Cover letter

Dear Giulio,

Thank you for the opportunity to revise our manuscript entitled “A physically consistent soil thickness map of the Qinghai-Tibet Plateau derived from coupled erosion mechanisms” (ESSD-2025-809).

We sincerely thank both referees for their constructive and insightful comments, which have significantly improved the quality and clarity of the paper. We have carefully addressed all concerns raised. The major revisions include:

- Comprehensive quantitative uncertainty and sensitivity analyses (Latin Hypercube Sampling + Sobol + One-At-A-Time) with spatial mapping of uncertainty;
- Explicit acknowledgment of shared observational data limitations in validation;
- Theoretical clarification of the steady-state assumption and the role of adjustable weighting coefficients;
- Improved discussion of model limitations, particularly regarding empirical erosion inputs in permafrost regions;

A detailed point-by-point response to each reviewer’s comments is provided in the attached document. All changes in the manuscript are marked in tracked-changes mode.

Thank you again for your time and consideration.

Shuping Zhao & Zhuotong Nan
On behalf of all authors

Response to RC1

We sincerely thank Referee #1 for dedicating time and effort devoted to reviewing our manuscript. We are highly encouraged by your positive assessment of our approach and the potential value of the dataset. Your constructive comments have been invaluable in improving the quality and clarity of the paper, particularly regarding the theoretical limitations of our mathematical framework. All of your concerns have been carefully addressed in the revised manuscript.
Below, your original comments are provided in **blue,** our responses are given in **black,** and the corresponding revisions added to the main text are highlighted in **red.**

This manuscript presents a highly relevant and novel approach to mapping solum thickness on the Qinghai-Tibet Plateau (QTP). By departing from the purely empirical and machine-learning paradigms that currently dominate national-scale soil mapping, the authors implement a physically consistent mass balance model. This work addresses a critical bottleneck in Earth system modeling: the structural bias of observational data in remote, high-altitude regions. Overall, this is a worthy effort that offers a robust, physics-driven alternative to pervasive data-driven methods, offering a promising path forward for data-scarce environments. I believe this paper makes a strong contribution to the field. However, several methodological assumptions require deeper discussion before publication.

We sincerely thank the reviewer for this highly positive and encouraging assessment of our manuscript. We greatly appreciate your recognition of our efforts to transition from purely empirical paradigms to a physics-driven mass balance model to address the structural biases inherent in data-scarce, high-altitude regions like the QTP. We also agree with your assessment that our methodological assumptions warrant a much more rigorous discussion. In the revised manuscript, we have expanded our uncertainty analyses and theoretical discussions to address these important points, as detailed in our point-by-point responses below.

1) The core of the mathematical solution fundamentally relies on the assumption of a geomorphic steady state. The QTP, however, is a highly transient landscape. While the authors acknowledge this limitation in the discussion, relying on adjustable weighting coefficients to mathematically force a balance may inadvertently mask true mass imbalances across the region. I hope that the authors expand the discussion to

thoroughly explore how this mathematical workaround affects the final outputs and the interpretation of the underlying geomorphic mechanisms.

We thank the reviewer for this critical and insightful comment. To address this concern comprehensively, we have implemented a two-fold revision:

1 Quantitative sensitivity analysis: We conducted a One-At-A-Time sensitivity analysis (added as Section 4.2 Uncertainty and Sensitivity Analysis). Based on calibrated 1 km baseline map, we applied a relative perturbation of ±20% to the α, β, and γ coefficients. We then evaluated the sensitivity by calculating the absolute change (Δh), relative percentage change ($\%\Delta h$), and elasticity (the ratio of relative change in thickness to relative change in the coefficient). The results of this analysis map exactly where and to what extent the effective scaling coefficients influence the simulated thickness. We applied ±20% perturbations to the erosion weighting coefficients following the reviewer's suggestion to vary key parameters by ±20–50%. We consider this range appropriate because the calibrated coefficients in most clusters are already small; even under this perturbation the resulting elasticities remain very low, confirming model robustness.

As demonstrated in the figures and table below, the sensitivity is strongly partitioned by regional geomorphic regimes. For instance, in the extreme relief of the Hengduan Mountains (Cluster 1), the model exhibits the highest sensitivity to gravitational weights (γ elasticity = -0.05), while in the interior permafrost zones (cluster 7), sensitivity is dominated by aeolian weights (β elasticity = -0.19). However, the plateau wide elasticities remain robustly low (α = -0.01, β = -0.02, γ = -0.08). These results confirm a highly stable model structure that is relatively insensitive to reasonable parameter perturbations, while still capturing local geomorphic extremes. The low sensitivities partly result from the calibration process assigning small weighting coefficients in most clusters, which effectively down-weights the empirical RUSLE and RWEQ erosion rates.

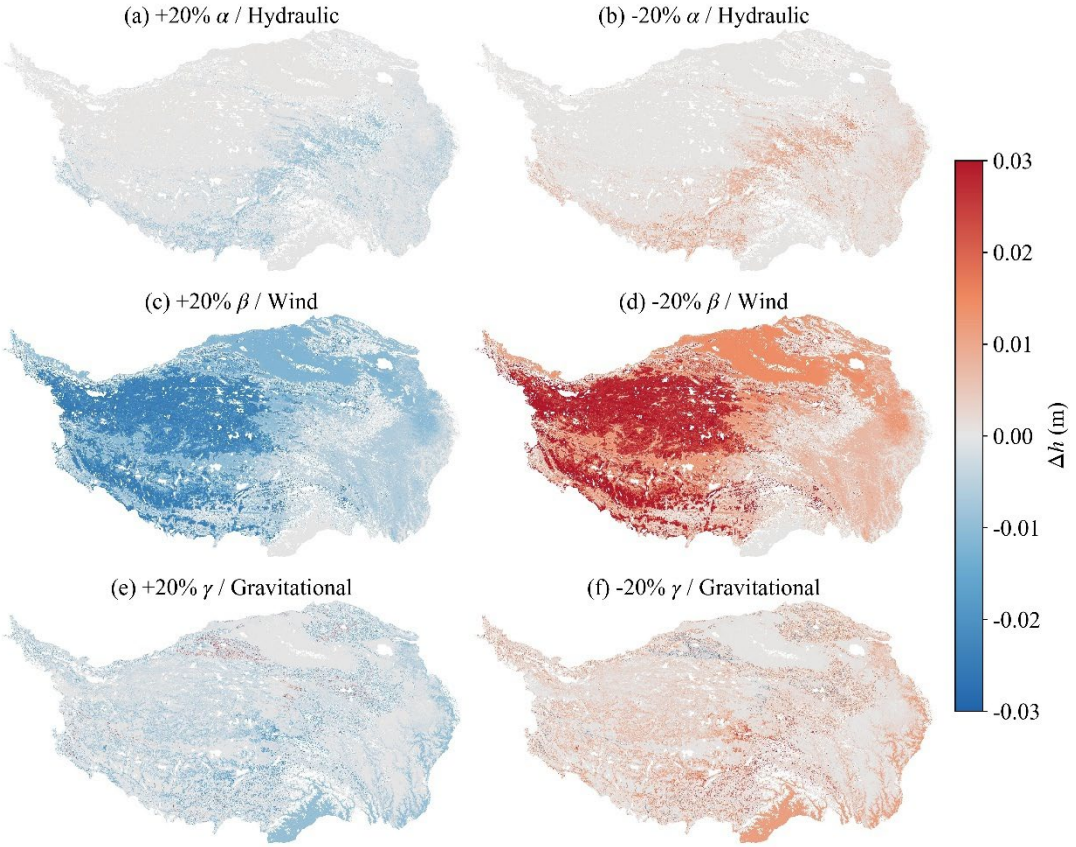


Figure 10. Spatial distribution of the sensitivity of simulated solum thickness to ±20% perturbations in erosion weighting coefficients (α, β, and γ). (a–b) Sensitivity to ±20% changes in the hydraulic erosion weight (α); (c–d) Sensitivity to ±20% changes in the aeolian erosion weight (β); (e–f) Sensitivity 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.

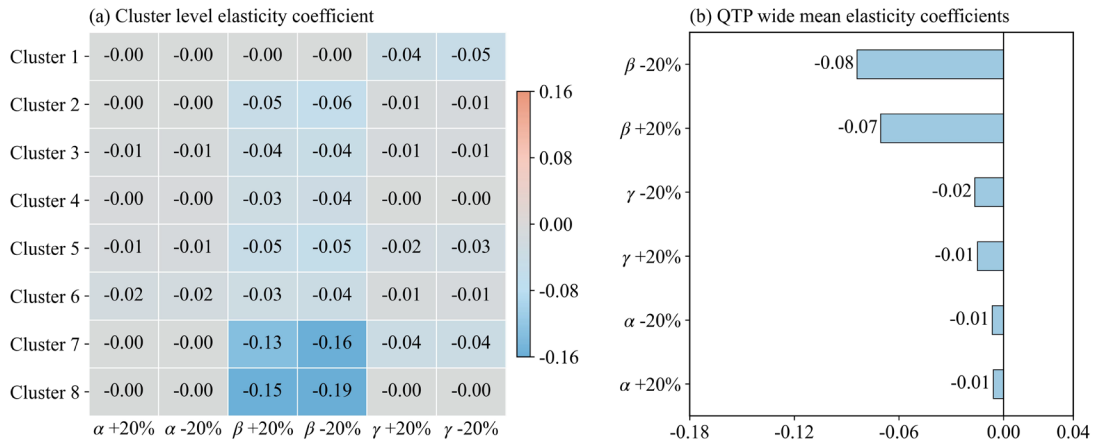


Figure C1 (Appendix C). Elasticity of simulated solum thickness to ±20% perturbation in erosion weighting coefficients. (a) Cluster-level elasticity coefficients showing the relative sensitivity of solum thickness to ±20% perturbations in α (hydraulic), β (aeolian), and γ (gravitational) for each of 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 (Appendix C). 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

2 Theoretical clarifications: We have revised Section 4.3 to explicitly state that the weighting coefficients act as “effective scaling factors”. They absorb multiple uncertainties, including empirical erosion biases, unresolved transient dynamics, and scale mismatches, rather than serving as direct proxies for true mass imbalances.

"...To partially address the limitations of such data, the steady-state assumption remains the most robust framework for capturing broad-scale spatial patterns. We addressed the potential bias of this assumption by introducing 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 (Section 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." (Lines 607-614 in the tracked revised manuscript)

We acknowledge that this approach provides a quasi-steady-state representation rather than a full transient reconstruction. Future work coupling our framework with landscape evolution models will be needed to explicitly simulate time-dependent soil thickness changes under ongoing uplift and permafrost degradation.

2) The model uses the Effective Energy and Mass Transfer framework to estimate potential weathering rates, which relies strictly on Mean Annual Temperature (MAT) and Mean Annual Precipitation (MAP). Consequently, the model largely overlooks the specific mechanics of freeze-thaw cycles and active layer dynamics. While I recognize the immense difficulty of implementing these complex cryogenic components into the current model structure, the authors must provide a more thorough and critical discussion of this limitation.

We thank the reviewer for highlighting this important point. We agree that the EEMT framework, being derived primarily from MAT and MAP, does not explicitly model the specific thermodynamics of freeze-thaw cycles and active layer dynamics. However, some of their integrated effects may be indirectly represented through the calibrated transport coefficients.

Physically, freeze-thaw primarily acts as a weathering mechanism that generates loose detrital material through frost shattering. The subsequent mass loss that determines solum thickness is executed by three modelled transport processes (gravitational, wind, and water erosion). In our framework, the cluster-calibrated weighting coefficients (α, β, γ) scale these transport fluxes to match the observed local solum thickness. Consequently, these coefficients implicitly account for the net effects of cryogenic weathering by adjusting the empirical erosion terms, even though an explicit freeze-thaw production term is not included. We recognize that this implicit treatment is a limitation of the current model. An explicit cryogenic production term would be a valuable future improvement.

To rigorously address this concern, we have made two major revisions:

1 Theoretical clarification: We expanded the discussion to explicitly explain the mechanism described above.

"... 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." (Lines 615-622 in tracked revised manuscript)

2 Global uncertainty and sensitivity analysis: Because this implicit approach relies on baseline empirical erosion inputs derived from RUSLE and RWEQ, which were not specifically designed for permafrost environments, we performed a comprehensive global uncertainty and sensitivity analysis by adding it in Section 4.2 Uncertainty and Sensitivity Analysis. For the uncertainty analysis, the empirical water and wind erosion inputs were independently perturbed within ±40% of their baseline values using Latin Hypercube Sampling, generating 150 ensemble simulations. This ±40% range is justified by the known limitations of these temperate-derived empirical models when extrapolated to high-altitude permafrost environments, where uncertainties are typically reported in the 30–60% range.

Based on these simulations, we quantified the spatial uncertainty in simulated solum thickness using the coefficient of variation (CV), standard deviation, 5th–95th percentile range, and the relative difference between the ensemble mean and the baseline simulation (Fig. 9 a-d). For the sensitivity analysis, to identify the dominant sources of uncertainty, we performed a separate Sobol total-order sensitivity analysis. The required samples were generated using the Saltelli sampling scheme, with E_{water} and E_{wind} sampled within ±40% of their baseline values. A total of 512 model evaluations were used to calculate the Sobol total-order indices. These indices quantify the overall contributions of E_{water} and E_{wind} to output uncertainty, including both their direct effects and interaction effects (Fig. 9 e, f).

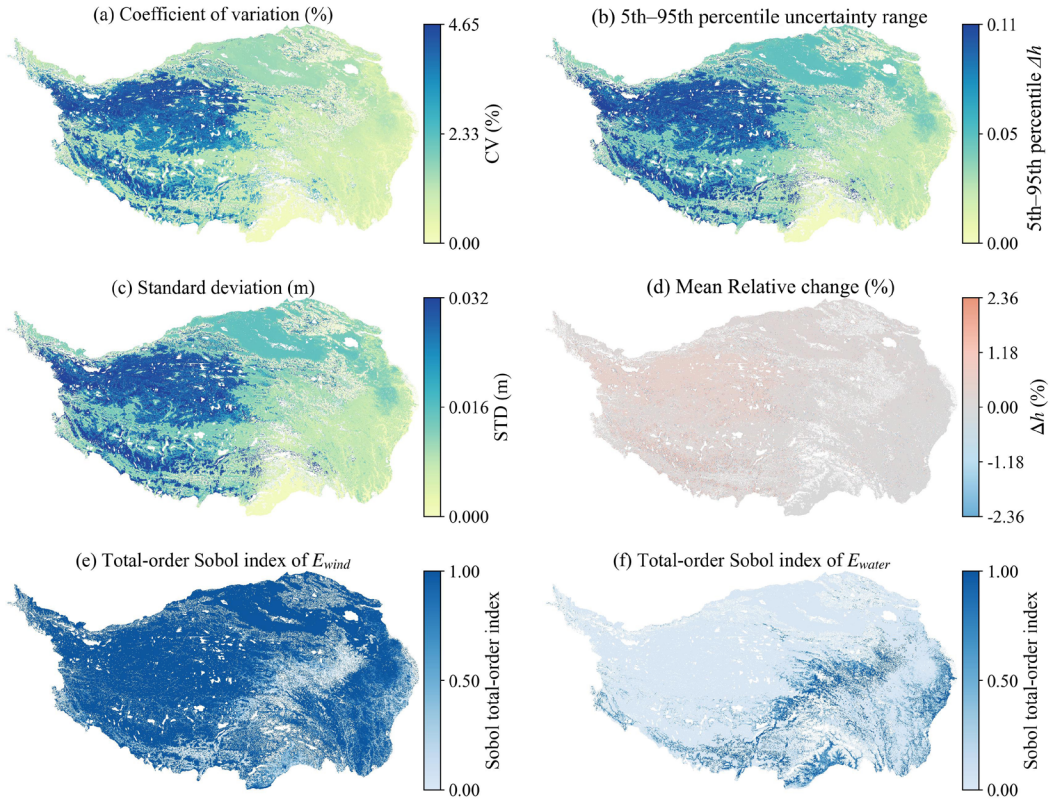


Figure 9. Spatial distribution of uncertainty and sensitivity in the simulated solum thickness across the QTP due to uncertainties in the empirical wind and water erosion inputs (RUSLE and RWEQ). (a) Coefficient of variation (CV, %), (b) 5th–95th percentile range, (c) Standard deviation of the 150 ensemble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

This CV map explicitly quantifies how the uncertainty of using these erosion models in permafrost-dominated zones propagates into our final solum thickness estimates. Notably, the uncertainty is highest in the plateau's permafrost core, directly highlighting where aeolian and cryogenic interactions deviate most from the empirical baselines.

Crucially, we have added a new paragraph to the discussion to discuss this uncertainty as a critical diagnostic feature for the Earth system modeling community, rather than merely a model limitation:

"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 $\pm 40\%$ 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–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." (Lines 534-551, tracked revised manuscript)

3) The model is driven by several external inputs, including a global machine-learning Depth-to-Bedrock product and high-resolution outputs from the Revised Universal Soil Loss Equation and the Revised Wind Erosion Equation. Because these are fundamentally empirical products, they often also fail to capture the complex sediment transport dynamics specific to alpine meadows, freeze-thaw eroded slopes, and discontinuous permafrost zones. The uncertainties inherent in these driving datasets will inevitably propagate into the final solum product. The authors should discuss whether there are methods to quantify, or ideally further reduce, the uncertainties introduced by these foundational inputs.

We agree that the uncertainties from these empirical driving datasets propagate into the final solum map. So, we performed two additional uncertainty analyses:

1 Quantifying uncertainty from empirical erosion forcing (water and wind erosion):

As the reviewer correctly points out, the empirical RUSLE and RWEQ inputs may not fully capture the complex cryogenic sediment transport dynamics of the QTP. Thus, we have introduced a comprehensive global uncertainty and sensitivity analysis using Latin Hypercube Sampling (LHS) and Saltelli sampling scheme.

We applied a conservative ±40% perturbation specifically to the empirical wind and water erosion inputs across 150 ensemble members (gravitational erosion was excluded from this perturbation, as it is an internally calculated mechanistic term driven by local topography rather than an external empirical dataset). This allowed us to generate a spatial Coefficient of Variation (CV) map (Fig. 9), which explicitly quantifies and maps the structural uncertainty propagated from the RUSLE and RWEQ datasets across different geomorphic zones. Furthermore, to distinguish the relative importance of the two empirical erosion inputs, we performed a separate Sobol total-order sensitivity analysis. In this analysis, the empirical wind and water erosion inputs were independently sampled within ±40% of their baseline values using the Saltelli sampling scheme, resulting in 512 model evaluations. The resulting Sobol total-order sensitivity indices for E_{wind} and E_{water} were then used to identify the dominant sources of uncertainty in simulated solum thickness.

2 Quantifying the impact of the DTB constraint:

The global machine-learning DTB product (Shangguan et al., 2017) serves a different function. In our revised mass balance framework, the solum thickness (hs*) is independently derived through the steady-state balance between soil production and surface erosion (Eq. 9). The DTB dataset is only used as a boundary condition to define the maximum available regolith (as a cutoff for hs*).

Therefore, uncertainties or underestimations in the global DTB product do not propagate into our calculated solum thickness unless the independently calculated solum depth exceeds the DTB threshold. To quantify this, we analyzed the intersection of our simulated solum depth with the DTB constraint. To maintain consistency with the perturbation ranges applied to E_{wind} and E_{water}, we also applied a ±40% perturbation to the DTB input and compared the results with those from the 150 ensemble simulations. Our analysis shows that, even under the −40% DTB perturbation scenario, solum thickness is truncated by the DTB constraint in only 0.001% of the total simulated area across the plateau.

"... The partitioning of the regolith into solum and saprolite ($h_c = DTB - h_s$) 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 $\pm 40\%$ 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). " (Lines 623-633, tracked revised manuscript)
"

4) In Section 3.3, the authors perform a Pearson correlation analysis and report MAT and elevation as the primary drivers of the regional thickness gradient. However, because the model's core soil production function is mathematically driven by MAT and MAP in the first place, the final results are practically guaranteed to correlate strongly with MAT and elevation. I recommend reframing this section. Rather than presenting these correlations as independent scientific discoveries about the region, they should simply be framed as an internal validation confirming that the model behaves as parameterized.

We agree with your assessment. To address this, we have completely reframed Section 3.3 Internal Validation of Environmental Drivers. We altered the section heading and revised the text to explicitly present the correlation analysis as an internal validation exercise. The purpose is now framed to confirm that the prescribed thermodynamic relationships were successfully preserved through the steady-state solver, and that the regional climatic signal was not entirely overridden by the local geomorphic noise introduced by the multi-process erosion modules.
I show some important revised text as follows:

“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. ” (Lines 361-381)

Specific comments

1. Please maintain consistency in terminology. The manuscript uses “solum thickness,” “solum thickness,” and “solum depth” interchangeably.

In the revised manuscript, we now consistently use "solum thickness" throughout the text, figures, and tables when referring to our specific simulated and measured metric (the vertical distance from the ground surface to the C horizon interface). We have removed all instances of "solum depth" to avoid confusion. In the few instances where "soil thickness" is retained for broader readability, such as in the title or general introductory statements, we have ensured it is explicitly and immediately defined as referring to the solum thickness.

2. Lines 163–165: The text initially states that freeze-thaw erosion is not considered as a separate category, but immediately follows up by describing two main freeze-thaw erosion processes. This creates a contradiction.

We thank the reviewer for pointing out this confusing phrasing. Our intent was to clarify the physical distinction between the weathering phase (frost shattering creating loose particles) and the transport phase. To resolve this, we have revised the text to avoid the term "freeze-thaw erosion" entirely. We now explicitly define frost action as a "weathering agent," clarifying that the loose material it generates is subsequently moved by the three modeled transport fluxes (wind, water, and gravity).

We revised text to clarify this confusion in Section 2:

“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.” (Lines 178-185)

3. Lines 160–170: The text repeatedly states that human activities are not considered, which is excessively lengthy.

We thank the reviewer for pointing out this redundancy. We have streamlined this paragraph, condensing the justification into a single, brief sentence.

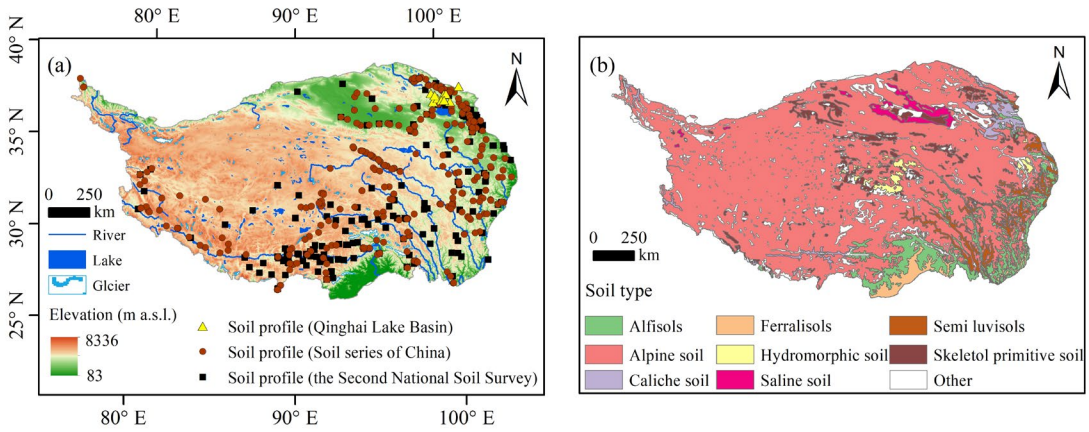
"Furthermore, given the vast area and sparse population of the QTP, the impact of direct human activities on regional-scale erosion patterns is considered negligible compared to natural forcing (Sun et al., 2020) and was therefore excluded from the model." (Lines 185-187)

4. Figure 1: Increase the font size to ensure clarity, and review the formatting of all remaining figures and tables to ensure consistency with the journal's requirements.

We have redrawn Figure 1 to significantly increase the font sizes of all labels, coordinates, and legends, ensuring clear legibility.

Furthermore, we have conducted a comprehensive review of all remaining figures and tables in the revised manuscript to verify that their layouts, resolutions, and text sizes are fully consistent with ESSD's publication guidelines.

Revised Figure 1



5. Table 2: What is the basis for the range of each parameter in Table 2? Please add supporting references.

We thank the reviewer for pointing out this omission. In the revised manuscript, we have updated Table 2 to include a "Basis / References" column. The ranges for the kinetic and transport parameters were established based on classic geomorphological compilations and previous applications of the EEMT framework.

The soil diffusion coefficient (D) is bounded based on established empirical ranges for hillslope transport (e.g., Roering, 2008), soil bulk density (ρ_s) is constrained according to the commonly reported range for mineral soils summarized by Rühlmann et al. (2006), the characteristic soil depth (h_0) and EEMT coefficients (a, b) are bounded based on the foundational works of Pelletier and Rasmussen (2009a) and Pelletier et al. (2013), and the weighting coefficients (α, β, γ) are mathematically constrained between 0 and 1 as they represent relative dimensionless scaling factors.

Table 2 Parameters of the revised mass balance model.

Parameter	Description	Value Range	Unit	Basis / References
ρ_s	Soil bulk density	2400-2900	$kg \cdot m^{-3}$	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	$m \cdot a^{-1}$	
b	Climate response coefficient	0-0.1	$m^2 \cdot a \cdot kJ^{-1}$	
α	Hydraulic erosion weight	0-1	-	Mathematically constrained
β	Aeolian erosion weight	0-1	-	
γ	Gravitational erosion weight	0-1	-	
D	Soil diffusion coefficient	0.00001-0.01	$m^2 \cdot a^{-1}$	Roering (2008)

6. Lines 272–274: The description of the distribution of cluster 6 is inaccurate.

We thank the reviewer for pointing out this geographical inaccuracy. We have carefully re-examined the spatial distribution of the clusters in Figure 2b and agree that grouping Clusters 5 and 6 together as broadly "northeastern and northern" was imprecise. While Cluster 5 is predominantly located in the northeastern QTP, Cluster 6 extends across the central and eastern-central portions of the plateau, including the Three Rivers Source Region.

We have corrected this description in the revised manuscript to accurately reflect the distinct spatial footprints of these two clusters.

"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." (Lines 291-294)

7. Figure 7: Insert the corresponding legend for soil thickness.

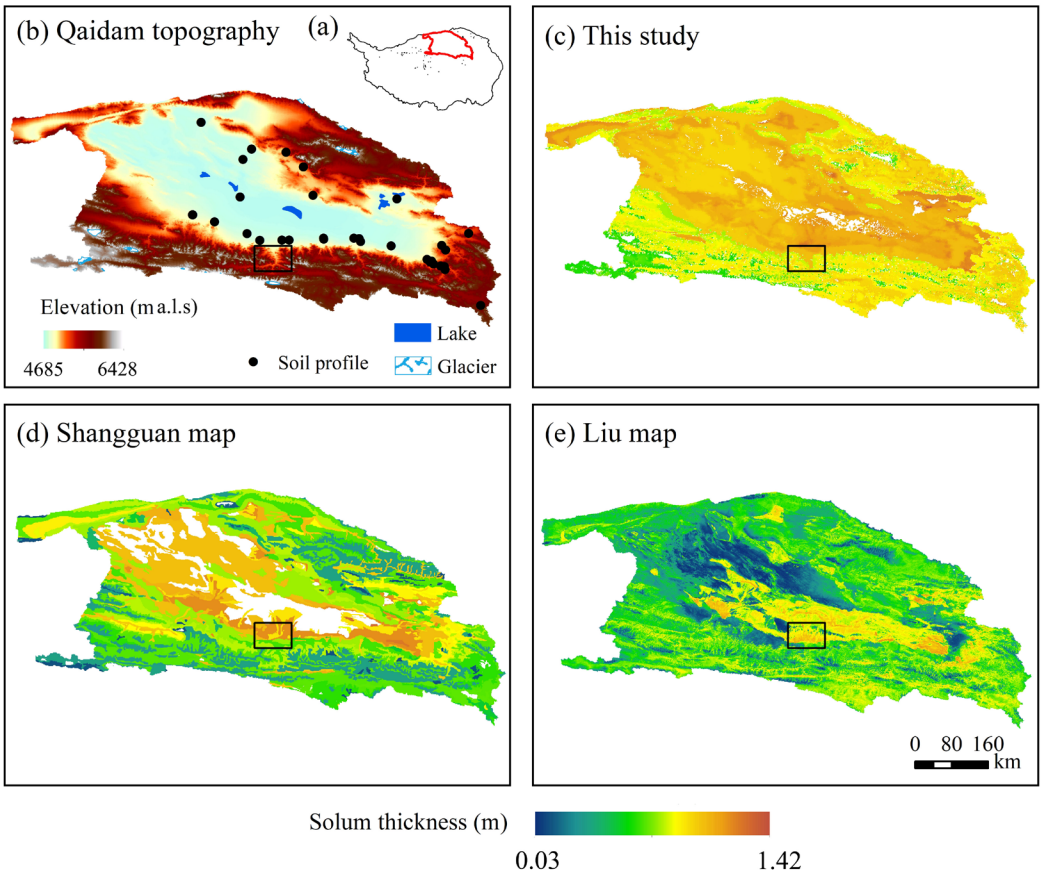
We thank the reviewer for pointing this out. In the original manuscript, the legend for solum thickness was located within panel (e). To maintain a clean and uncluttered visual layout, we prefer not to repeat the identical legend across subplots (c), (d), and

(e). However, we agree that its previous placement may not have clearly indicated that the scale applied to all three maps.

To resolve this, we have redrawn Figure 7. We extracted the solum thickness color bar from the individual panel and enlarged it into a single, prominent, unified legend positioned to clearly serve the spatial maps in panels (c), (d), and (e).

Note to Reviewer: The same unified legend formatting has also been applied to Figure 8 for consistency.

Revised Figure 7

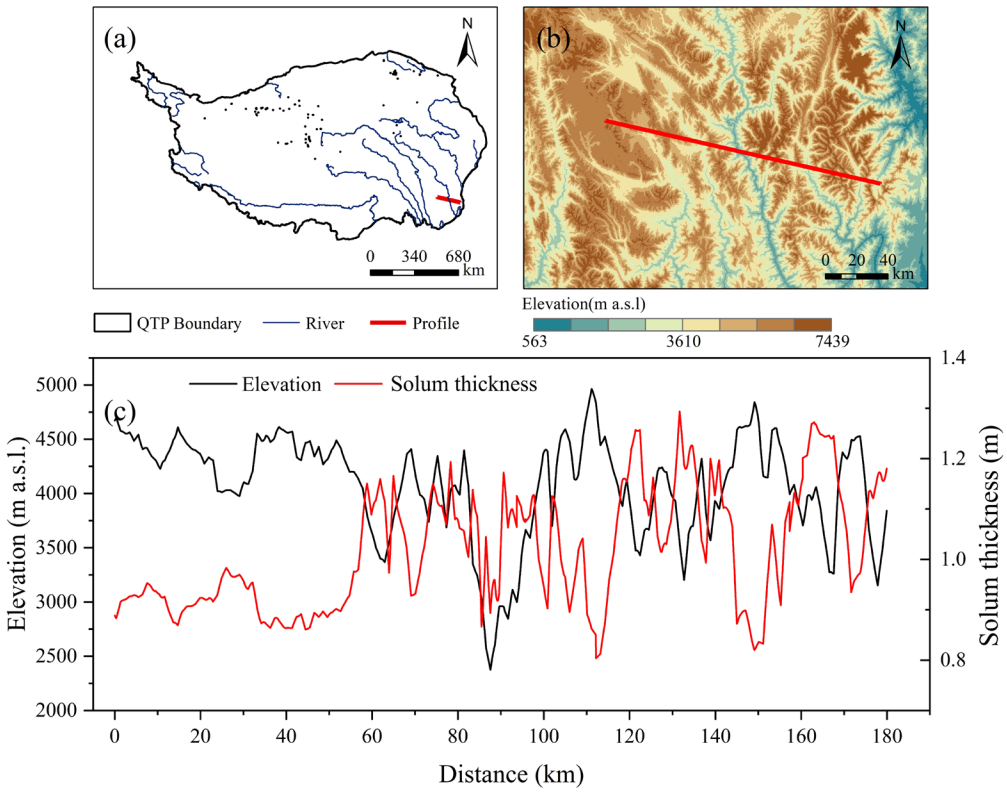


8. Figure 5-8: Please add legend to Fig. 5 and carefully review the captions for Figs. 6–8.

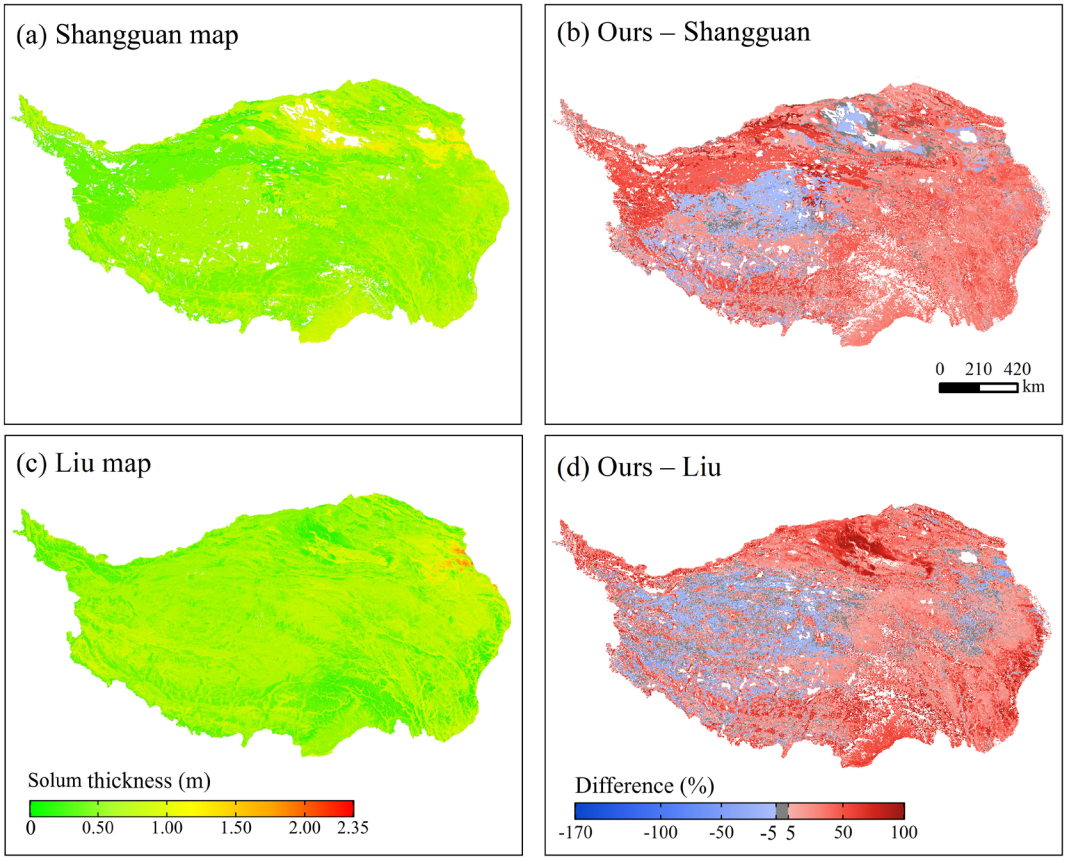
We thank the reviewer for their careful attention to the figures. We agree that the omission of the elevation legend in Figure 5b made the topographic transect difficult to interpret. In the revised manuscript, we have updated Figure 5 to include a clear legend in panel (b).

Furthermore, we have thoroughly reviewed and revised the captions for Figures 6, 7, and 8. We have expanded these captions to provide more explicit context for the spatial patterns shown, and make sure the use of consistent "solum thickness" as established in our response to Specific Comment #1.

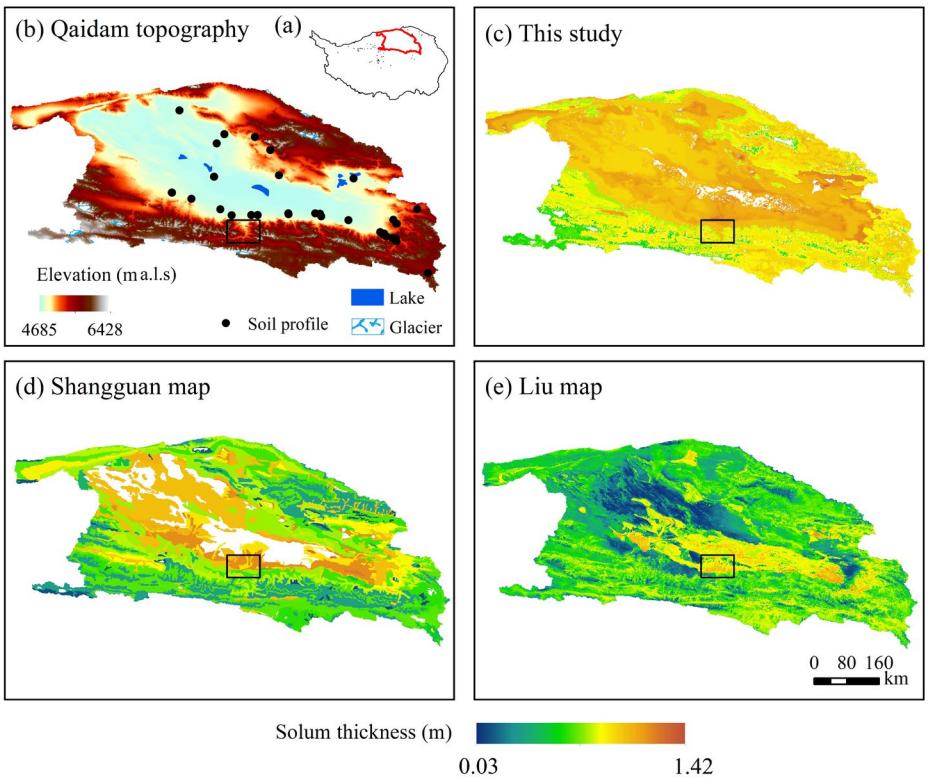
Revised Figure 5



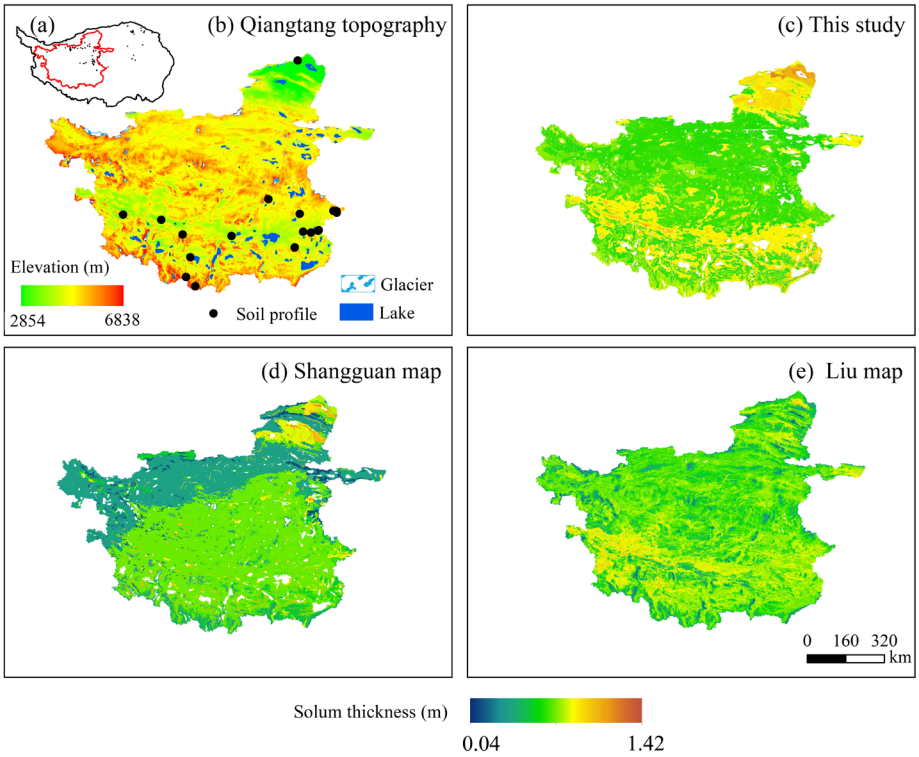
Revised Figure 6



Revised Figure 7



Revised Figure 8


(a)
(b) Qiangtang topography
(c) This study
(d) Shangguan map
(e) Liu map
Elevation (m)
2854
6838
Soil profile
Glacier
Lake
0 160 320
km
Solum thickness (m)
0.04
1.42

Response to RC2

We sincerely thank the Referee #2 for their constructive evaluation and for recognizing the scientific merit, novelty and potential value of our map. We deeply appreciate your rigorous critique of our initial uncertainty quantification and validation strategy. In response, we have implemented major revisions, most notably reframing our validation claims to explicitly acknowledge the shared-data limitations, adding new spatial pattern analyses, and introducing quantitative global uncertainty analysis using Latin Hypercube Sampling.

Below, we provide a detailed, point-by-point response to your comments. Your original comments are presented in **blue**, our responses are given in **black**, and the corresponding revisions added to the manuscript are provided in **red.**

This manuscript presents a new 1 km resolution solum thickness dataset for the Qinghai-Tibet Plateau (QTP) developed using a revised geomorphic mass balance model. The topic is highly relevant and timely. The methodological innovations represent a thoughtful and non-routine adaptation of classical mass balance models to the QTP's unique conditions. This work has strong potential value for the soil science, cryospheric, and Earth system modeling communities.

Overall, the scientific core has clear merit and novelty. However, several non-trivial weaknesses in validation strategy and uncertainty quantification prevent acceptance in the current form. With careful attention to the major and minor points below, a revised version has a good chance of acceptance in ESSD.

We thank the reviewer for this highly encouraging assessment of our manuscript's scientific significance. We fully acknowledge your concerns regarding the initial validation strategy and uncertainty quantification. As detailed in our point-by-point responses below, we have dedicated our efforts to directly resolving these two issues. We believe these major revisions have fundamentally strengthened the manuscript.

Major Comments

1 The performance assessment relies on 4-fold cross-validation using the same 552 soil profiles that overlap significantly with the data sources used to construct the benchmark maps (Shangguan and Liu). Therefore, the claims of outperforming existing products by 10–17% in RMSE and MRE are weakened by the lack of truly independent test data. Please provide a clearer acknowledgment of this shared-data limitation. Additionally, spatial pattern analysis (e.g., correlation of modeled thickness with topographic

curvature, topographic position index, or slope) might be helpful in enhancing assessment.

We thank the reviewer for this critical observation. We strongly agree that this shared-data limitation must be explicitly addressed.

To clarify the nature of the data structure: as detailed in Section 2.2, our 552 soil profiles are an integrated compilation of two historical archives: the Second National Soil Survey (1979–1985) and the National Soil Series Survey (2009–2019), and supplemented by additional independent regional samples. Because our reported performance metrics were generated using a 4-fold cross-validation on this entire compiled dataset, the test folds inevitably contained data shared with both the Shangguan map (which relied on the 1979–1985 data) and the Liu map (which relied on the 2009–2019 data).

While our dataset does contain a substantial number of independent samples relative to each individual baseline map, the reviewer is strictly correct that the final cross-validated metrics themselves lack true independence. However, compiling a massive, 100% independent dataset is currently impossible due to limited soil profile available on the extreme physically inaccessible QTP. Therefore, all large-scale mapping efforts have to utilize the same limited legacy archives.

To comprehensively address this limitation, we have implemented a two-part revision:

1 Acknowledging the share-data limitation. We have revised the Section 4.3 (Limitations and Future Directions) and the discussion about Table 5 to explicitly acknowledge this constraint and temper our performance claims. We now strictly frame the statistical comparison as a relative assessment under shared observational constraint rather than definitive proof of generalization to untested regions. In Section 2.5, we added:

"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. " (Lines 264-266)

In Section 4.3, we edited:

"As noted in Section 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." (Lines in 423-425)

2 Deepening the spatial pattern analysis. We now explicitly use the topographic correlation (Slope Position r = 0.25, TRI r = 0.24, and Slope r = 0.25 from Figure 4)

and the Hengduan Mountains transect (Figure 5) as additional spatial validation to verify that the model correctly reproduces physically plausible, terrain-driven differentiation at the hillslope scale. Updated the spatial pattern validation part to reference the exact Hengduan Mountains transect explanation and curvature correlation text. Key revision in Section 3.3:

"... 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." (Lines 382-399)

2 The steady-state assumption is also a critical simplification on a tectonically active and permafrost-degrading plateau. While limitations are discussed, there is no quantitative propagation of uncertainty into the final dataset. I suggest the authors can add some sensitivity analysis. For example, vary key parameters ±20–50% or using Monte Carlo sampling for weighting coefficients and present spatial uncertainty estimates, then discuss how uncertainty varies across clusters or geomorphic zones.

As already detailed in our reply to Referee #1 (Comment 1), we performed a comprehensive uncertainty and sensitivity analysis, including:

1 One-At-A-Time (OAT) ±20% perturbations of erosion weighting coefficients (α, β, γ) (Fig. 10, Appendix C)

2 Latin Hypercube Sampling ±40% perturbation of empirical erosion inputs (Fig. 9)

3 Sobol total-order sensitivity analysis

Full details, figures, and discussion of implications for the steady-state assumption are provided in our response to Referee #1 and in Section 4.2-4.3 of the revised manuscript.

Minor Comments

Abstract

- Lines 24–25: Specify the metric(s) in the performance claim (e.g., "by approximately 10–17% in RMSE and MRE").

In the revised abstract, we have specified the exact metrics achieved by the model (RMSE of 0.34 m and MRE of 0.78). Furthermore, consistent with our revisions regarding the shared validation data (detailed in our response to Major Comment #1), we have carefully rephrased this performance claim. Rather than claiming absolute independent outperformance, we explicitly state these metrics represent a relative reduction under the available cross-validation framework.

"The resulting dataset reveals that solum thickness ranges from 0.39 m 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 compared with existing national-scale products under shared observational constraints." (Lines 24-30)

Study Area and Data

- Line 105: Briefly mention known limitations of the Shangguan et al. (2017) DTB dataset in the QTP interior when it is first introduced.

We thank the reviewer for this helpful suggestion. We have revised the text in Section 2.1 to explicitly note the limitations of the Shangguan et al. (2017) dataset regarding the QTP interior. In section 2.1:

"... 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." (Lines 118-122)

Methods

- Lines 165–170: Justification for excluding human activities is too brief. Note that effects are assumed negligible at 1 km resolution.

In the revised manuscript, we have expanded this explanation to explicitly clarify that, given the 1 km spatial resolution and the sparsely populated nature of the QTP, geomorphic and climatic factors overwhelmingly dominate large-scale erosion patterns.

"... 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." (Lines 185-187)

- Table 1: Move "Parent Material of Soil Formation" to a new "Geological" category.

Thanks. We've updated Table 1 in response to your suggestion.

Results

- Table 3: Add "(–)" for the unitless NDVI column.

Many thanks and we've fixed this formatting problem.

- Table 4: Replace "-" for Cluster 8 RMSE with "N/A (n=4 insufficient for cross-validation)". Add a footnote explaining *n*.

We have updated Table 4 in the revised manuscript to replace the "-" with "N/A". We explained n as the number of soil profile observations used for calibration in each cluster. N/A: the sample size of 4 is insufficient for reliable cross-validation.

- Fig. 3: Consider adding a histogram inset or basin-specific summary statistics.

In the revised manuscript, we have updated Figure 3 to include a new histogram inset (panel b) and a basin-level boxplot of solum thickness (panel c). Furthermore, we have updated the main text to explicitly interpret this distribution, highlighting that the model predominantly produces moderately developed soils, with thin and thick extremes cleanly partitioned by geomorphic regimes. We also added a description of the differences in solum thickness among different basins. For example, the Yarlung Zangbo Basin in southeastern QTP exhibits the greatest solum thickness, but also the largest standard deviation due to its complex terrain. In contrast, the permafrost-dominated basins in the inner QTP show the smallest solum thickness and the lowest internal variability.

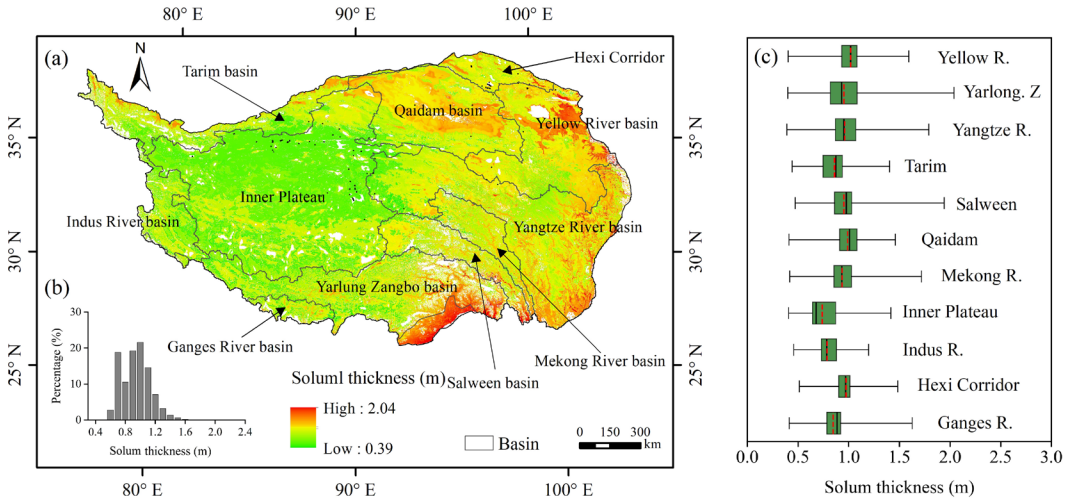


Figure 3. Spatial pattern and frequency distribution of simulated solum thickness across the QTP. (a) Spatial distribution of simulated solum thickness on the QTP. The delineated areas for the Mekong, Salween, and Yarlung Zangbo (Brahmaputra) rivers represent the catchment sections located within China. White patches indicate 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.

Table B1 (Appendix B). 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

- Lines 330–340: explain why curvature shows near-zero correlation at plateau scale yet drives local differentiation in the model.

We thank the reviewer for highlighting this apparent paradox. we have addressed it as part of our broader revision to Section 3.3. In the revised text, we explain that general curvature fluctuates rapidly between positive (divergent ridges) and negative (convergent valleys) at the local scale. When aggregated across the immense 2.5 million km² area of the plateau, this high-frequency spatial signal statistically averages out to near-zero, masked by the dominant, low-frequency climatic gradients. However, at the hillslope scale, the physical mass balance model correctly preserves and responds to these local terrain features. In section 3.3:

"... 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." (Lines 382-399)

Discussion and Comparison

- Table 5: Explicitly note the shared validation data limitation. explain the higher MAE.

We thank the reviewer for this important suggestion. We have addressed both points in the revised text accompanying Table 5.

1 Consistent with our comprehensive revisions for Major Comment #1, we have updated the text introducing Table 5 to explicitly acknowledge the overlapping nature of the validation profiles. We now strictly frame these metrics as a relative assessment under shared observational constraints rather than a fully independent validation.

2 We have expanded the discussion to explain why our model exhibits a slightly higher Mean Absolute Error (MAE of 0.13 m) compared to the Shangguan map (0.06 m).

The legacy statistical maps tend to produce spatially smoothed estimates. While statistical smoothing mathematically minimizes average absolute errors, it artificially masks sharp, terrain-driven physical extremes. In contrast, our physical mass balance model intentionally preserves this high-frequency spatial heterogeneity (e.g., thin regolith on steep ridges and thick accumulations in valleys). Therefore, the slightly higher MAE is a natural byproduct of capturing a wider, more physically realistic range of solum thicknesses. In Section 3.4.2:

"As noted in Section 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 (Fig. 3 and Fig. 5). The reference maps, by contrast, tend to predict spatially smoothed values, which artificially minimizes absolute error but fails to represent the landscape's physical 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." (Lines 423-432)

- Lines 470–485: Strengthen discussion of potential RUSLE/RWEQ biases in permafrost areas and suggest future use of cryogenic-specific erosion models.

In the revised manuscript, we have significantly strengthened the discussion regarding potential biases in RUSLE and RWEQ estimates. We explicitly acknowledge that these models may not fully account for the complex interactions of seasonal freeze-thaw cycles and ice-rich permafrost layers on sediment mobility. Furthermore, we have added a dedicated recommendation for the future development and integration of QTP-specific cryogenic parameterizations for erosion and weathering to reduce these systematic uncertainties. In Section 4.2:

"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. " (Lines 534-539)

"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. "(Lines 617-622)

- Lines 500–510: Rephrase the sentence about weighting coefficients mitigating steady-state bias to "partially mitigated".

We thank the reviewer for this important correction. In the revised manuscript, we have rephrased this section to explicitly state that the weighting coefficients only "partially" reduce the biases associated with the steady-state assumption.