are primarily concentrated in the southeastern and eastern margins of the plateau, particularly within the Yellow River Basin, where the average thickness reaches 1.16 m. This area corresponds mainly to Cluster 1, which features the most favorable moisture and thermal regimes on the plateau, fostering dense vegetation cover (NDVI > 0.6) and rapid soil development. Crucially, the high-resolution simulation resolves distinct geomorphic signatures that are often obscured in coarser empirical maps. The eastern drainage systems (Yangtze, Mekong, and Salween basins) display distinct dendritic patterns of solum thickness (Fig. 3). Thicker colluvial deposits (indicated by yellow to orange bands in Fig. 3) extend along the river valleys, contrasting sharply with the thinner solum on the adjacent bedrock-dominated mountain ridges. This ridge-valley differentiation provides geomorphic support for the model's gravitational erosion module, which simulates the transport of material from steep slopes to valley floors.

A similar depositional signature is evident in the Qaidam Basin (Cluster 2), which appears as a coherent "island" of moderately thick solum amidst the rugged terrain of the

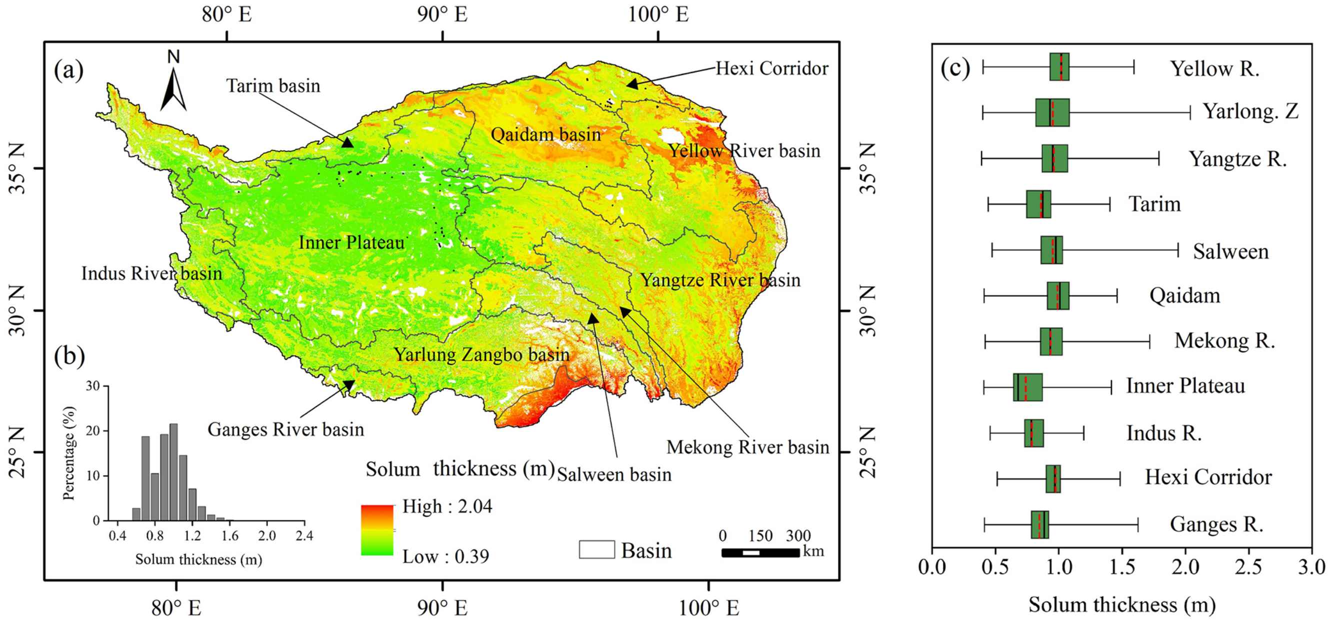


Figure 3. Spatial pattern and frequency distribution of simulated solum thickness across the QTP. **(a)** Spatial distribution of simulated solum thickness on the QTP. The delineated areas for the Mekong, Salween, and Yarlung Zangbo (Brahmaputra) rivers represent the catchment sections located within China. White patches denote masked non-soil areas (glaciers and water bodies) where solum thickness was not simulated. **(b)** The inset histogram shows the frequency distribution of simulated solum thickness across the QTP. **(c)** Boxplot showing the distribution of solum thickness across different basins of the QTP.

surrounding Qilian and Altun Mountains. This reflects the basin's nature as a stable depositional zone. Similarly, in the south, a continuous linear band of high solum thickness marks the Yarlung Zangbo River (Brahmaputra) valley. This depositional corridor stands in stark contrast to immediately adjacent high-altitude zones, the Himalayas to the south and the Gangdise Mountains to the north, which exhibit some of the lowest thickness values on the plateau due to intense glacial and periglacial erosion.

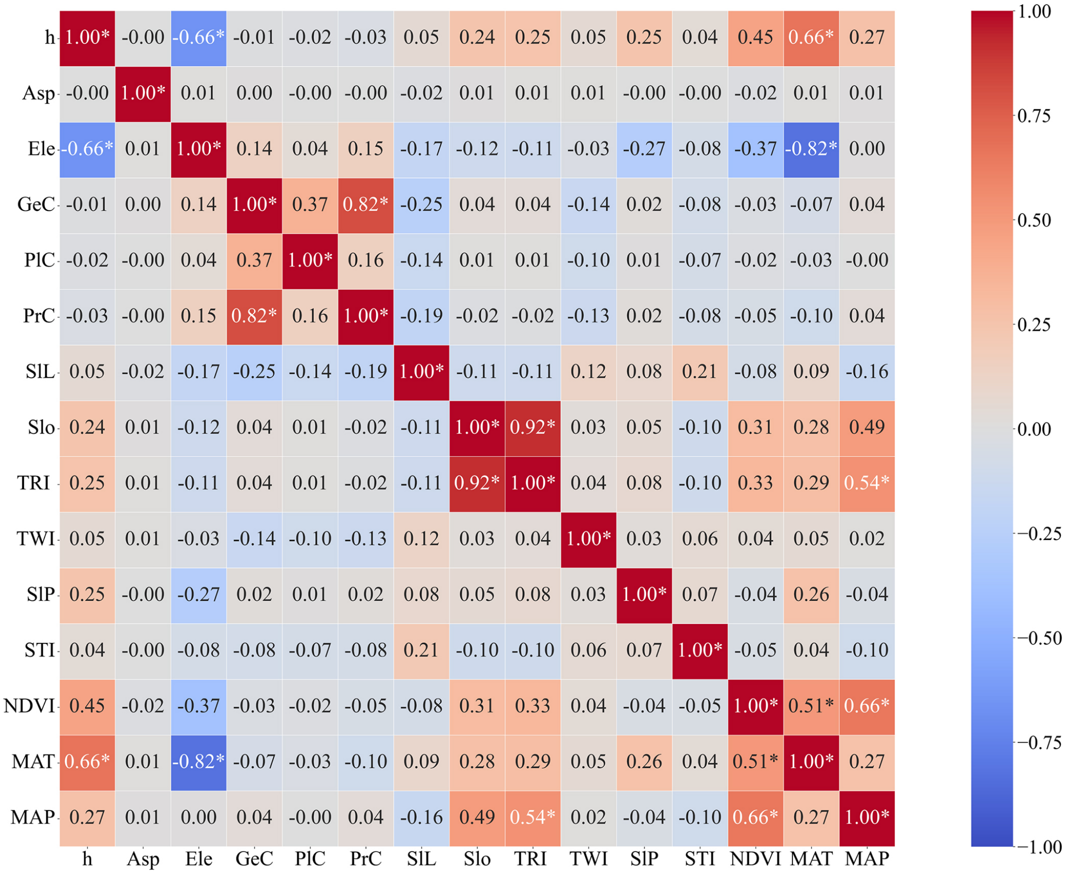
Conversely, the vast interior of the plateau (Clusters 7 and 8) is dominated by uniformly low solum thickness ($< 0.60\,\mathrm{m}$). This broad expanse reflects the core permafrost zone, where extreme cold and aridity kinetically inhibit pedogenesis, and the lack of vegetation leaves the surface vulnerable to deflation by strong westerly winds.

The frequency distribution of simulated solum thickness (Fig. 3b) shows that the model predominantly produces moderately developed soils (0.6–1.2 m), with thin and thick extremes largely partitioned by geomorphic regimes. Basin-level analysis (Fig. 3c and Table B1) further reveals clear differences among major catchments. The Yarlung Zangbo Basin in the southeastern QTP exhibits the greatest mean solum thickness and the largest internal variability (standard deviation), consistent with its complex terrain and strong topographic gradients. In contrast, the permafrost-dominated basins in the inner QTP (e.g., Inner Plateau) show the smallest mean solum thickness and the lowest variability, reflecting kinetically limited pedogenesis under cold and arid conditions.

3.3 Internal validation of environmental drivers

To verify that the model's underlying thermodynamic parameterizations were preserved in the final steady-state solutions, we conducted a Pearson correlation analysis between simulated solum thickness and key environmental variables (Fig. 4). Because the initial soil production function (P_0) is mathematically driven by the EEMT framework, the final results are structurally expected to correlate with both primary climatic and ecological inputs.

As anticipated, MAT ($r = 0.66$) and elevation ($r = -0.66$) emerge as the primary correlates of the regional solum thickness gradient. This strong coupling serves as an internal validation of our modeling framework. It confirms that the broad, regional-scale thermodynamic forcing is preserved through the steady-state mass balance equations. Vegetation exerts an equally important control. NDVI ($r = 0.45$) shows a clear positive correlation with simulated solum thickness, confirming that areas with denser vegetation cover experience reduced erosion rates and therefore support greater solum accumulation. This relationship validates the multi-process erosion framework, particularly the role of the hydraulic and aeolian weighting coefficients in dampening erosion under higher vegetation conditions.



	h	Asp	Ele	GeC	PlC	PrC	SlL	Slo	TRI	TWI	SlP	STI	NDVI	MAT	MAP
h	1.00*	-0.00	-0.66*	-0.01	-0.02	-0.03	0.05	0.24	0.25	0.05	0.25	0.04	0.45	0.66*	0.27
Asp	-0.00	1.00*	0.01	0.00	-0.00	-0.00	-0.02	0.01	0.01	0.01	-0.00	-0.00	-0.02	0.01	0.01
Ele	-0.66*	0.01	1.00*	0.14	0.04	0.15	-0.17	-0.12	-0.11	-0.03	-0.27	-0.08	-0.37	-0.82*	0.00
GeC	-0.01	0.00	0.14	1.00*	0.37	0.82*	-0.25	0.04	0.04	-0.14	0.02	-0.08	-0.03	-0.07	0.04
PlC	-0.02	-0.00	0.04	0.37	1.00*	0.16	-0.14	0.01	0.01	-0.10	0.01	-0.07	-0.02	-0.03	-0.00
PrC	-0.03	-0.00	0.15	0.82*	0.16	1.00*	-0.19	-0.02	-0.02	-0.13	0.02	-0.08	-0.05	-0.10	0.04
SlL	0.05	-0.02	-0.17	-0.25	-0.14	-0.19	1.00*	-0.11	-0.11	0.12	0.08	0.21	-0.08	0.09	-0.16
Slo	0.24	0.01	-0.12	0.04	0.01	-0.02	-0.11	1.00*	0.92*	0.03	0.05	-0.10	0.31	0.28	0.49
TRI	0.25	0.01	-0.11	0.04	0.01	-0.02	-0.11	0.92*	1.00*	0.04	0.08	-0.10	0.33	0.29	0.54*
TWI	0.05	0.01	-0.03	-0.14	-0.10	-0.13	0.12	0.03	0.04	1.00*	0.03	0.06	0.04	0.05	0.02
SlP	0.25	-0.00	-0.27	0.02	0.01	0.02	0.08	0.05	0.08	0.03	1.00*	0.07	-0.04	0.26	-0.04
STI	0.04	-0.00	-0.08	-0.08	-0.07	-0.08	0.21	-0.10	-0.10	0.06	0.07	1.00*	-0.05	0.04	-0.10
NDVI	0.45	-0.02	-0.37	-0.03	-0.02	-0.05	-0.08	0.31	0.33	0.04	-0.04	-0.05	1.00*	0.51*	0.66*
MAT	0.66*	0.01	-0.82*	-0.07	-0.03	-0.10	0.09	0.28	0.29	0.05	0.26	0.04	0.51*	1.00*	0.27
MAP	0.27	0.01	0.00	0.04	-0.00	0.04	-0.16	0.49	0.54*	0.02	-0.04	-0.10	0.66*	0.27	1.00*

Figure 4. Pearson correlation matrix between simulated solum thickness (h) and environmental variables. Red represents positive correlations and blue represents negative correlations. Asterisks denote statistically significant correlations ($p < 0.05$). Refer to Table 1 for the full definitions of the variable abbreviations.

Conversely, the near-zero plateau-scale correlation with general curvature ($r = -0.01$) should not be interpreted as evidence of negligible topographic influence. Rather, it indicates that curvature-related mass redistribution fluctuates rapidly around zero at the hillslope scale, and this high-frequency signal is statistically masked by the immense, low-frequency climatic gradients when aggregated across the entire plateau.

At the hillslope scale, however, topography-driven differentiation remains clearly visible. This is well illustrated in the transect analysis across the Hengduan Mountains (Fig. 5), a region of extreme relief and vertical zonation. Here, solum thickness varies systematically with terrain position: steep ridges act as sediment sources where active erosion restricts accumulation, whereas gentler slopes and valley bottoms function as depositional zones. Consequently, a distinct inverse relationship is observed (Fig. 5c), where solum thickness peaks in the warm, vegetated valley bottoms and diminishes rapidly towards the cold, barren ridges. These spatial patterns confirm that terrain attributes successfully drive local differentiation within the model, even though their signal is averaged out at the plateau scale.

3.4 Accuracy assessment and comparative analysis

3.4.1 Validation against field observations

The model performance was evaluated using a 4-fold cross-validation approach based on 552 available soil profile observations. The results indicate an overall RMSE of 0.34 m for the entire QTP. At the cluster level, RMSE values range from 0.29 to 0.40 m (Table 4). Cluster 7 (interior permafrost zone) exhibited the lowest RMSE (0.29 m). While this suggests high accuracy, it likely reflects the relatively low environmental heterogeneity of the central plateau combined with a limited sample size ($n = 15$), which warrants a cautious interpretation. Note that cross-validation was not performed for Cluster 8 due to an insufficient sample size of $n = 4$. Conversely, slight increases in RMSE were observed in the transitional zones (Clusters 1, 3, and 4) where complex topography and active monsoon climates introduce greater variability in the solum thickness. Overall, the consistency of RMSE values across diverse environmental zones suggests that the calibrated mass balance model effectively generalizes the relationship between soil formation processes and the chosen environmental covariates.

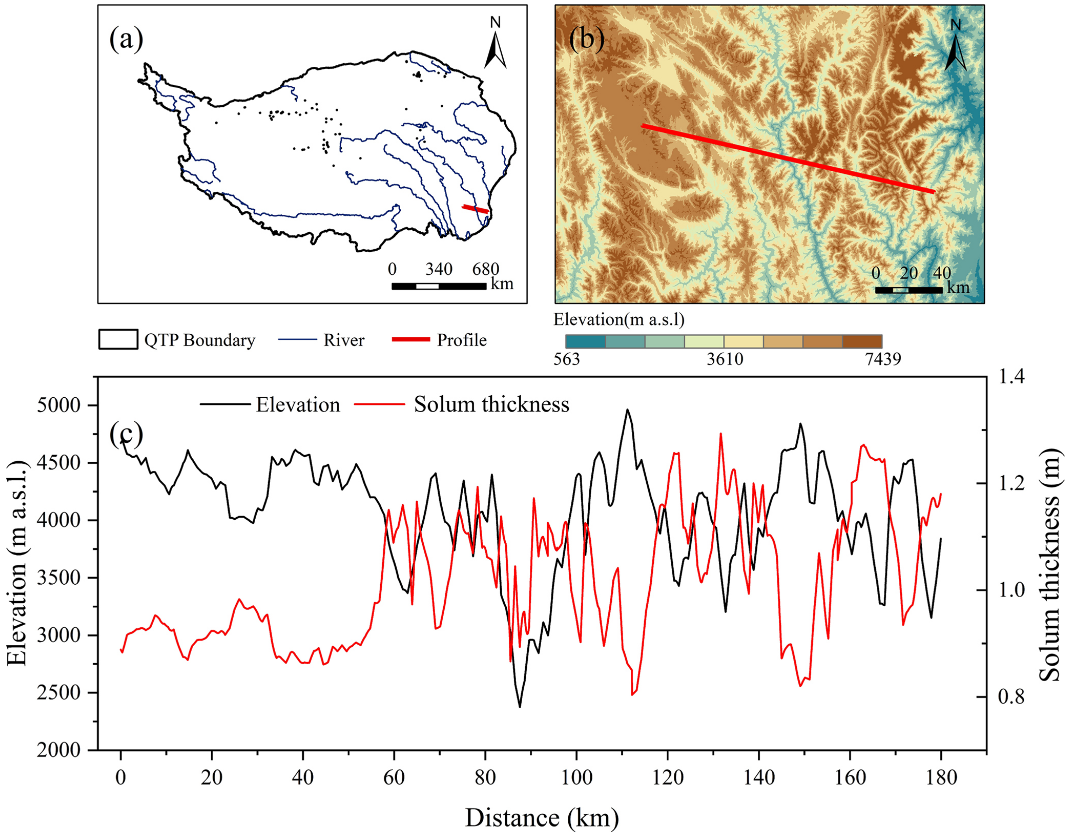


Figure 5. Topographic controls on solum thickness along a transect in the Hengduan Mountains. **(a)** Location of the study transect on the southeastern QTP. **(b)** Local topography of the transect area. **(c)** Cross-sectional profile illustrating the inverse relationship between elevation (black line) and simulated solum thickness (red line), where thickness peaks in valley floors and minimizes on ridge crests.

3.4.2 Comparison with existing solum thickness datasets

We compared our simulated solum thickness against two widely used regional datasets (the Shangguan map and the Liu map) (Table 5). In terms of magnitude, our model estimates a mean of 0.89 m (range: 0.39–2.04 m) across the QTP, which is notably higher than both the Shangguan map (mean: 0.65 m) and the Liu map (mean: 0.66 m).

To quantitatively assess accuracy, we calculated error metrics for all three datasets against the 552 compiled soil profile observations (Table 5). As noted in Secti. 2.5, these observations partially overlap with the source data used to generate the reference maps, the performance comparison should be viewed as a relative assessment under shared observational constraints. Despite this limitation, the revised mass balance model still yielded the lowest RMSE (0.34 m) and MRE (0.78), outperforming both the Shangguan map (RMSE = 0.45 m, MRE = 0.98) and the Liu map (RMSE = 0.42 m, MRE = 0.91). While our model exhibited a slightly higher MAE (0.13 m) compared to the Shangguan map (0.06 m), this difference is likely attributable to our model's ability to capture the full physical range of solum thickness, from negligible depths on ridges to thick deposits in the valleys (Figs. 3 and 5). The reference maps, by contrast, tend to predict spatially smoothed values, which artificially minimizes absolute error but does not explicitly resolve some geomorphic extremes (Fig. 6a and b). This results in a wider error distribution (higher MAE) but a better overall structural fit to the landscape (lower RMSE) for our physically based model.

Table 5. Statistical comparison of solum thickness and error metrics between this study and existing datasets. Min, Max and Mean refer to the solum thickness values (m) across the QTP. Error metrics (RMSE, MAE, MRE) were calculated against the 552 soil profiles. Note that these observations partially overlap with the source data used to generate the Shangguan and Liu maps; thus the comparison should be interpreted as a relative assessment under shared observational constraints.

	Min (m)	Max (m)	Mean (m)	RMSE (m)	MAE (m)	MRE
This study	0.39	2.04	0.89	0.34	0.13	0.78
Shangguan map	0.12	1.60	0.65	0.45	0.06	0.98
Liu map	0	2.35	0.66	0.42	0.11	0.91

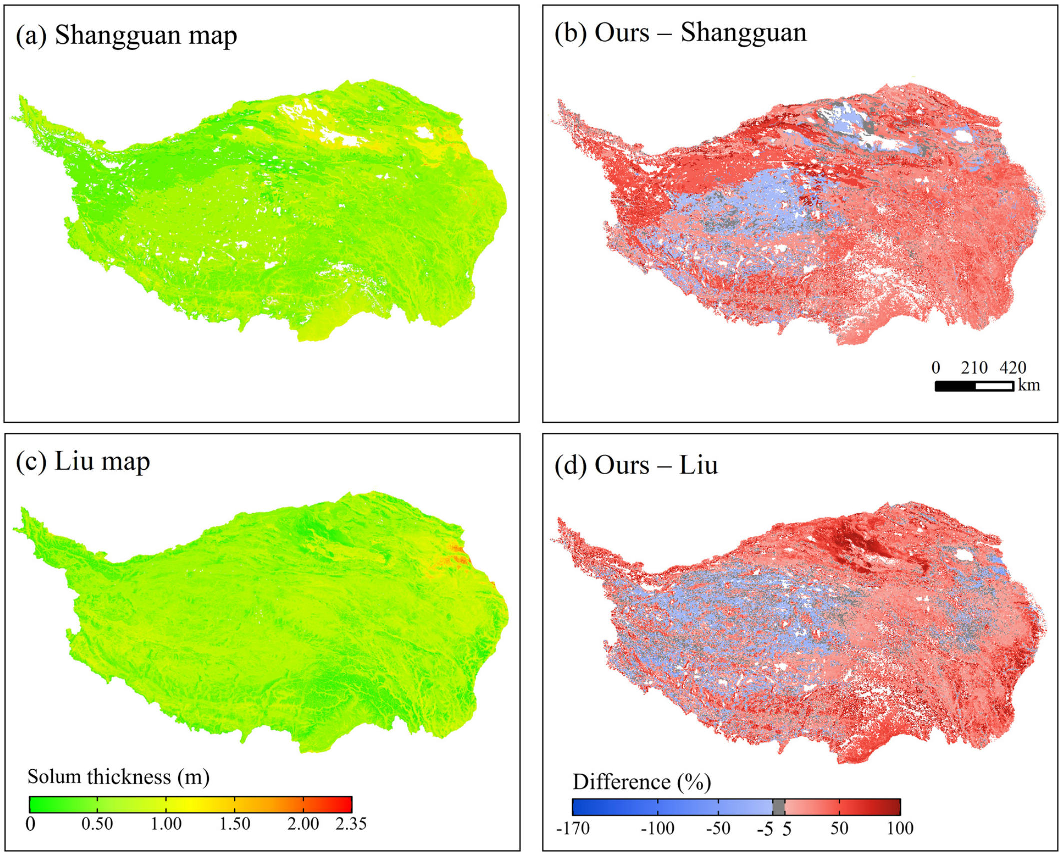


Figure 6. Comparison of solum thickness datasets across the QTP. Spatial distribution of solum thickness derived from **(a)** the Shangguan map and **(c)** the Liu map. Relative percentage difference between the simulated results of this study and **(b)** the Shangguan map and **(d)** he Liu map. Positive values indicate that our model predicts thicker solum than the reference datasets, while negative values indicate thinner solum. White patches represent null values (no data). In the Shangguan map, these correspond to defined non-soil units, while in our study, these correspond to masked glaciers and lakes.

Spatially, the difference maps (Fig. 6b and d) reveal where our process-based approach diverges from the previous maps. The grey masked regions indicate areas of strong agreement where the difference between our simulation and the reference datasets is within ±5 %. Comparison with the Shangguan map (Fig. 6b) reveals substantial negative differences in the Qaidam Basin and the inner QTP, indicating that the Shangguan map predicts significantly thicker solum in these arid, erosional landscapes than our model. This aligns with our earlier finding that the Shangguan map tends to overestimate solum thickness on steep slopes and ridges. Conversely, distinct positive differences are concentrated in the deep valleys of the southeastern Hengduan Mountains, suggesting our model identifies thicker colluvial deposits where the Shangguan map likely underestimates accumulation. The difference map for the Liu dataset (Fig. 6d) displays a higher degree of spatial fragmentation, reflecting the stochastic nature of the machine learning extrapolation used in the Liu map. Similar to the Shangguan comparison, negative differences are prevalent in the inner QTP, suggesting the Liu map also overestimates thickness in immature soil regions. In contrast, our model predicts higher thickness in the active depositional valleys of the Yarlung Zangbo and eastern river basins and lower thickness in the zone of inhibited pedogenesis in the interior of the QTP.

3.4.3 Performance in representative zones

Although the difference maps (Fig. 6) reveal substantial discrepancies between the datasets, numerical differences alone do not determine which map better represents the actual solum thickness. We conducted a detailed examination of two representative regions, the Qaidam Basin and the Qiangtang Plateau, where the divergence between our model and the reference datasets is pronounced. In the absence of dense independent validation data in these remote areas, we evaluated the results based on physical plausibility, specifically assessing whether the simulated spatial patterns align with the expected geomorphological processes of erosion and deposition.

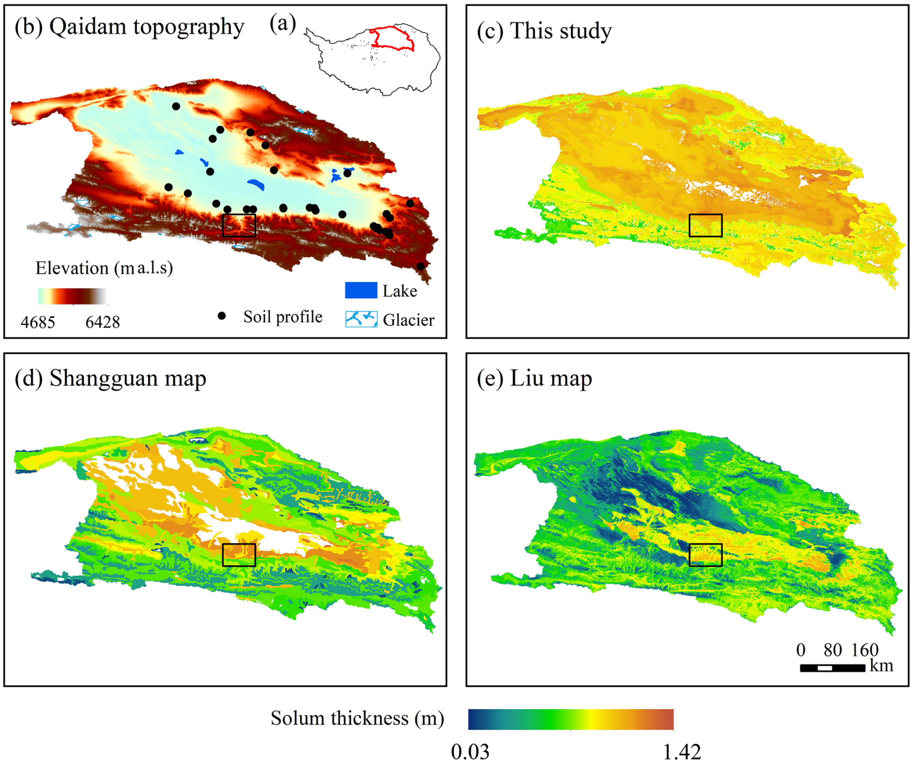


Figure 7. Comparison of solum thickness datasets in the Qaidam Basin. **(b)** Topography of the basin with an inset **(a)** showing its location on the QTP. Spatial distribution of solum thickness simulated by **(c)** this study, **(d)** the Shangguan map, and **(e)** the Liu map. The black box highlights the topographic gap in the southern mountain rim discussed in the text. White patches represent no data.

The Qaidam Basin (Fig. 7), located in the northeastern QTP, serves as a critical test case for depositional environments. Physically, this region is a stable, flat basin surrounded by high mountain ranges, where long-term sedimentation should result in a relatively continuous and uniform soil layer. Our model simulates a continuous layer of moderately thick solum (mean: 0.99 m) across the basin floor, with a low standard deviation (0.13 m) that is geomorphologically consistent with the basin's depositional nature. In contrast, the Liu map (Fig. 7e) displays anomalous patches of extreme high and low values (standard deviation: 0.24 m), resulting in a fragmented spatial pattern that contradicts the expected uniformity of a lacustrine basin floor. Similarly, the Shangguan map (Fig. 7d), while spatially continuous, predicts extremely thick solum (up to 1.36 m) extending onto the erosional ridges of the surrounding mountains, suggesting an overestimation of thickness in steep terrain.

Our results most accurately resolve the sharp physical boundary between the depositional basin (thick solum) and the erosional mountain ridges (thin solum). A clear example is visible in the southern mountain rim (boxed area in Fig. 7), where a distinct topographic gap breaks the continuity of the ridge. Our model (Fig. 7c) successfully captures this geomorphological feature, showing a corresponding break in solum thickness that mirrors the terrain. In contrast, the Shangguan map (Fig. 7d) and Liu map (Fig. 7e) smooth over this gap with a continuous block of thick solum, failing to reflect the topographic interruption.

The Qiangtang Plateau (Fig. 8) represents the harsh, high-altitude permafrost core of the QTP. This region is characterized by an average elevation of 4964 m and extremely sparse vegetation (mean NDVI: 0.18). In this environment, soil formation is physically controlled by the coupled effects of low temperature, sparse vegetation and topography. Statistically, the three datasets yield relatively comparable ranges, though our model predicts slightly thicker solum overall. Our simulation estimates a mean solum thickness of 0.74 m (range: 0.41–1.42 m), compared to a mean of 0.58 m (range: 0.12–1.22 m) for the Shangguan map and 0.67 m (range: 0.04–1.16 m) for the Liu map.

Despite these moderate numerical differences, the spatial distributions exhibit notable divergences. Our revised mass balance model captures a clear topographic dependency: solum thickness decreases systematically from the lower-elevation northeast to the high-altitude southwest (Fig. 8a and b). This gradient mirrors the spatial patterns of NDVI and elevation, with the lowest thickness values correctly located in the central high-altitude permafrost areas where weathering is kinetically limited. Conversely, the reference datasets exhibit patterns that are less consistent with expected geo-

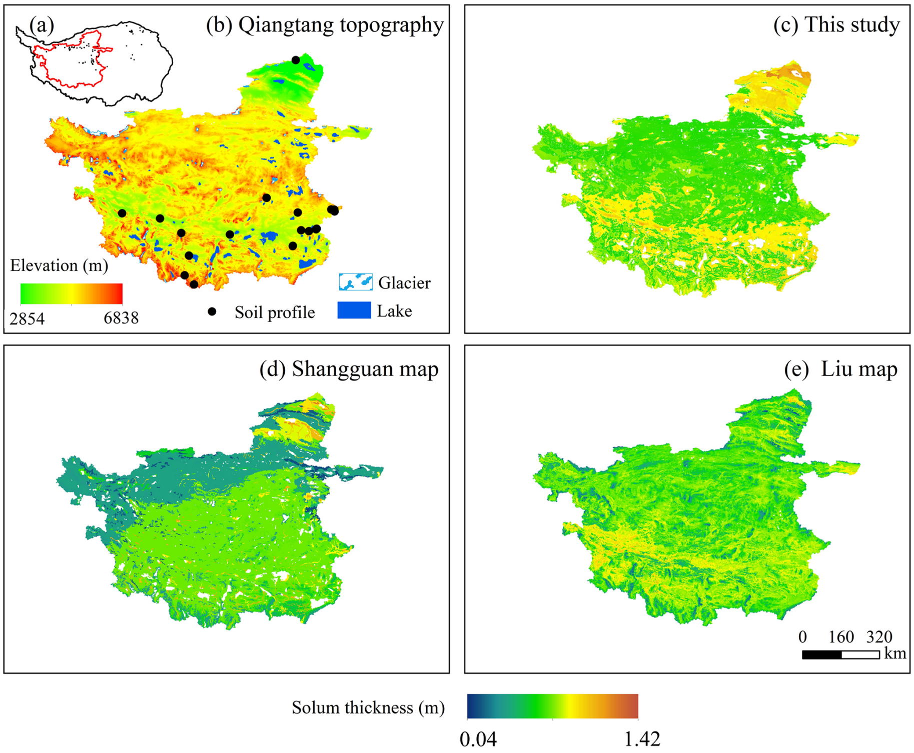


Figure 8. Comparison of solum thickness datasets in the Qiangtang Plateau. **(b)** Topography of the region with an inset **(a)** showing its location on the QTP. Spatial distribution of solum thickness simulated by **(c)** this study, **(d)** the Shangguan map, and **(e)** the Liu map. White patches represent no data.

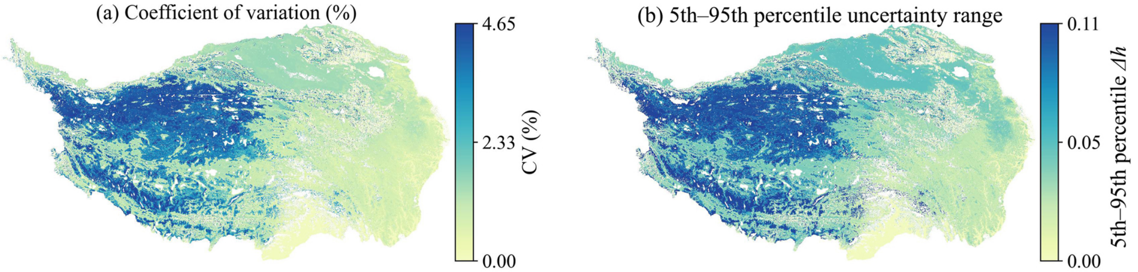
morphic controls. The Shangguan map (Fig. 8d) displays a distinct layered artifact, while the Liu map (Fig. 8e) shows a spatially uniform distribution that correlates poorly with the local topography. By effectively reproducing the elevation-dependent thinning of solum, our model demonstrates a superior capacity to represent solum thickness in extreme environments where empirical and machine learning models often fail due to the scarcity of training data (Jordan and Mitchell, 2015; Li and Heap, 2011).

4 Discussion

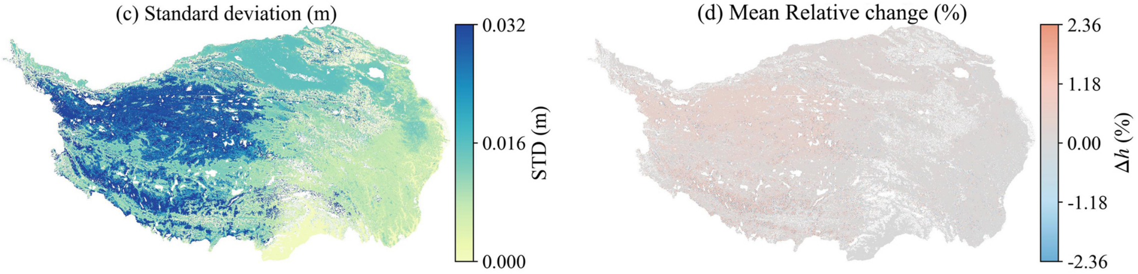
4.1 Advantages of process-based modeling in data-scarce regions

The accuracy of spatially explicit soil models on the QTP has long been constrained by the scarcity and bias of observational data. As evidenced by the distribution of the 552 soil profiles used in this study (Fig. 1a), sampling is predominantly clustered in accessible, lower-elevation regions, while high-altitude zones, which are characterized by cryogenic conditions, low atmospheric pressure, and rugged terrain, remain significantly undersampled. This spatial imbalance poses a fundamental challenge for empirical and machine learning approaches (e.g., the Shangguan and Liu maps), which operate on the assumption that the training data statistically represent the entire domain. When trained on biased datasets where hillslopes are underrepresented, these models tend to overfit the characteristics of flat terrain, leading to poor generalization in complex topography. This bias explains the artifacts observed in the reference datasets, such as the noisy fragmentation in the Qaidam Basin (Fig. 7e) or the artificial layered patterns in the Qiangtang Plateau (Fig. 8d).

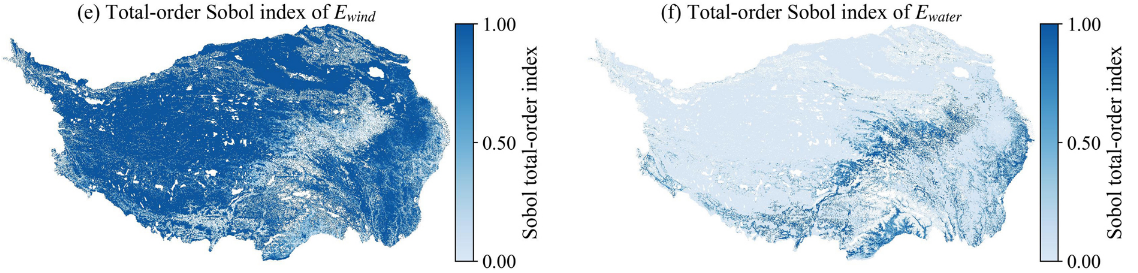
Despite utilizing the same biased observational dataset for calibration, the revised mass balance model offers an important advantage that stems from its reliance on physical mechanisms of pedogenesis rather than statistical correlation. By grounding the simulation in the mass balance equations (Eqs. 1 and 2), the model reduces dependence on dense field observations, and the structural constraint forces solum thickness to respond to topographic curvature and slope-dependent erosion, even under uneven training data distributions. Specifically, even if the model parameters are calibrated using data primarily from flat areas, the inclusion of the gravitational erosion term ensures that the model systematically removes solum from convex ridges and accumulates it in concave valleys, even in areas with no calibration data. This physically constrained behavior allows the model to infer plausible spatial patterns in undersampled regions and reproduce realistic geomorphic patterns (e.g., sharp basin-ridge boundaries in the Qaidam Basin and elevation-dependent gradients in the Qiangtang Plateau).


(a) Coefficient of variation (%)

(b) 5th–95th percentile uncertainty range


(c) Standard deviation (m)

(d) Mean Relative change (%)


(e) Total-order Sobol index of E_{wind}

(f) Total-order Sobol index of E_{water}

Figure 9. Spatial distribution of uncertainty and sensitivity in the simulated solum thickness across the QTP due to uncertainties in the empirical wind and water erosion inputs (RUSLE and RWEQ). **(a)** Coefficient of variation (CV, %), **(b)** 5th–95th percentile range. **(c)** Standard deviation of the 150 ensemble simulations. **(d)** Relative difference between ensemble mean and baseline. **(e, f)** Sobol total-order sensitivity indices for E_{wind} and E_{water}, respectively, quantifying the overall contribution of each input (including interactions) to output uncertainty. The legends for **(a)**–**(d)** show the 2nd–98th percentile range of values.

4.2 Uncertainty and sensitivity analysis

While the mass balance framework ensures physical consistency, the accuracy of the simulation remains partially contingent on the quality of the environmental forcing data and internal model parameters. A primary source of uncertainty stems from the empirical RUSLE and RWEQ erosion rates. These models were originally developed for low-altitude temperate zones and may have reduced applicability on the QTP due to its extreme altitude, widespread permafrost, frequent freeze-thaw cycles, and distinct surface conditions.

To quantify how uncertainty in these inputs propagates into simulated solum thickness, we independently perturbed E_{water} and E_{wind} within ±40 % of their baseline values using Latin Hypercube Sampling (LHS). This perturbation range was selected because previous studies have reported typical uncertainties of 30 %–60 % when these empirical models are applied outside their original calibration domains, particularly in cold, high-altitude, or permafrost-affected regions (Wei et al., 2025; Schürz et al., 2020). A total of 150 ensemble simulations were generated.

As shown in Fig. 9, the coefficient of variation (Fig. 9a), 5th–95th percentile uncertainty range (Fig. 9b), and standard deviation (Fig. 9c) consistently indicate higher uncertainty in the central and western permafrost core. The mean relative change between the ensemble mean and the baseline simulation (Fig. 9d) remains small over most of the plateau, suggesting that the baseline result is generally representative. Sobol total-order sensitivity indices (Fig. 9e and f), which measure the total contribution of an input (including its interactions with other inputs) to the variance of the simulated solum thickness, reveal that aeolian erosion exerts a stronger influence on output uncertainty across large areas, especially in arid and permafrost zones, while hydraulic erosion contributes more in the humid southeastern margins.

Despite these limitations, these RUSLE and RWEQ datasets remain the best available spatially explicit erosion products for the QTP. Previous validations have demonstrated their reliability, with reported coefficients of determination of 0.93 for aeolian erosion and 0.81 for hydraulic erosion against ^{137}Cs measurements (Teng et al., 2018, 2021).

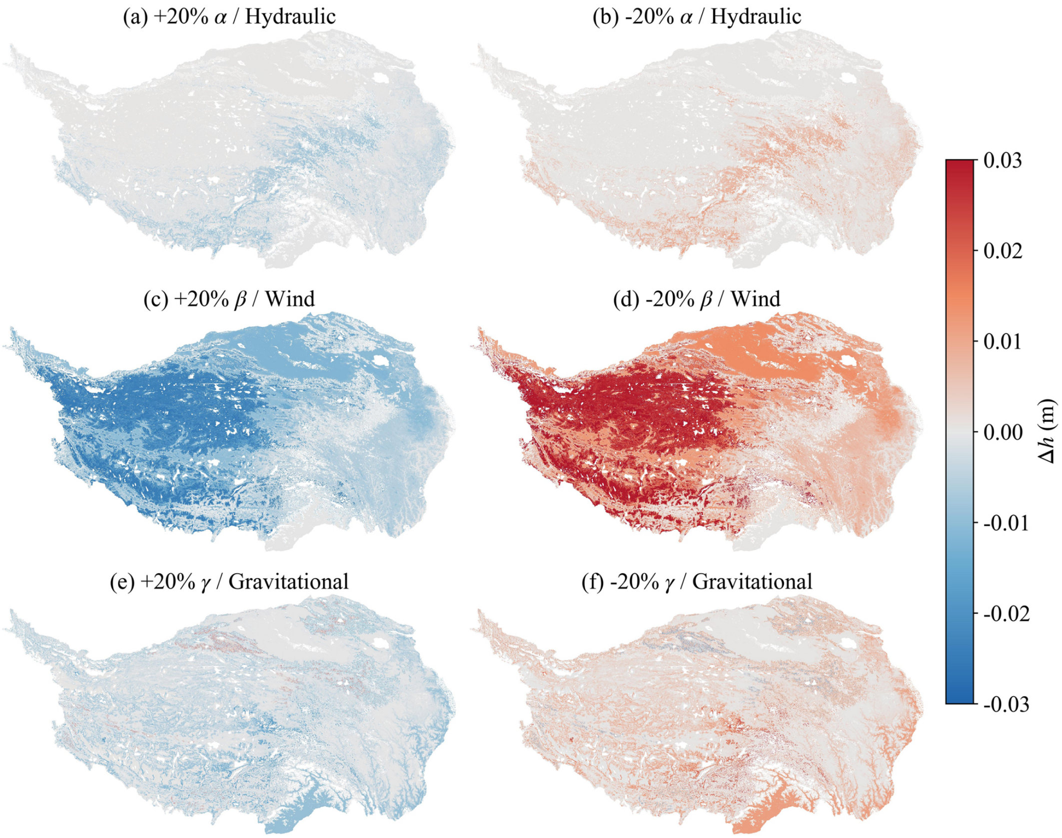


Figure 10. Spatial distribution of the sensitivity of simulated solum thickness to ±20 % perturbations in erosion weighting coefficients (α, β, and γ). **(a, b)** Sensitivity to ±20 % changes in the hydraulic erosion weight (α); **(c, d)** sensitivity to ±20 % changes in the aeolian erosion weight (β); **(e, f)** sensitivity to ±20 % changes in the gravitational erosion weight (γ). The legend indicates the 2 %–98 % range of values across the plateau. See Appendix C for cluster-level details.

Moreover, our model structure mitigates these input uncertainties through cluster-based calibration of the weighting coefficients (α, β, γ). By allowing these coefficients to adjust during calibration, the inverse modeling process functions as a physically constrained bias correction. Consequently, while the model relies on the spatial pattern of the erosion inputs, their magnitudes are rigorously scaled to better match local observations.

We further evaluated the sensitivity of simulated solum thickness to the erosion weighting coefficients through a One-At-A-Time (OAT) analysis by applying ±20 % perturbations to α, β, and γ based on the calibrated baseline map. The overall response is small across most of the plateau (Fig. 10). Detailed cluster-level elasticity coefficients and summary statistics are provided in Fig. C1 and Table C1. Plateau-wide elasticities, defined as the ratio of the relative change in simulated solum thickness to the relative change in the input parameter, remained consistently low ($\alpha \approx -0.01$, $\beta \approx -0.07$ to -0.08, $\gamma \approx -0.01$ to -0.02), indicating that the model is robust to reasonable variations in these parameters. A value close to zero indicates low sensitivity, while negative values indicate that solum thickness decreases as erosion increases. However, sensitivity exhibits spatial variation: gravitational weighting (γ) shows relatively higher influence in high-relief areas (e.g., Hengduan Mountains), while aeolian weighting (β) is more sensitive in the arid interior and permafrost zones (Clusters 7 and 8).

Overall, the combined uncertainty and sensitivity analyses demonstrate that the revised mass balance model produces stable solum thickness estimates under reasonable perturbations of both empirical inputs and internal parameters. The elevated uncertainty observed in the permafrost core primarily reflects limitations in the empirical erosion forcing rather than instability in the model itself.

4.3 Limitations and future directions

Although the revised mass balance model provides a physically consistent approach to mapping solum thickness across the QTP, its application relies on several important assumptions that warrant critical discussion.

The first assumption is the use of a steady-state framework. Our solution for solum thickness (Eq. 9) assumes that soil production is balanced by surface erosion over long timescales (Dietrich et al., 1995). On the tectonically active QTP (Ding et al., 2019; Pan et al., 2022), this assumption is an idealization. Soil systems are continuously adjusting in response to ongoing uplift (Zhao et al., 2023; Wei et al., 2025), climatic fluctuations, and permafrost degradation (Li et al., 2012). Localized non-linear processes, such as landslides, debris flows, and retrogressive thaw slumps (Luo et al., 2022), can also cause rapid changes in solum thickness that are not captured by the current linear diffusion-based formulation (Dixon et al., 2009; Dietrich et al., 1995; Roering et al., 1999).

To partially address the limitations of the steady-state assumption, we introduced adjustable weighting coefficients (α, β, γ). These coefficients should be interpreted as effective scaling factors rather than direct proxies for deviations from geomorphic equilibrium. As demonstrated in our uncertainty and sensitivity analyses (Sect. 4.2), they absorb multiple sources of uncertainty, including biases in the empirical erosion estimates (E_{water} and E_{wind}), unresolved transient dynamics, and scale mismatches between processes.

A second important limitation arises from the implicit representation of cryogenic processes. Potential soil production is estimated using the EEMT framework driven by MAT and MAP. While this provides a robust thermodynamics basis for regional patterns, the formulation does not explicitly account for freeze-thaw cycles and active layer dynamics, which are critical in permafrost regions. Similarly, the empirical RUSLE and RWEQ erosion inputs do not fully account for the complex interactions between seasonal freeze-thaw cycles, ice-rich permafrost, and sediment mobility. As quantified by our LHS analysis, these combined omissions introduce additional uncertainty that is most pronounced in the permafrost core. Further research should therefore prioritize the development of QTP-specific cryogenic parameterizations for both weathering and erosion.

The partitioning of the regolith into solum and saprolite ($h_{\mathrm{c}} = \mathrm{DTB} - h_{\mathrm{s}}$) is structurally constrained by the global depth-to-bedrock product of Shangguan et al. (2017). Although this dataset provides the best available estimate of the total weathered mantle, it is limited by the scarcity of deep borehole data across much of the plateau interior. To evaluate the influence of potential DTB uncertainty on our results, we applied a ±40 % perturbation to the DTB values. Even under a -40 % perturbation scenario, solum thickness is truncated by the DTB constraint in only 0.001 % of the total simulated area. This demonstrates that uncertainties in the DTB dataset have negligible impact on the reported solum thickness, which is derived independently through the erosion-production balance (Eq. 9).

Looking forward, coupling the current mass balance framework with landscape evolution models and time-dependent permafrost simulations would enable a more explicit representation of transient responses. Such integration would allow the simulation of how solum thickness may evolve under future climate scenarios, including potential deepening of the active layer and changes in erosion regimes associated with permafrost thaw.

5 Data availability

The resulting high-resolution (1 km) solum thickness dataset for the Qinghai–Tibet Plateau generated in this study, alongside supplementary datasets necessary for reproducing the analysis is freely available at figshare via https://doi.org/10.6084/m9.figshare.30925358 (Chen et al., 2025). The soil profile observations used for model calibration and validation were obtained from the National Earth System Science Data Center (https://doi.org/10.11666/00003.ver1.db) (Pan and Shi, 2015), the Chinese Soil Series Chronicle (Zhao and Li, 2020; Li et al., 2020; Yuan, 2020; Huang and Lu, 2020; Wu et al., 2020; Yang and Zhang, 2020) and regional studies (Zhang et al., 2024b). The existing national-scale solum thickness datasets used for comparative analysis are available from their respective publications (Shangguan et al., 2013; Liu et al., 2021). All other environmental covariates and input datasets used to drive the mass balance model are detailed in Appendix A.

6 Conclusions

This study presents a new 1 km resolution solum thickness dataset for the QTP derived from a revised geomorphic mass balance model. By explicitly partitioning erosion into hydraulic, aeolian, and gravitational components and implementing cluster-specific calibration, the model achieves a physically consistent representation of soil development across the plateau's diverse geomorphic and climatic regimes. This resulting dataset offers a reliable alternative to purely statistical products in high-altitude, data scare environments. The key findings are summarized as follows:

1. The simulated solum thickness on the QTP ranges from 0.39 to 2.04 m, with a plateau-wide mean of 0.89 m. The distribution exhibits a pronounced southeast-to-northwest decreasing gradient that aligns with the regional climatic and vegetation transitions from the warm, humid, vegetated margins to the cold, arid, barren interior.

2. Comprehensive validation against 552 soil profiles yielded an RMSE of 0.34 m and an MRE of 0.78. Compared to existing national-scale products under shared observational constraints, the process-based approach demonstrated improved performance, reducing RMSE by approximately 10 %–17 % relative to the Shangguan and Liu maps.

3. A critical advantage of this model is its ability to reproduce realistic geomorphic patterns, including sharp

ridge-valley differentiation and elevation-dependent gradients in data-sparse regions, features that empirical and machine learning products often fail to capture. Quantitative uncertainty and sensitivity analyses confirm that the model is robust to reasonable perturbations in key parameters and inputs, while also mapping spatially structured uncertainty that is highest in the permafrost core.

Appendix A: Sources of model input

The specific sources and access links for these input variables to the revised mass balance model are listed below:

Meteorological forcing. The mean annual air temperature (1979–2018) and precipitation (1901–2024) data were sourced from the National Tibetan Plateau Data Center. These datasets are available at https://doi.org/10.11888/Meteoro.tpdc.270339 (Ding et al., 2022) and https://doi.org/10.12041/geodata.192891852410344.ver1.db (Peng, 2025), respectively.

Erosion forcing. The hydraulic erosion data generated via the RULSE (Teng et al., 2018) and the aeolian erosion via RWEQ (Teng et al., 2021) are available upon request to the respective authors of those studies.

Vegetation and land use. Vegetation conditions were parameterized using the MODIS-derived NDVI product (2001–2020), accessible at https://doi.org/10.11888/Geogra.tpdc.270050 (Zhu, 2022). Land use data for the year 2015 was obtained from https://doi.org/10.11888/Geogra.tpdc.270198 (Xu, 2019).

Geological and topographic drivers. Soil parent material information was extracted from the Dataset of Soil Parent Material of China, available at https://doi.org/10.12072/ncdc.westdc.db3894.2023 (Li, 2020). Topographic attributes were derived from the NASADEM dataset, which is available through the National Aeronautics and Space Administration's Earthdata platform.

Hydrological masks. To mask non-soil areas, glacier extents were derived from the Second Chinese Glacier Inventory at https://doi.org/10.3972/glacier.001.2013.db (Guo et al., 2015b), and water bodies were identified using the China Lake Dataset at https://doi.org/10.11888/Hydro.tpdc.270302 (Zhang et al., 2019).

Appendix B: Statistics of solum thickness for QTP basins

Table B1. Statistics of solum thickness at the basin level across the QTP.

Basin	Mean (m)	Min (m)	Max (m)	SD (m)
Ganges River	0.85	0.41	1.63	0.15
Hexi Corridor	0.97	0.51	1.48	0.12
Indus River	0.79	0.46	1.20	0.12
Inner Plateau	0.74	0.41	1.42	0.12
Mekong River	0.94	0.42	1.72	0.13
Qaidam Basin	0.99	0.41	1.46	0.13
Salween Basin	0.95	0.47	1.94	0.17
Tarim Basin	0.86	0.44	1.40	0.16
Yangtze River	0.96	0.39	1.79	0.17
Yarlung Zangbo	0.95	0.40	2.04	0.22
Yellow River	1.02	0.40	1.59	0.16

Appendix C: OAT sensitivity analysis of erosion weighting coefficients (α, β, γ)

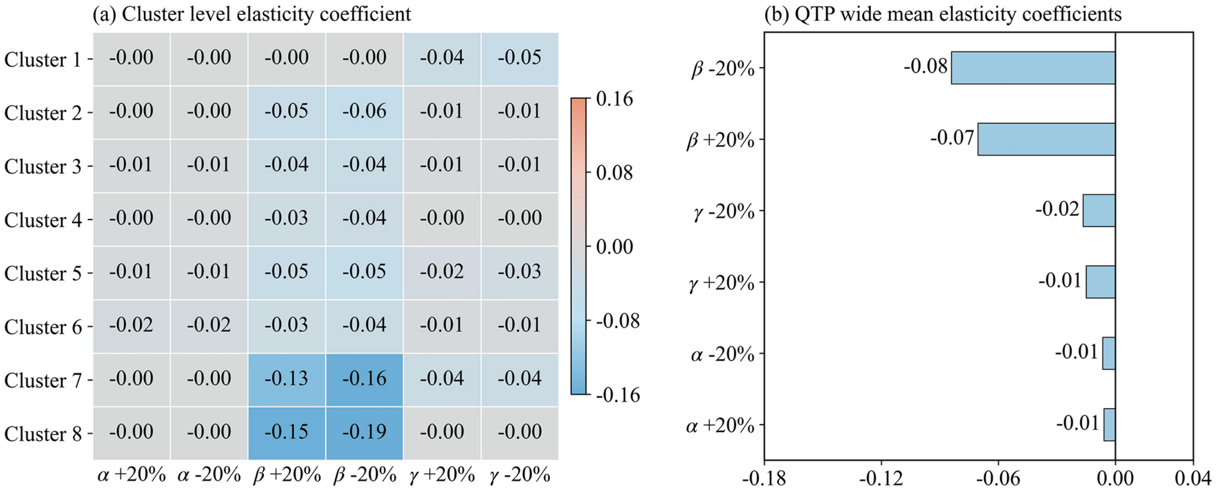


Figure C1. Elasticity of simulated solum thickness to ±20 % perturbation in erosion weighting coefficients. **(a)** Cluster-level elasticity coefficients showing the relative sensitivity of solum thickness to ±20 % perturbations in α (hydraulic), β (aeolian), and γ (gravitational) for each of the eight environmental clusters. **(b)** QTP-wide mean elasticity coefficients. Elasticity is negative because solum thickness responds inversely to erosion rate (Eq. 9).

Table C1. Summary statistics of the One-At-A-Time sensitivity analysis. Mean relative change (%), mean elasticity, and percentage of area with absolute change > 10 % are reported for each environmental cluster and for the entire QTP under ±20 % perturbations of the erosion weighting coefficients (α, β, γ).

Parameter	Scenario	Cluster	Mean relative change (%)	Mean elasticity	Area with > 10 % change (%)
α	−20 %	1	0.00	0.00	0
	−20 %	2	0.03	0.00	0
	−20 %	3	0.11	−0.01	0
	−20 %	4	0.08	0.00	0
	−20 %	5	0.19	−0.01	0.14
	−20 %	6	0.37	−0.02	0
	−20 %	7	0.03	0.00	0
	−20 %	8	0.07	0.00	0
	−20 %	Plateau wide	0.13	−0.01	0.01
α	+20 %	1	0.00	0.00	0
	+20 %	2	−0.03	0.00	0
	+20 %	3	−0.10	−0.01	0
	+20 %	4	−0.08	0.00	0
	+20 %	5	−0.19	−0.01	0.14
	+20%	6	−0.32	−0.02	0
	+20 %	7	−0.03	0.00	0
	+20 %	8	−0.06	0.00	0
	+20 %	Plateau wide	−0.12	−0.01	0.01
β	−20 %	1	0.00	0.00	0
	−20 %	2	1.16	−0.06	0
	−20 %	3	0.89	−0.04	0
	−20 %	4	0.71	−0.04	0
	−20 %	5	1.08	−0.05	0.90
	−20 %	6	0.74	−0.04	0
	−20 %	7	3.12	−0.16	0
	−20 %	8	3.73	−0.19	0
	−20 %	Plateau wide	1.68	−0.08	0.07
β	+20 %	1	0.00	0.00	0
	+20 %	2	−0.96	−0.05	0
	+20 %	3	−0.75	−0.04	0
	+20 %	4	−0.59	−0.03	0
	+20 %	5	−0.95	−0.05	0.61
	+20 %	6	−0.63	−0.03	0
	+20 %	7	−2.63	−0.13	0
	+20 %	8	−3.06	−0.15	0
	+20 %	Plateau wide	−1.41	−0.07	0.05
γ	−20 %	1	0.92	−0.05	0
	−20 %	2	0.14	−0.01	0
	−20 %	3	0.21	−0.01	0
	−20 %	4	0.04	0.00	0
	−20 %	5	0.54	−0.03	0.72
	−20 %	6	0.13	−0.01	0
	−20 %	7	0.77	−0.04	0
	−20 %	8	0.00	0.00	0
	−20 %	Plateau wide	0.33	−0.02	0.06
γ	+20 %	1	−0.76	−0.04	0
	+20 %	2	−0.12	−0.01	0
	+20 %	3	−0.19	−0.01	0
	+20 %	4	−0.04	0.00	0
	+20 %	5	−0.46	−0.02	0.68
	+20 %	6	−0.12	−0.01	0
	+20 %	7	−0.72	−0.04	0
	+20 %	8	0.00	0.00	0
	+20 %	Plateau wide	−0.30	−0.01	0.05

Author contributions. Conceptualization, ZN and SZ; Methodology, ZN and XD; Formal Analysis, LC, XD and KF; Writing – original draft, LC, ZN, XD, and SZ; Writing – review and editing, ZN, LC, SZ, FN and XD; Supervision, SZ and ZN; Funding acquisition, SZ and ZN.

Competing interests. The contact author has declared that none of the authors has any competing interests.

Acknowledgements. The authors are grateful to Jinyan Zhao from Beijing Normal University and Zhou Shi from Zhejiang University for providing the aeolian and hydraulic erosion datasets, respectively.

Financial support. This research has been supported by the National Natural Science Foundation of China (grant nos. 42571149 and 42171125).

Review statement. This paper was edited by Giulio G. R. Iovine and reviewed by two anonymous referees.

References

Berhe, A. A., Barnes, R. T., Six, J., and Marín-Spiotta, E.: Role of Soil Erosion in Biogeochemical Cycling of Essential Elements: Carbon, Nitrogen, and Phosphorus, Annu. Rev. Earth Planet. Sci., 46, 521–548, https://doi.org/10.1146/annurev-earth-082517-010018, 2018.

Cao, Z., Nan, Z., Hu, J., Chen, Y., and Zhang, Y.: A new 2010 permafrost distribution map over the Qinghai–Tibet Plateau based on subregion survey maps: A benchmark for regional permafrost modeling, Earth Syst. Sci. Data, 15, 3905–3930, https://doi.org/10.5194/essd-15-3905-2023, 2023.

Chen, L., Liu, L., Qin, S., Yang, G., Fang, K., Zhu, B., Kuzyakov, Y., Chen, P., Xu, Y., and Yang, Y.: Regulation of priming effect by soil organic matter stability over a broad geographic scale, Nat. Commun., 10, https://doi.org/10.1038/s41467-019-13119-z, 2019.

Chen, L., Ding, X., Zhao, S., and Nan, Z.: Dataset associated with the paper "A physically consistent soil thickness map of the Qinghai–Tibet Plateau derived from coupled erosion mechanisms", figshare [dataset], https://doi.org/10.6084/m9.figshare.30925358, 2025.

Dietrich, W. E., Reiss, R., Hsu, M. L., and Montgomery, D. R.: A process-based model for colluvial soil depth and shallow landsliding using digital elevation data, Hydrol. Process., 9, 383–400, https://doi.org/10.1002/hyp.3360090311, 1995.

Ding, J., Wang, T., Piao, S., Smith, P., Zhang, G., Yan, Z., Ren, S., Liu, D., Wang, S., Chen, S., Dai, F., He, J., Li, Y., Liu, Y., Mao, J., Arain, A., Tian, H., Shi, X., Yang, Y., Zeng, N., and Zhao, L.: The paleoclimatic footprint in the soil carbon stock of the Tibetan permafrost region, Nat. Commun., 10, 4195, https://doi.org/10.1038/s41467-019-12214-5, 2019.

Ding, L., Zhou, J., Wang, W., and Ma, J.: Dataset of 0.01° Surface Air Temperature over Tibetan Plateau (1979–2018), National Tibetan Plateau Data Center [data set], https://doi.org/10.11888/Meteoro.tpdc.270339, 2022.

Dixon, J. L., Heimsath, A. M., Kaste, J., and Amundson, R.: Climate-driven processes of hillslope weathering, Geology, 37, 975–978, https://doi.org/10.1130/G30045A.1, 2009.

Duan, Q. Y., Gupta, V. K., and Sorooshian, S.: Shuffled complex evolution approach for effective and efficient global minimization, J. Optimiz. Theory Appl., 76, 501–521, https://doi.org/10.1007/BF00939380, 1993.

Fryrear, D. W., Bilbro, J. D., Saleh, A., Schomberg, H., Stout, J. E., and Zobeck, T. M.: RWEQ: Improved wind erosion technology, J. Soil Water Conserv., 55, 183–189, https://doi.org/10.1080/00224561.2000.12436392, 2000.

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

Gabet, E. J. and Mudd, S. M.: A theoretical model coupling chemical weathering rates with denudation rates, Geology, 37, 151–154, https://doi.org/10.1130/G25270A.1, 2009.

Guo, B., Zhou, Y., Zhu, J., Liu, W., Wang, F., Wang, L., and Jiang, L.: An estimation method of soil freeze-thaw erosion in the Qinghai–Tibet Plateau, Nat. Hazards, 78, 1843–1857, https://doi.org/10.1007/s11069-015-1808-5, 2015a.

Guo, J., Zhou, J., Zou, Q., Liu, Y., and Song, L.: A Novel Multi-Objective Shuffled Complex Differential Evolution Algorithm with Application to Hydrological Model Parameter Optimization, Water Resour. Manage., 27, 2923–2946, https://doi.org/10.1007/s11269-013-0324-1, 2013.

Guo, W., Liu, S., Xu, J., Wu, L., Shangguan, D., Yao, X., Wei, J., Bao, W., Yu, P., and Liu, Q.: The second Chinese glacier inventory: data, methods and results, J. Glaciol., 61, 357–372, https://doi.org/10.3189/2015JoG14J209, 2015b.

Gupta, K., Satyam, N., and Segoni, S.: A comparative study of empirical and machine learning approaches for soil thickness mapping in the Joshimath region (India), Catena, 241, 108024, https://doi.org/10.1016/j.catena.2024.108024, 2024.

Hall, K., Thorn, C. E., Matsuoka, N., and Prick, A.: Weathering in cold regions: Some thoughts and perspective, Prog. Phys. Geogr., 26, 577–603, https://doi.org/10.1191/0309133302pp353ra, 2002.

Hickey, R., Smith, A., and Jankowski, P.: Slope length calculations from a DEM within ARC/INFO grid, Comput. Environ. Urban Syst., 18, 365–380, https://doi.org/10.1016/0198-9715(94)90017-5, 1994.

Huang, B. and Lu, S.: Soil Series of China-Yunnan, Science Press, Beijing, China, ISBN 978-7-5088-5704-6, 2020.

Huang, Z.: Extensions to the *k*-means algorithm for clustering large data sets with categorical values, Data Min. Knowl. Discov., 2, 283–304, https://doi.org/10.1023/A:1009769707641, 1998.

Jordan, M. I. and Mitchell, T. M.: Machine learning: Trends, perspectives, and prospects, Science, 349, 255–260, https://doi.org/10.1126/science.aaa8415, 2015.

Ke, L., Lei, N., Zhang, S., Yin, C., Lu, Y., Wang, L., Tan, Q., Zhao, Y., and Wang, Q.: Estimation of blue carbon stock in the Liaohe Estuary wetland based on soil thickness and multi-scenario modeling, Ecol. Indic., 171, 113201, https://doi.org/10.1016/j.ecolind.2025.113201, 2025.

Korup, O. and Montgomery, D. R.: Tibetan plateau river incision inhibited by glacial stabilization of the Tsangpo gorge, Nature, 455, 786–789, https://doi.org/10.1038/nature07322, 2008.

Li, D., Zhao, X., and Zhang, G.: Soil Series of China-Qinghai, Science Press, Beijing, China, ISBN 978-7-5088-5702-2, 2020.

Li, J. and Heap, A. D.: A review of comparative studies of spatial interpolation methods in environmental sciences: Performance and impact factors, Ecol. Inform., 6, 228–241, https://doi.org/10.1016/j.ecoinf.2010.12.003, 2011.

Li, X.: Comprehensive data set of Cryosphere in the Qinghai Tibet Plateau, National Cryosphere Desert Data Center [data set], https://doi.org/10.12072/ncdc.westdc.db3894.2023, 2020.

Li, X., Jin, R., Pan, X. D., Zhang, T. J., and Guo, J. W.: Changes in the near-surface soil freeze-thaw cycle on the Qinghai–Tibetan Plateau, Int. J. Appl. Earth Obs. Geoinf., 17, 33–42, https://doi.org/10.1016/j.jag.2011.12.002, 2012.

Liu, F., Wu, H., Zhao, Y., Li, D., Yang, J., Song, X., Shi, Z., Zhu, A., and Zhang, G.: Mapping high resolution national soil information grids of China, Sci. Bull., 67, 328–340, https://doi.org/10.1016/j.scib.2021.10.013, 2021.

Liu, J., Chen, X., Lin, H., Liu, H., and Song, H.: A simple geomorphic-based analytical model for predicting the spatial distribution of soil thickness in headwater hillslopes and catchments, Water Resour. Res., 49, 7733–7746, https://doi.org/10.1002/2013WR013834, 2013.

Luo, J., Niu, F., Lin, Z., Liu, M., Yin, G., and Gao, Z.: Inventory and Frequency of Retrogressive Thaw Slumps in Permafrost Region of the Qinghai–Tibet Plateau, Geophys. Res. Lett., 49, https://doi.org/10.1029/2022GL099829, 2022.

Montgomery, D. R.: Valley formation by fluvial and glacial erosion, Geology, 30, 1047–1050, https://doi.org/10.1130/0091-7613(2002)030<1047:Vfbfag>2.0.Co;2, 2002.

Moreno, M., Levachkine, S., Torres, M., and Quintero, R.: Geomorphometric Analysis of Raster Image Data to detect Terrain Ruggedness and Drainage Density, Prog. Pattern Recog. Speech Image Anal., 643–650, https://doi.org/10.1007/978-3-540-24586-5_79, 2003.

Pan, G., Wang, L., Yin, F., Geng, Q., Li, G., and Zhu, D.: Researches on geological-tectonic evolution of Tibetan Plateau: A review, recent advances, and directions in the future, Sediment. Geol. Tethyan Geol., 42, 151–175, https://doi.org/10.19826/j.cnki.1009-3850.2022.05004, 2022.

Pan, X. and Shi, J.: National Second Soil Census Typical Soil Profile Database (1980–1996), National Earth System Science Data Center [data set], https://doi.org/10.11666/00003.ver1.db, 2015.

Pelletier, J. D. and Rasmussen, C.: Quantifying the climatic and tectonic controls on hillslope steepness and erosion rate, Lithosphere, 1, 73–80, https://doi.org/10.1130/L3.1, 2009a.

Pelletier, J. D. and Rasmussen, C.: Geomorphically based predictive mapping of soil thickness in upland watersheds, Water Resour. Res., 45, https://doi.org/10.1029/2008WR007319, 2009b.

Pelletier, J. D., Barron-Gafford, G. A., Breshears, D. D., Brooks, P. D., Chorover, J., Durcik, M., Harman, C. J., Huxman, T. E., Lohse, K. A., Lybrand, R., Meixner, T., McIntosh, J. C., Papuga, S. A., Rasmussen, C., Schaap, M., Swetnam, T. L., and Troch, P. A.: Coevolution of nonlinear trends in vegetation, soils, and topography with elevation and slope aspect: A case study in the sky islands of southern Arizona, J. Geophys. Res.-Earth, 118, 741–758, https://doi.org/10.1002/jgrf.20046, 2013.

Peng, S.: 1-km monthly precipitation dataset for China (1901–2024), National Earth System Science Data Center [data set], https://doi.org/10.12041/geodata.192891852410344.ver1.db, 2025.

Phillips, J. D.: The convenient fiction of steady-state soil thickness, Geoderma, 156, 389–398, https://doi.org/10.1016/j.geoderma.2010.03.008, 2010.

Piao, S., Zhang, X., Wang, T., Liang, E., Wang, S., Zhu, J., and Niu, B.: Responses and feedback of the Tibetan Plateau's alpine ecosystem to climate change, Chinese Sci. Bull., 64, 2842–2855, https://doi.org/10.1360/TB-2019-0074, 2019.

Rasmussen, C. and Tabor, N. J.: Applying a quantitative pedogenic energy model across a range of environmental gradients, Soil Sci. Soc. Am. J., 71, 1719–1729, https://doi.org/10.2136/sssaj2007.0051, 2007.

Rasmussen, C., Southard, R. J., and Horwath, W. R.: Modeling energy inputs to predict pedogenic environments using regional environmental databases, Soil Sci. Soc. Am. J., 69, 1266–1274, https://doi.org/10.2136/sssaj2003.0283, 2005.

Renard, K. G., Foster, G. R., Yoder, D. C., and McCool, D. K.: RUSLE revisited: Status, questions, answers, and the future, J. Soil Water Conserv., 49, 213–220, https://doi.org/10.1080/00224561.1994.12456855, 1994.

Roering, J. J.: How well can hillslope evolution models "explain" topography? Simulating soil transport and production with high-resolution topographic data, Geol. Soc. Am. Bull., 120, 1248–1262, https://doi.org/10.1130/B26283.1, 2008.

Roering, J. J., Kirchner, J. W., and Dietrich, W. E.: Evidence for nonlinear, diffusive sediment transport on hillslopes and implications for landscape morphology, Water Resour. Res., 35, 853–870, https://doi.org/10.1029/1998WR900090, 1999.

Rosin, N. A., Mello, D. C. d., Bonfatti, B. R., Hartemink, A. E., Ferreira, T. O., Silvero, N. E. Q., Poppiel, R. R., Mendes, W. d. S., Veloso, G. V., Francelino, M. R., Alves, M. R., Falcioni, R., and Demattê, J. A. M.: Mapping soil thickness using a mechanistic model and machine learning approaches, Catena, 249, 108621, https://doi.org/10.1016/j.catena.2024.108621, 2025.

Rühlmann, J., Körschens, M., and Graefe, J.: A new approach to calculate the particle density of soils considering properties of the soil organic matter and the mineral matrix, Geoderma, 130, 272–283, https://doi.org/10.1016/j.geoderma.2005.01.024, 2006.

Schürz, C., Mehdi, B., Kiesel, J., Schulz, K., and Herrnegger, M.: A systematic assessment of uncertainties in large-scale soil loss estimation from different representations of USLE input factors – a case study for Kenya and Uganda, Hydrol. Earth Syst. Sci., 24, 4463–4489, https://doi.org/10.5194/hess-24-4463-2020, 2020.

Shangguan, W., Dai, Y. J., Liu, B. Y., Zhu, A. X., Duan, Q. Y., Wu, L. Z., Ji, D. Y., Ye, A. Z., Yuan, H., Zhang, Q., Chen,

D. D., Chen, M., Chu, J. T., Dou, Y. J., Guo, J. X., Li, H. Q., Li, J. J., Liang, L., Liang, X., Liu, H. P., Liu, S. Y., Miao, C. Y., and Zhang, Y. Z.: A China data set of soil properties for land surface modeling, J. Adv. Model. Earth Syst., 5, 212–224, https://doi.org/10.1002/jame.20026, 2013.

Shangguan, W., Hengl, T., Mendes de Jesus, J., Yuan, H., and Dai, Y.: Mapping the global depth to bedrock for land surface modeling, J. Adv. Model. Earth Syst., 9, 65–88, https://doi.org/10.1002/2016ms000686, 2017.

Smith, P., Cotrufo, M. F., Rumpel, C., Paustian, K., Kuikman, P. J., Elliott, J. A., McDowell, R., Griffiths, R. I., Asakawa, S., Bustamante, M., House, J. I., Sobocká, J., Harper, R., Pan, G., West, P. C., Gerber, J. S., Clark, J. M., Adhya, T., Scholes, R. J., and Scholes, M. C.: Biogeochemical cycles and biodiversity as key drivers of ecosystem services provided by soils, Soil, 1, 665–685, https://doi.org/10.5194/soil-1-665-2015, 2015.

Sun, Y., Liu, S., Shi, F., An, Y., Li, M., and Liu, Y.: Spatio-temporal variations and coupling of human activity intensity and ecosystem services based on the four-quadrant model on the Qinghai–Tibet Plateau, Sci. Total Environ., 743, https://doi.org/10.1016/j.scitotenv.2020.140721, 2020.

Teng, H., Liang, Z., Chen, S., Liu, Y., Viscarra Rossel, R. A., Chappell, A., Yu, W., and Shi, Z.: Current and future assessments of soil erosion by water on the Tibetan Plateau based on RUSLE and CMIP5 climate models, Sci. Total Environ., 635, 673–686, https://doi.org/10.1016/j.scitotenv.2018.04.146, 2018.

Teng, Y. M., Zhan, J. Y., Liu, W., Sun, Y. X., Agyemang, F. B., Liang, L., and Li, Z. H.: Spatiotemporal dynamics and drivers of wind erosion on the Qinghai–Tibet Plateau, China, Ecol. Indic., 123, https://doi.org/10.1016/j.ecolind.2021.107340, 2021.

Vrugt, J. A., Gupta, H. V., Bouten, W., and Sorooshian, S.: A Shuffled Complex Evolution Metropolis algorithm for optimization and uncertainty assessment of hydrologic model parameters, Water Resour. Res., 39, https://doi.org/10.1029/2002wr001642, 2003.

Wang, W., Zhao, Y., Zhang, T., Wang, R., Wei, Z., Sun, Q., and Wu, J.: Regional soil thickness mapping based on stratified sampling of optimally selected covariates, Geoderma, 400, https://doi.org/10.1016/j.geoderma.2021.115092, 2021.

Wang, Y., Lv, W., Xue, K., Wang, S., Zhang, L., Hu, R., Zeng, H., Xu, X., Li, Y., Jiang, L., Hao, Y., Du, J., Sun, J., Dorji, T., Piao, S., Wang, C., Luo, C., Zhang, Z., Chang, X., Zhang, M., Hu, Y., Wu, T., Wang, J., Li, B., Liu, P., Zhou, Y., Wang, A., Dong, S., Zhang, X., Gao, Q., Zhou, H., Shen, M., Wilkes, A., Miehe, G., Zhao, X., and Niu, H.: Grassland changes and adaptive management on the Qinghai–Tibetan Plateau, Nat. Rev. Earth Environ., 3, 668–683, https://doi.org/10.1038/s43017-022-00330-8, 2022.

Wei, P., Du, J., Bahadur, A., Zhang, H., Wang, S., Wu, T., and Chen, S.: Soil erosion and risk assessment on the Qinghai–Tibetan Plateau, Commun. Earth Environ., 6, 365, https://doi.org/10.1038/s43247-025-02355-5, 2025.

Wu, K., Wu, H., and Ju, B.: Soil Series of China-Xijiang, Science Press, Beijing, China, ISBN 978-7-5088-5717-3, 2020.

Xiao, T., Segoni, S., Liang, X., Yin, K., and Casagli, N.: Generating soil thickness maps by means of geomorphological-empirical approach and random forest algorithm in Wanzhou County, Three Gorges Reservoir, Geosci. Front., 14, hhttps://doi.org/10.1016/j.gsf.2022.101514, 2023.

Xu, E.: Land use of the Tibet Plateau in 2015 (Version 1.0), National Tibetan Plateau Data Center [data set], https://doi.org/10.11888/Geogra.tpdc.270198, 2019.

Yamashita, N., Ohnuki, Y., Iwahashi, J., and Imaya, A.: National-scale mapping of soil-thickness probability in hilly and mountainous areas of Japan using legacy and modern soil survey, Geoderma, 446, https://doi.org/10.1016/j.geoderma.2024.116896, 2024.

Yang, J. and Zhang, G.: Soil Series of China-Gansu, Science Press, Beijing, China, ISBN 978-7-5088-5891-3, 2020.

Yang, Y., Fang, J., Fay, P. A., Bell, J. E., and Ji, C.: Rain use efficiency across a precipitation gradient on the Tibetan Plateau, Geophys. Res. Lett., 37, https://doi.org/10.1029/2010GL043920, 2010.

Yu, Y., You, Q., Zhang, Y., Jin, Z., Kang, S., and Zhai, P.: Integrated warm-wet trends over the Tibetan Plateau in recent decades, J. Hydrol., 639, https://doi.org/10.1016/j.jhydrol.2024.131599, 2024.

Yuan, D.: Soil Series of China-Sichuan, Science Press, Beijing, China, ISBN 978-7-5088-5707-7, 2020.

Zhang, G., Yao, T., Chen, W., Zheng, G., Shum, C., Yang, K., Piao, S., Sheng, Y., Yi, S., and Li, J.: Regional differences of lake evolution across China during 1960s–2015 and its natural and anthropogenic causes, Remote Sens. Environ., 221, 386–404, https://doi.org/10.1016/j.rse.2018.11.038, 2019.

Zhang, J., Wang, S., Fu, Z. Y., Wang, F., Wang, K. L., and Chen, H. S.: Soil thickness influences the control effect of micro-topography on subsurface runoff generation in the karst hillslope critical zone, Catena, 239, https://doi.org/10.1016/j.catena.2024.107957, 2024a.

Zhang, Q., Yuan, R., Singh, V. P., Xu, C. Y., Fan, K., Shen, Z., Wang, G., and Zhao, J.: Dynamic vulnerability of ecological systems to climate changes across the Qinghai–Tibet Plateau, China, Ecol. Indic., 134, https://doi.org/10.1016/j.ecolind.2021.108483, 2022.

Zhang, S., E, C., Ji, X., Li, P., Peng, Q., Zhang, Z., and Zhang, Q.: Soil Genesis of Alluvial Parent Material in the Qinghai Lake Basin (NE Qinghai–Tibet Plateau) Revealed Using Optically Stimulated Luminescence Dating, Atmosphere, 15, 1066, https://doi.org/10.3390/atmos15091066, 2024b.

Zhao, W., Cheng, Y., Jian, J., Jiao, J., Cheng, C., Li, J., and Chen, T.: Water erosion changes on the Qinghai–Tibet Plateau and its response to climate variability and human activities during 1982–2015, Catena, 229, 107207, https://doi.org/10.1016/j.catena.2023.107207, 2023.

Zhao, Y. and Li, D.: Soil Series of China–Xizang, Science Press, Beijing, China, ISBN 978-7-5088-5890-6, 2020.

Zhu, J.: Dataset of Normalized Difference Vegetation Index over Tibetan Plateau From 2001 to 2020, National Tibetan Plateau Data Center [data set], https://doi.org/10.5067/MODIS/MOD13A2.006, 2022.