The study site appears to be the bottom and adjacent toe slopes of a valley that was glaciated in late Quaternary time. I write "appears to be" because I do not know the glacial geology of this region but merely note that large moraines are present elsewhere in the valley in Google Earth imagery. The equations of this paper are applicable to hillslopes dominated by soil production and colluvial and overland flow sediment transport processes and where an approximate balance between the erosion caused by those processes and rock uplift has been achieved. Glacial erosion and/or deposition is not considered, yet these are possibly the dominant processes in this study site.

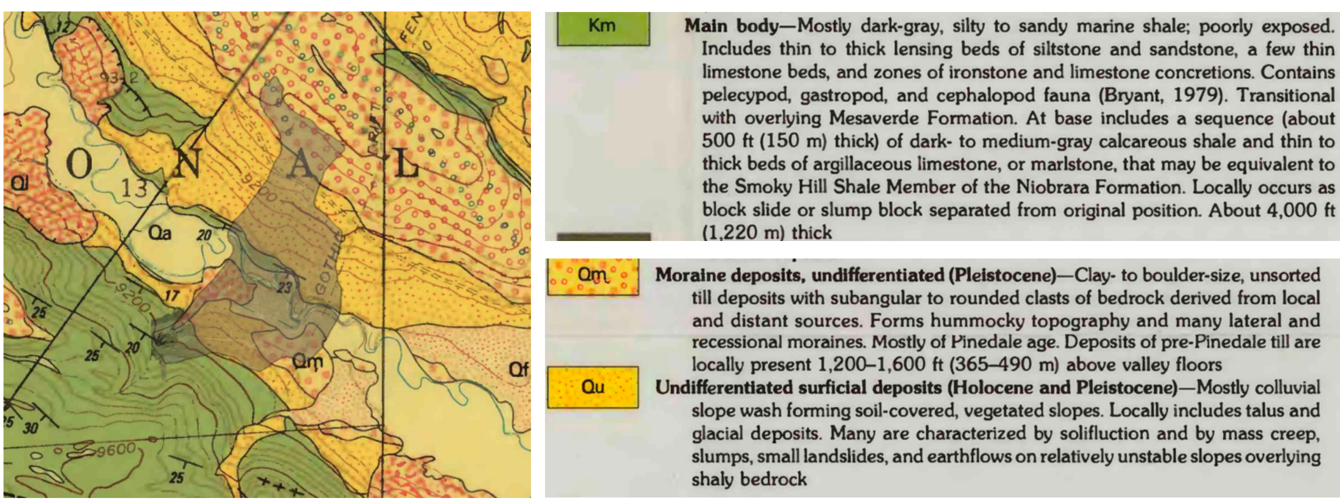
<u>*Our response:*</u>
We thank the reviewer for pointing this out. Glacial deposits are mapped throughout the watershed though they are rather isolated in that they have a limited spatial extent and are not characteristic for the depositional environment. In most boreholes that have been drilled there was no such layer. For the areas we analyzed, the amount of moraine deposits is indeed small (see the figure below). We add the following paragraph in Section 3:

> *"The last glacial advancing and retreating in the Upper Colorado River Basin is dated between 16.1 and 20.8 ka (Brugger, 2010). Glacial deposits are mapped at many locations throughout the watershed (Gaskill, 1991), but they are rather isolated and have a limited spatial extent, including in the area analysed in this study. In most boreholes that have been drilled in the studied area, the material beneath the soil layer is weathered shale (or saprolite) in the size roughly between 0.2 cm to 5 cm with light brown color (Fig. S3)."*

Reference:

Brugger, K. A.: Climate in the Southern sawatch range and Elk Mountains, Colorado, U.S.A., during the last glacial maximum: Inferences using a simple degree-day model, Arctic, Antarct. Alp. Res., 42(2), 164–178, doi:10.1657/1938-4246-42.2.164, 2010.

Gaskill, D. L.: Geologic map of the Gothic quadrangle, Gunnison County, Colorado., doi: 10.3133/gq1689, 1991.



Indirect measurements of soil thickness can be useful in augmenting measurement of soil thickness in excavated soil pits, but it is inadvisable to use such methods in isolation. Augering and penetration methods are generally considered to be minimum values in rocky soils because the auger or penetrometer can be stopped by gravel. The study would be strengthened by reporting observations from soil pits at representative locations so that the nature of the soil-saprolite contact and the presence/absence of any glacial tills in the study site can be ascertained.

Our response:

We acknowledge that measuring soil thickness at numerous locations is not simple and error-free. At this study site, for both auger and CPT measurements, we obtained measurements more than once at sampling sites where the results were suspicious. In addition, at sampling sites where the auger or CPT were stopped at a relatively shallow depth, we measured multiple times at nearby spots within about 1m diameter. Further, we tested the accuracy of the CPT measurement and found that the CPT i) shows largest change in resistance when entering weathered bedrock, and ii) is stopped very sharply only in the presence of a boulder, in which case the resistance is so strong that the measurement was deemed suspicious and repeated nearby. Because the CPT may not clearly identify the potential presence of moraine deposits, we also visually inspected the soil and saprolite materials extracted by the auger (see field images). We believe our measurement is relatively accurate and efficient, which provides a consistent assessment of soil thickness over space in comparison to other existing methods. We include the following images in the supplementary information and revise the last paragraph in Section 3 in the main text:

> *"At this study site, we used and compared both auger and CPT measurements to estimate soil thickness. The CPT measurements provide a vertical profile of soil resistance for a soil column. We tested the accuracy of the CPT measurement and found that the CPT i) shows largest change in resistance when entering weathered bedrock, and ii) can be stopped very sharply only in the presence of a boulder, in which case the resistance is so strong that the measurement was deemed suspicious and repeated nearby. Because the CPT may not clearly identify the potential presence of moraine deposits, we also visually inspected the soil and saprolite materials extracted by the auger. From the auger, the transition zone from soil to the saprolite or bedrock is based on the material size and color of retrieved samples (Fig. S3). When the auger reaches the bedrock shale, it cannot penetrate easily. We believe our measurement is relatively accurate and efficient, which provides a consistent assessment of soil thickness over space in comparison to other existing methods. Figures 1b-1e show the relationship between soil thickness estimated from auger, CPT, and local elevation. There is a high variation in soil thickness from local to hillslope scales. To fully take advantage of all the sampling data, we used auger data to fit values for CPT (Fig. S4) since more locations are sampled from auger drilling than CPT. The CPT and auger data are mostly in agreement. For soil thicknesses less than ~0.5 m, the CPT data are slightly higher than the auger, and for soil thickness larger than ~0.5 m, the CPT data are slightly lower."*

The authors state that the study site is in a soil-thickness steady state condition but Fig. S2 shows that this not to be the case. That figure shows the average soil thickness in a simulation increasingly steadily at the time (20 ky) when a steady state condition has purportedly been achieved. In stating that this figure demonstrates soil thickness steady state, I presume that the authors are referring to the fact that the rate of soil production is in decline. However, the rate of soil production is still far from zero and the rate of soil production for the simple case of a horizontal surface will result in a similar decline even as the soil thickness approaches infinity given sufficient time. Moreover, the average soil thickness should not be used to infer steady state because decreasing soil thickness in one area may be balanced out by increasing soil thickness elsewhere.

Our response:
We thank the reviewer's comment on this issue. We agree that the average soil thickness is not appropriate to infer the steady state condition. At depositional areas, the soil gradually accumulates on the land surface, and meanwhile, the soil formatting slows at the bottom, therefore, the soil thickness is supposed to gradually increase (Dietrich et al, 1995). Due to the complexity of soil deposits, such as expansion or compression of soils, we consider that using an empirical relationship is appropriate for the soil thickness at depositional areas. At erosional sites, the erosion from the land surface can be balanced out by the soil formation from the bottom. Therefore, we only apply the mass conservation method in erosional 2-D grid cells (see answer to the next question below). We also extend the simulation to a longer period than in the previous simulation to see if the model can reach steady state. Even though it takes almost 25 ka, which is longer than the last great glacial history that ends about 16-20 ka; we believe it is still acceptable because (1) the elevation data that we input in the model is the current lidar DEM rather than the historical topography, and (2) the parameters are from field calibration which follows ambient soil thickness rather than the one when the glacial period ended. The Figure S5 has been updated as:

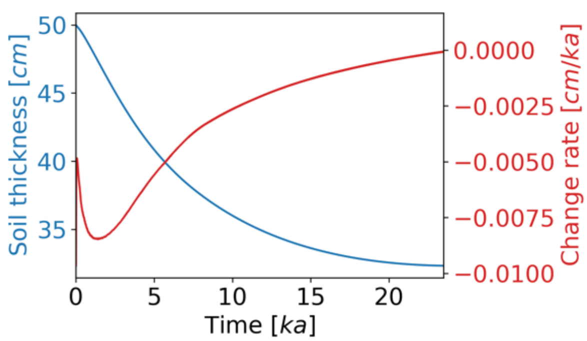


Figure S5: Spatial mean values of erosional sites of soil thickness evolution over time. The initial soil thickness is 0.5 m, time step is 1 year, and the initial elevation is the current DEM data. The boundary condition is Neumann boundary condition, the surface transport fluxes around the edge is zero. The time step is 1 yr, and the diffusion coefficient is m^2/yr for the north-facing hillslope and m^2/yr for the south-facing hillslope.

The revised sentences in Section 4.3 are:

"*At erosional sites, the erosion from the land surface can be balanced out by the soil formation from the bottom, therefore, the soil thickness may reach a steady state condition. By coupling soil thickness with landscape evolution, we found that the soil thickness reaches a dynamic steady-state after approximately 25 kyr at this study site (Fig. S5), which is consistent with other studies in mountainous areas (Dietrich et al., 1995; Vanwalleghem et al., 2013). This implies that the current soil thickness in the East River Watershed may have already reached a steady-state condition since the last glacial legacy. Here, we only focus on the steady state condition at erosional sites because they are where we apply the mass conservation equation. For depositional sites, the soil gradually accumulates on the land surface;*

and meanwhile, the soil weathers slowly at the bottom; therefore, the soil thickness is supposed to continuously increase (Dietrich et al, 1995). Due to the complexity of soil depositional environments, such as expansion or compression of soils, we consider that using an empirical relationship is appropriate for the soil thickness at depositional areas. "

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

Vanwalleghem, T., Stockmann, U., Minasny, B. and Mcbratney, A. B.: A quantitative model for integrating landscape evolution and soil formation, J. Geophys. Res. Earth Surf., 118(November 2011), 331–347, doi:10.1029/2011JF002296, 2013.

The authors state (line 35) that the mass conservation method can return no finite soil thickness but this has not been demonstrated. Without such a demonstration application of the Patton method seems premature. If the model is applicable to the study site and correctly parameterized it should return a finite soil thickness value. Perhaps the model was unable to achieve such a steady state because the actual study site is not in steady state (e.g. recently glaciated) or because the authors used an inappropriate method for computing discharge by overland flow (d8 or steepest descent, as shown in Fig 1.)). In any case the paper would be strengthened by applying the model to the study site with appropriate assumptions first (see Pelletier et al., JGR, 2011 for examples of soil production and transport modeling without a steady state assumption) and then using the Patton method if and only if the model can be demonstrated to produce no adequate solution for any reasonable set of parameter values.

<u>*Our response:*</u>

We agree that soil production and transport rate modeling do not require a steady-state assumption. For the mass conservation equation, if we know the initial value of soil thickness and proper boundary conditions, then the steady-state assumption is unnecessary because the mass conservation equation is a partial different equation that can be solved numerically. The steady-state assumption is only needed for the soil thickness estimation in the assumption that the regolith production balances the physical erosion, as used in other studies (Pelletier and Rasmussen, 2009; Pelletier et al, 2011; Dietrich et al, 1995).

The equation to calculate soil thickness, h:

$$\frac{\rho_r}{\rho_s}\frac{1}{cos\theta_s}P_o e^{-hcos\theta/h_o} - \nabla \cdot q_d - \nabla \cdot q_s = 0 \qquad (*)$$

can be rearranged as:

$$\frac{\rho_r}{\rho_s}\frac{1}{cos\theta_s}P_o e^{-hcos\theta/h_o} = \nabla \cdot q_d + \nabla \cdot q_s \qquad (**)$$

Because the soil production rate is always larger than or equal to zero under any conditions, expressed as $\frac{\rho_r}{\rho_s}\frac{1}{cos\theta_s}P_o e^{-hcos\theta/h_o} \geq 0$,

from here, the only way for the soil thickness 'h' to have a real number is that $\nabla \cdot q_d + \nabla \cdot q_s > 0$.

In a landscape evolution model, the values calculated by $\nabla \cdot q_d + \nabla \cdot q_s$ on each 2-D grid cell can be either positive, meaning erosional process; or negative, meaning a depositional process. In depositional cells where $\nabla \cdot q_d + \nabla \cdot q_s < 0$, then there is no real number for 'h' in equation (**). For this reason, we apply Patton's method on cells whichever undergoes depositional process.

We compute our overland flow following the shallow water overland flow equation as shown in Equation 4: $\frac{\partial H_w}{\partial t} = \frac{\partial}{\partial x}\left(D_h \frac{\partial H_w}{\partial x}\right) + \frac{\partial}{\partial y}\left(D_h \frac{\partial H_w}{\partial y}\right)$. The details of this equation can be found in Lal (1998) and Yan et al (2019).

We deleted "*... can return no finite soil thickness...*" and revised the sentence in the updated manuscript:

> "*…there is no real number for soil thickness if* $\nabla \cdot q_d + \nabla \cdot q_s < 0$..."

We include the following in Section 2.1.4 in the revised manuscript:

> *" **2.1.4 Combine the mass conservation method with the empirical method***
> *For the mass conservation equation, the steady-state assumption is needed for the soil thickness estimation in the assumption that the regolith production balances the physical erosion, as used in other studies (Pelletier and Rasmussen, 2009; Pelletier et al, 2011; Dietrich et al, 1995). Therefore, the mass conservation method with the steady-state assumption can be used to solve the soil thickness at erosional sites but has limitations at depositional sites (Eqn. 7, Dietrich et al, 1995). Patton's method is better adapted to depositional sites. However, it can provide negative values of soil thickness at zones with high negative-curvature values where erosion is the main process. Also, in a low gradient and divergent area, if the soil transport rate is assumed as a linear relationship with curvature (i.e.,* $\nabla \cdot q_d = -K_d \nabla \cdot \nabla \eta$, *and* $\nabla \cdot q_s = 0.0$*), then Equation 7 can be further simplified in that the soil thickness has a natural logarithm relationship with curvature (i.e.,* $h = -m * ln(\nabla \cdot \nabla \eta) + C$, *where* m *and* C *are constant parameters that can be calibrated from field sampling data). However, Patton's method (Patton et al., 2018) always assumes a linear relation. This may be why his empirical relationship does not work very well in the erosional areas. However, this can be compensated for by using the mass-conservation method."*

Reference:

Lal, A. M. W.: Performance comparison of overland flow algorithms, J. Hydraul. Eng., 124(4), 342–349, doi:10.1061/(ASCE)0733-9429(1998)124:4(342), 1998.

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

Pelletier, J. D. and Rasmussen, C.: Geomorphically based predictive mapping of soil thickness in upland watersheds, Water Resour. Res., 45(9), 1–15, doi:10.1029/2008WR007319, 2009.

Pelletier, J. D., et al. (2011), Calibration and testing of upland hillslope evolution models in a dated landscape: Banco Bonito, New Mexico, J. Geophys. Res., 116, F04004, doi:10.1029/2011JF001976.

Yan, Q., Le, P. V. V, Woo, D. K., Hou, T., Filley, T. and Kumar, P.: Three-Dimensional Modeling of the Coevolution of Landscape and Soil Organic Carbon, Water Resour. Res., 55(2), 1218–1241, doi:10.1029/2018WR023634, 2019.

Some examples of unclear or incorrect methodology, incorrect units, parameter values that are very different from the literature, etc.:

Our response:

We appreciate the reviewer's suggestion and address the questions in the following one-by-one.

1) please state how eqn. (7) was solved to determine h; this is the heart of the modeling and it is difficult to evaluate the paper without any mention of the solution method,

Our response:

In Equation 7:

$$\frac{\rho_r}{\rho_s}\frac{1}{cos\theta_s}P_o e^{-hcos\theta/h_o} - \nabla \cdot q_d - \nabla \cdot q_s = 0$$

By rearranging equation 7, we get

$$h = -\frac{h_o}{cos\ cos\ \theta}\ ln\ ln\ \left[\frac{cos\theta_s(\nabla \cdot q_d + \nabla \cdot q_s)}{P_o}\frac{\rho_s}{\rho_r}\right]$$

From equation 3a and 3b:

$$\nabla \cdot q_s = \frac{q_{s,out} - \Sigma q_{s,in}}{d_s}$$

$$q_s = K_s H_w^\alpha S^\beta$$

From equation 2:

$$q_d = -\frac{K_d \nabla\eta}{1 - \left(\frac{\nabla\eta}{S_c}\right)^2}$$

$$\nabla \cdot q_d = -\nabla \cdot \frac{K_d \nabla\eta}{1 - \left(\frac{\nabla\eta}{S_c}\right)^2}$$

In our method, $\nabla \cdot q_d$ and $\nabla \cdot q_s$ are independent of soil thickness h. We add the following after equation 7:

"*The soil thickness value* h *can be directly solved here because* $\nabla \cdot q_d$ *and* $\nabla \cdot q_s$ *are independent of soil thickness*".

2) the drainage map in Fig. S1 suggests that D8 or steepest descent is used to determine sediment flux by overland flow; this method is incapable of modeling discharge by overland flow; D-infinity or another multiple flow direction routing algorithm must be used,

Our response:

Regarding the overland flow, as explained above, we use the diffusive overland flow equation ($\frac{\partial H_w}{\partial t} = \frac{\partial}{\partial x}\left(D_h \frac{\partial H_w}{\partial x}\right) + \frac{\partial}{\partial y}\left(D_h \frac{\partial H_w}{\partial y}\right)$), which is Equation 4 in the manuscript. We only use ArcGIS (the D8 algorithm) for the purpose to delineate the boundary of our study site in Fig S1.

3) the h_0 value of 0.1-0.125 m is much smaller than other studies; h_0 typically ranges from 0.2-0.5 m, see papers by Heimsath et al.;

Our response:

We appreciate that the reviewer pointed this out. The values were typed wrongly. It should have been 0.2 and 0.18 for the north-facing and south-facing hillslope, respectively. The values have been updated in Table 1.

We agree that among a series of Heimsath's papers and Dietrich et al (1995), h_o ranges from 0.2-0.5 in the form of $e^{(-\frac{h}{h_o})}$. One possible reason that our value is slightly smaller than this range is that we defined h as the distance along the norm direction to the land surface, which give $e^{-hcos\theta/h_o}$, where θ is the slope of the land surface in degree (Pelletier and Rasmussen, 2009). In our study site, $cos\theta$ ranges from 0.77-1.0; and among the sampling sites, $cos\theta$ ranges from 0.85-0.98. The value of h_o is calibrated based on the sampling data. Specifically, $\frac{cos\theta}{h_o} = \frac{1}{h_o/cos\theta}$, and $h_o/cos\theta$ gives the range between 0.18/0.77 (=0.20) to 0.2/0.98 (=0.23).

References:

Heimsath, A. M., Dietrich, W. E., Nishiizumi, K. and Finkel, R. C.: The soil production function and landscape equilibrium, Nature, 388(July), 358–361, 1997.

Heimsath, A. M., Chappell, J., Dietrich, W. E., Nishiizumi, K. and Finkel, R. C.: Soil production on a retreating escarpment in southeastern Australia, Geology, 28(9), 787–790, doi:10.1130/0091-7613(2000)28<787:SPOARE>2.0.CO;2, 2000.

Heimsath, A. M., Furbish, D. J. and Dietrich, W. E.: The illusion of diffusion: Field evidence for depth-dependent sediment transport, Geology, 33(12), 949–952, doi:10.1130/G21868.1, 2005.

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

Pelletier, J. D. and Rasmussen, C.: Geomorphically based predictive mapping of soil thickness in upland watersheds, Water Resour. Res., 45(9), 1–15, doi:10.1029/2008WR007319, 2009.

4) what is the relation of Bp to P_0? only P_0 appears in the equations yet only Bp appears in Fig. 3; if these are related by the bulk density ratio, as I would have thought, why is one larger on the n-facing side while the other is larger on the s-facing side?

Our response:

We appreciate that the reviewer pointed this mistake. In the older version, $B_p = \frac{\rho_r}{\rho_s} P_o$, where ρ_r and ρ_s are rock and soil bulk density, respectively. B_p is used for the purpose of simplicity for calculation and thus should not be considered as an extra parameter in this work. In Figure 3, it is actually P_o not B_p. In the revised manuscript, we delete B_p from Table 1, and revise Figure 3 accordingly.

5) values of alpha and beta are not reported,

Our response:

We thank the reviewer for pointing this mistake. The values and references for the beta have been added in the revised manuscript. Because beta and alpha share the same value, we use alpha only. The revised equation and texts are shown below:

$$q_s = K_s(H_w S)^\beta \quad (3b)$$

where K_s is the soil erodibility coefficient $[L^{2-\alpha}/T]$, S is the slope along flow direction [-], and β is an empirical constant for surface erosion, where $\beta = 1.68$ (Papanicolaou et al., 2015, Yan et al., 2019)."

Yan, Q., Le, P. V. V., Woo, D. K., Hou, T., Filley, T. R. and Kumar, P.: 3-D modeling of the co-evolution of landscape and soil organic carbon, Water Resour. Res., doi:10.1029/2018WR023634, 2019.

Papanicolaou, A. N., Wacha, K. M., Abban, B. K., Wilson, C. G., Hatfield, J. L., Stanier, C. O. and Filley, T. R.: From soilscapes to landscapes: A landscape-oriented approach to simulate soil organic carbon dynamics in intensively managed landscapes, J. Geophys. Res. Biogeosciences, 120, 979–988, doi:10.1002/2014JG002802, 2015.

6) the parameter a has incorrect units (must be L^2 since the product of a and curvature (units of 1/L) results in units of L).

Our response:

We agree and thank the reviewer's comment. We have revised the unit as "m^2" in Table 1.

7) In Fig. 7, what is the meaning of negative soil transport rate?

Our response:

The negative value means erosion, the higher value means faster rate. We have revised the caption in the updated manuscript and shown below:

"Positive values of transport rate represent deposition, and negative values represent erosion."

==

Second reviwer:

I have read the manuscript, "Hybrid data-model-based mapping of soil thickness in a mountainous watershed" by Yan and colleagues. The authors present a new approach to predicting soil thickness that utilizes the strengths of both numerical and empirical relationships within a portion of the East River, CO watershed. The new data presented here are 78 auger and 54 CPT measurements for 78 locations across two aspects. Their work produces a high-resolution (0.5 m) map of soil thickness, production rates, and transport rates for the two dominate aspects. I found this paper a pleasure to read. I thought it was interesting and provided a creative approach for predicting soil thickness where other approaches have limitations. This document is well written and has a logical flow that is easy to follow; however there are sections that could use more clarification to strengthen the approach and conclusions. Overall, the work was of good quality and falls within the scope of Earth Surface Dynamics target audience. Below I have provided a brief list of major and minor comments to the manuscript. In addition, I have submitted a PDF with a more complete and thorough in-text comments.

Our response:

We thank the reviewer for constructive comments and suggestions on this study. We address the comments one-by-one below.

Major Comments:

Methods clarity- **Though I generally understood how your models work, it was difficult to follow the step-by-step methods (i.e., when each variable/equation is used). Could you more explicitly describe what equations (EQ1-10) and all the necessary variables (7) that the reader would need to use in your approach? At the present, I cannot tell if you calculate your 7 variables using OAT or your model, or if they already have been determine in a past study. Please clarify. Lastly, I would recommend making a diagram in the supplementary information that highlights the workflow and points to the exact equations and variables that are mentioned within the text.**

Our response:

We calculate the values of our seven *parameters* (not variables) by calibrating our models with field sampling data. We used a range of parameter values based on the literature for the sensitively analysis. The OAT method is used for sensitivity analysis of the seven parameters. We stated the following sentence in the methodology section in our original submission: "*We also introduce the Morris one-step-at-a-time (OAT) method for a sensitivity analysis of parameters used in the hybrid model*" and "*Given the uncertainty of the input parameters, we applied the Morris OAT method to quantify parameter sensitivity (Campolongo, et al. 2007; Morris, 1991).*" We add the following sentence in the Methodology section in the revised manuscript:

"*In this hybrid model, seven parameters (Table 1) need to be calibrated for a specific hillslope area.*"

A diagram is included in the supplementary information. The following sentence is added in the Methodology section in the revised manuscript:

"*A diagram that highlights the workflow is shown in Figure S1.*"

We also reorganized the the structure of the methodology section to make the organization clear and easier to follow:

2.1 Hybrid modeling approach:
- 2.1.1 Mass conservation method
- 2.1.2 An empirical approach for depositional areas
- 2.1.3 Investigation of the LiDAR DEM smoothing range for curvature
- 2.1.4 Combine the mass conservation method with the empirical method

2.2 Sensitivity analysis of the model parameters

2.3 Random Forest regression

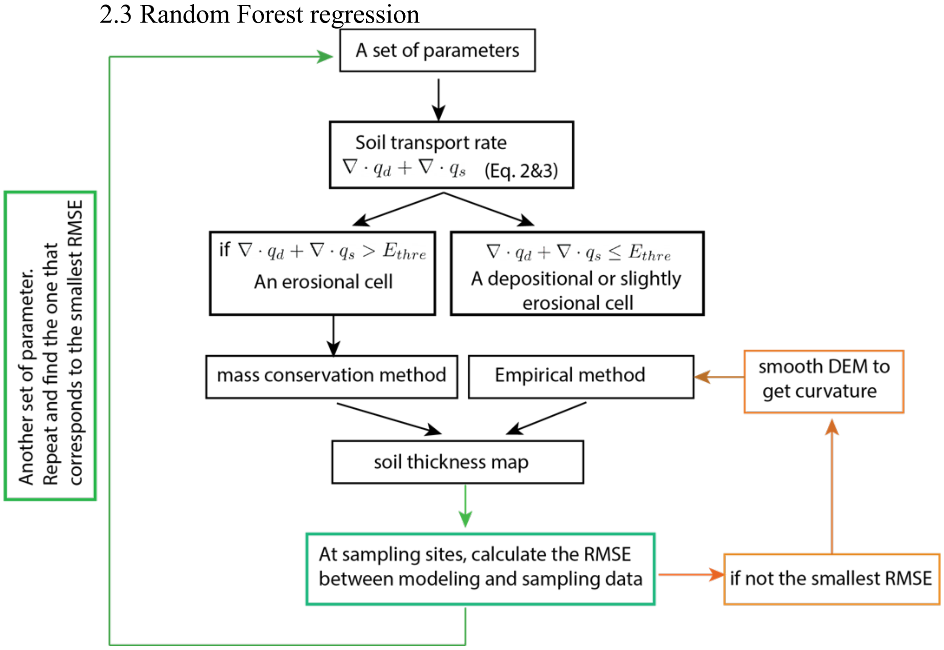


Figure S1: The workflow of the hybrid method.

***Smoothing and resampling grid size-* The different methods for smoothing the landscape is interesting but it's still unclear if it's extremely important for your study and possibly removes the focus away from the main findings. Curvature can be calculated at any resolution but what I am gathering from Sections 2.3 and 4.1 is that the authors want the highest resolution with the lowest RMSD, hence why they selected there smoothing over time approach. To my knowledge, I have not seen any studies which smooth elevation data using diffusion equations and since the authors did not mention any previous studies, I am assuming this is new. If it is not, please provide some references. I do have some concerns with this smoothing approach:**

Our response:

The spatial resolution is always 0.5 m for this study. With such a high-resolution, it can cause 'noise' for curvature, therefore, even though with the same resolution (0.5 m), we smooth the topography to calculate curvature. We tried three approaches to smooth the lidar DEM and keep the 0.5 m resolution all the time. We use the original lidar DEM (0.5m) for other calculations such as soil transport and overland flow estimation. To the authors' knowledge, smoothing the elevation using the diffusion equation is new and original in the study. We revise the following sentence in Section 2.3 in the updated manuscript for clarification:

''... To the authors' knowledge, smoothing over time approach is new and original in the study."

We also revised the manuscript to address the smoothing over space technique more clearly:

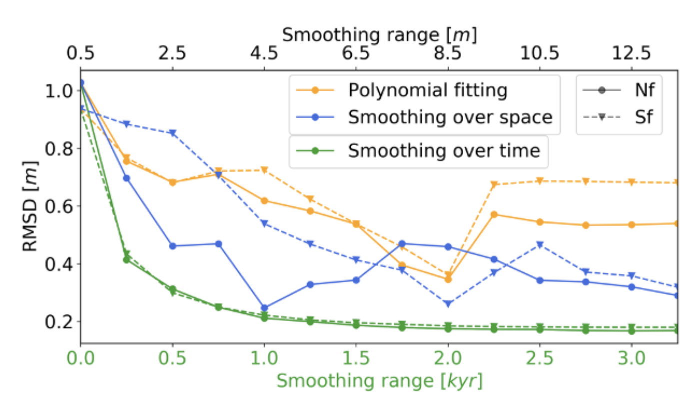
"Smoothing of the DEM over space is done by replacing the value of a 2-D grid cell with the mean value of its surrounding neighbours. The range of its neighbour cells is calculated by $3\Delta x$ *(8 neighbors),* $5\Delta x$ *(24*

neighbors), 7Δx (48 neighbors), …, $(2N+1)\Delta x$ *(*$(2N+1)^2 - 1$ *neighbors) times, respectively; where* Δx *is the resolution (i.e., 0.5 m), and N is an integer; then a moving window replaces the value of every single 2-D grid cell in the 0.5 m lidar."*

- **The smoothing of the DEM uses a linear sediment transport equation when in your hybrid model you utilize non-linear sediment transport (EQ2). If the East River watershed is governed by non-linear sediment transport than your current smoothing equation is inappropriate. Could you provide your reasoning for selecting this equation?**

  Our response:

  We thank the reviewer's comment on this issue and agree that smoothing DEM over time should use the non-linear sediment transport equation. We replot figure 2, and the results are very similar as the one which was achieved by using the linear sediment transport equation. The reason could be that the study area has relatively low local relief. The non-linear sediment transport equation will show very similar results as linear sediment transport equation usually unless the it is a high gradient region. And we chose the nonlinear equation for this study because this equation is commonly used for mountainous areas as we stated in the text: "O*n steep slopes, the following nonlinear slope-dependent transport law is often used for topographic analysis and numerical experiments and has been successfully demonstrated by field studies and laboratory experiments (Andrews and Bucknam, 1987; Perron, 2011; Roering et al., 1999, 2001)*".



- **What values of soil diffusion coefficient (K) and time-steps are you using? I see that your model calculated K value but this happens after the original smoothing occurred. Could you clarify?**

  Our response:

  To smooth the lidar DEM, time step is 1 year. The goal of smoothing DEM is to find the curvature, however, when we calculate the soil thickness, the K is recalculated for the mass conservation method after calibrating to the field data. We use the K value which corresponds to the smallest RMSE (see the work flow above). The curvature is used for the imperial relationship from Patton's method, and time step are independent from the hybrid model. We add the following sentences in section 2.3 in the revised manuscript:

  "The smoothed DEM is for calculating curvature used in the empirical method only, and the rest of all other calculations still use the original lidar DEM as the input."

  We include the following in the caption of Figure 2:

  "… For smoothing over time approach, the time step is one year; the diffusion coefficient, Kd, is 1.1×10^{-3} *m2/yr and* 1.8×10^{-3} *for the north-facing and south-facing hillslopes, respectively …"*

  in the caption of figure S5, we also include:

  "The time step is 1 yr, and the diffusion coefficient is 1.1×10^{-3} *m2/yr for the north-facing hillslope and* 1.8×10^{-3} *m2/yr for the south-facing hillslope."*

- **Though this approach is interesting, I believe it would introduce more uncertainty and unnecessary complexity into your elevation data. For instance, if you were to propagate the error with every time-step (error in original DEM and K) the uncertainty would be much larger than if you were to resample or use a smoothing window. If you were to propagate the error in all your smoothing methods and provide that uncertainty with your Figure 2, it may highlight a more appropriate smoothing method and resolution.**
  <u>Our response:</u>
  We agree that the purpose of smoothing the topography is unclear. Topographic curvature is a key variable. The 0.5 m DEM gives the noise. We need to define the resolution of DEM for curvature. The goal is to determine the optimal resolution to match with the sampling data. The smoothing method is only for the calculation of curvature. We add the following in section 4.1 about the purpose of the smoothing:

  > *"The topographic curvature is the key variable for estimating the soil thickness for the empirical approach. However, curvature is an inherently resolution-dependent topographic feature that is derived from a DEM. A 0.5 m DEM can provide 'noises' for the results of curvature. The goal here is to determine the optimal DEM resolution for curvature to match with the sampling data, and the smoothing methods provided here is only for the calculation of curvature not the equations in the mass conservation method."*

***Sensitivity analysis*- I have read through Section 2.4 and Results 4.2 several times, but I am still having a difficult time wrapping my head around the 7 variables, associated uncertainties, and subsequent sensitivity. Below are my two major questions. More clarification would be greatly appreciated. Lastly, it would benefit the general audience who may not have much expertise in the OAT method (such as myself) to provide a brief explanation on how to interpret Figure 3.**

<u>Our response:</u>

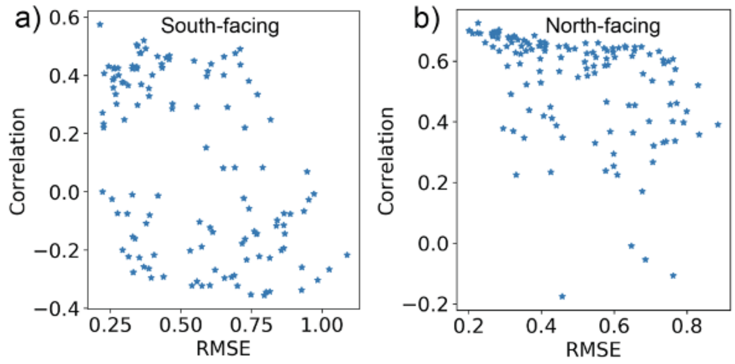
We thank the reviewer's suggestion and revised section 4.2 as following:

> *"We apply the Morris OAT method to investigate the global sensitivity of the seven parameters (Table 1) in the hybrid model. For each parameter, we calculate the "absolute of the mean elementary effect,"* $|\mu|$*, in that the higher number represents higher importance; and the standard deviation of the elementary effect,* σ*, which represent the nonlinearity effect or interactions with other parameters (Fig. 3). Each dot represents an evaluation of one parameter at one sampling site. In general, the parameters in the mass-conservation model have higher* $|\mu|$ *values, meaning that they have more significant impact on soil thickness than the parameters in the empirical model. The diffusion coefficient,* K_d*, is the most important factor (high* $|\mu|$ *value) and high nonlinearity (*σ*) and thus should be carefully calibrated. It represents the soil diffusive-like process such as soil creeping and biogenic activities. The normalized soil depth (*h_o*) is also has higher | high* $|\mu|$ *value, which suggests that it is a very important factor because, but more linear than* K_d *due to relatively small* σ*. These imply that on surface of a soil layer, the diffusive process is the most important transport mechanism for hillslope soil erosion rather than the soil erosion from overland flow (Dietrich et al., 1995; Nicótina et al., 2011; Roering et al., 1999, 2001); and at the bottom of the soil layer, the normalized soil depth is the most important parameter for estimating the soil production rate. The two parameters from the empirical method,* α *and* $\bar{h}$*, are used for soil depositional areas. The sensitivity of* α *and* $\bar{h}$ *are nearly linear (since □ is close to zero), but when the soil thickness reaches."*

- **Did you personally calculate values, acquire from prior studies, or did your model generate them?**
  <u>Our response:</u>
  We thank the reviewer's comments on this issue. We calibrate the seven parameters (not variables) by comparing with the field sampling data. We first provide a range of values for each parameter by using a model named iTOUGH2, then run the model with each set of 7 parameters. The root-minimum-square-error (RMSE) between sampling and modeling are provided. These seven parameters used for the two hillslopes are obtained by comparing model soil thickness with the sampling data and using the to select

the corresponding set of parameters. In Fig. S6 (shown below), each dot corresponds to one set of parameters. We add the following in the revised manuscript:

> "*Here, we use iTOUGH2 (Wainwright et al., 2014) to generate sets of parameters and then sample……*"



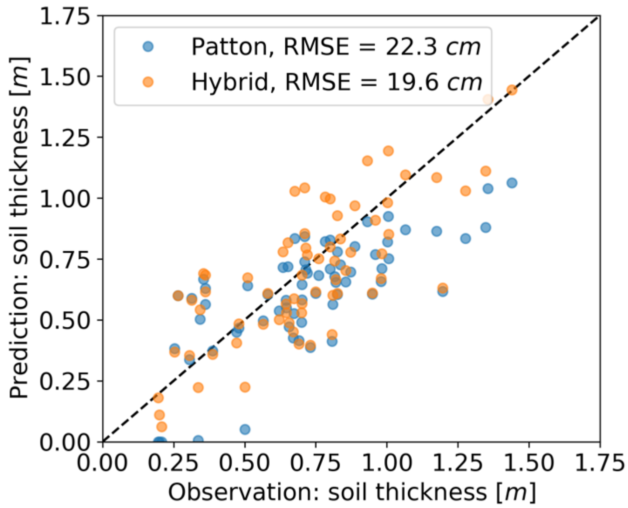
- **When you apply the OAT method to determine sensitivity, was this just for the 0.5 DEM smoothed with time?**
  <u>*Our response:*</u>
  Yes. And again, the smoothed DEM is for calculating curvature only. The sensitivity analysis is performed for the specific study site, so we choose the curvature whichever provides the minimum error between the modeling and sampled data.

Minor Comments:

Model comparison- You predict soil thickness through your hybrid approach and the random forest approach but why not compare it with the components of your model (i.e., just predicting soil thickness using the conservation of mass models and the Patton et al. method)? In the methods you nicely lay out their limitations but you could also demonstrate it. I am particularly interested in how your model will compare with the Patton et al. method because, like your model, it can determine soil thickness across the full topography. By adding a direct comparison you might be able to see additional pros and cons of the models. At the moment there are some clear benefits of your model and worth highlighting such as: you can account for the full landscape where the conservation of mass equations cannot, the Patton et al method is limited to a 5 m resolution, and your model can determine soil production and transport rates.
<u>*Our response:*</u>

I appreciate for this suggestion. The comparison is shown below. We include the comparison between our approach and Patton's approach in the supplementary information. At a thin soil layer where there is usually a divergent topography, Patton's method can generate a negative value, and we forced it to be zero. As we stated in Section 2.2, "*the negative soil thickness values predicted with this method [empirical method] can be compensated for by using the mass-conservation method.*" We included the following in Section 4.4 in the revised manuscript:

> *"This hybrid method also provides higher accuracy than Patton's method in this study site, particularly at very thin or thicker soil layers, because of introducing the mass conservation method and taking parameter 'a' as an independent parameter. (Fig. S7)."*



Model validation- **I have no doubt that your hybrid approach is appropriate for other locations, given your reasonable results, but this has not been actually tested. A cross site comparison would be beneficial, specifically in sites that the Patton et al. approach is limited (watersheds with broad distributions of curvatures and available data) (i.e., Gordon Gulch, CO; Coos Bay, OR; Marshal Gulch, AZ). This would validate your models versatility and provide an additional comparison between the models.**

Our response:

We appreciate the reviewer's suggestion. In this study, we focus on the approaches of generating a soil thickness map and the analysis of the parameters within the hybrid model. This approach requires building a hydrological model. In the next paper, we will apply this model to other sites for comparison.

Defining- **Many of your symbols (i.e., E_{thre}) and specialized vocabulary (i.e., curvature) are missing definitions. Geomorphology is becoming increasingly interdisciplinary. Please insure you have clearly defined the words and the equations used. See PDF for in-text examples.**

Our response:

We thank the reviewer's recommendation. The definitions are added accordingly in the main text and listed below:

> *"Many studies have used curvature—defined as the second order derivative of elevation—as an empirical proxy for soil thickness"*
>
> ……
>
> *"where the threshold, E_{thre}, is a condition of the soil erosion rate and equal or larger than zero value. If $\nabla \cdot q_d + \nabla \cdot q_s > E_{thre}$ at a 2-D grid cell, then this cell must be an erosional site; if $\nabla \cdot q_d + \nabla \cdot q_s \leq E_{thre}$, then this cell can be either a depositional site (if $\nabla \cdot q_d + \nabla \cdot q_s \leq 0$) or a slightly erosional site (if $0 < \nabla \cdot q_d + \nabla \cdot q_s \leq E_{thre}$). In most of areas, a divergent topography corresponds to erosional areas and vice versa for depositional areas. But here we use the transport rate instead of the curvature as the criteria to choose between the two methods because there are possibly sites which are convergent but erosional where overland flow erosion is stronger than the diffusive deposition. In other words, areas where $\nabla \cdot q_d + \nabla \cdot q_s \leq 0$, it must be a convergent area and undergoing deposition, but if it is a convergent area, it is unnecessary to meet $\nabla \cdot q_d + \nabla \cdot q_s \leq 0$. Also, we assign $E_{thre} \geq 0$ instead of equal to zero, aiming to provide a more flexibility to switch between the two methods. Overall, E_{thre} is supposed to be very close to zero."*

***Figures*- Overall, the figures are helpful to understand and progress the reader through the manuscript; however, minor edits will greatly benefit their readability. Please see comments in PDF for figure comments.**
Our response:

We thank you for your comments. We respond to them one-by-one as listed below.

The supplement to comments:

Page 1 title. Your research is more of a "hillslope" study, not so much a "watershed" study.
Our response:

OK, change the title to:

> "*Hybrid data-model-based mapping of soil thickness in mountainous hillslopes*"

Line 12: "two aspects"
Our response:

OK, the sentence is replaced as:

> *"We apply this model to two aspects of hillslopes (southwest- and northeast-facing, respectively)"*

Line 13: **How? I do agree that it shows it has versatilely across aspects but not necessary across to other landscape.**
Our response:

We compare the model results and field sample results in Figure 4. In this study, we focuse on the East River Watershed. In our future work, we'll expend to other sites. OK, the following is deleted as suggested by the reviewer:

> *"that validates the effectiveness of the model".*

Line 21: Seven parameters dose not seem trivial.
Our response:

Whether seven parameters is trivial or significant is subjective. Climate and biogeochemical models can easily have dozens to hundreds of parameters. To avoid the confusion, we delete 'only', and revised the sentence as:

> *"With seven parameters in total for calibration,……"*

Line 22: After reading through your paper, I believe that your model is robust and would likely do a great job in other watersheds; however, you never demonstrated this outside your study site.
Our response:

You are right, we will apply our method in other study sites as the future work. We delete the following phrases:

> *"at other study sites"*

Line 23: How many samples are required? Were you able to calculate the minimum you would need to determine this?
Our response:

With seven parameters, at least seven samples are required, which needs to cover a wide range of curvature values. In reality, 14, 21, or even more would improve the accuracy of the prediction. We are reluctant to provide a specific number at this point.

Line 33: A citation would be good here. See below:
Carvalhais, N., Forkel, M., Khomik, M., Bellarby, J., Jung, M., Migliavacca, M., ... & Reichstein, M. (2014). Global covariation of carbon turnover times with climate in terrestrial ecosystems. Nature, 514(7521), 213-217.

Pelletier, J. D., Broxton, P. D., Hazenberg, P., Zeng, X., Troch, P. A., Niu, G. Y., ... & Gochis, D. (2016). A gridded global data set of soil, intact regolith, and sedimentary deposit thicknesses for regional and global land surface modeling. Journal of Advances in Modeling Earth Systems, 8(1), 41-65.

Fan, Y., Miguez-Macho, G., Jobbágy, E. G., Jackson, R. B., & Otero-Casal, C. (2017). Hydrologic regulation of plant rooting depth. Proceedings of the National Academy of Sciences, 114(40), 10572-10577.

Patton, N. R., Lohse, K. A., Seyfried, M. S., Godsey, S. E., & Parsons, S. B. (2019). Topographic controls of soil organic carbon on soil-mantled landscapes. Scientific reports, 9(1), 1-15.
<u>*Our response:*</u>

OK, added.

Line 36: What is local? Is this at the pedon or watershed scale?
<u>*Our response:*</u>

We agree that local is a vague word. We replace it with "hillslopes" in the revised manuscript.

Line 45: Do you mean "bell-shaped" or the "humped" soil production function?
<u>*Our response:*</u>

Yes, it has been corrected accordingly.

Line 50: Hillslopes have both depositional and erosional portions (ie. hollows/swales/valleys and ridges/noses/crest, respectively) and they are not "low-land areas" The presence of both is not a good indicator.
<u>*Our response:*</u>

We agree with this comment. "i.e., a low-land area" has been deleted.

A good indicator (when slopes are low and follow the linear soil transport) is curvature (C) where under a conservation of mass and steady-state soil thickness where C<0 is convex and C >0 is concave.
I would suggest removing the bit about Lidar and placing it lower in the paper and focus on why these convergent (depositional) areas are such a problem with the numerical equations.
<u>*Our response:*</u>

The sentence has been revised as:

> *"In areas where the topography is a mixture of divergent (mostly erosional) and convergent (mostly depositional) zones, which are commonly revealed in a Light Detection and Ranging (LiDAR) based digital elevation model (DEM) for high spatial-resolution modeling, these mechanistic models fail to capture the soil thickness distribution."*

Line 52: "...fail to capture the *full* soil thickness distribution."
<u>*Our response:*</u>

OK, "full" is added in the revised manuscript.

Line 53: Hybrid approach with what? At this point of the paper you have not mentioned the empirical approach (which you ultimately combine with the numerical approach to create your hybrid model). I would move this sentence after you go through the empirical soil thickness models.
<u>*Our response:*</u>

We added the following phrase in the sentence:

> *"a hybrid approach that couples mechanistic and empirical methods"*

Line 55: You need to define curvature and which one you are using.
<u>*Our response:*</u>

We added the following in the revised manuscript:

"Many studies have used curvature—defined as the second order derivative of elevation—……"

Line 58: This is already been said above.
Our response:

OK, deleted.

Line 58: delete "studies focusing on surface curvature may not be sufficient for predicting soil thickness. For example"
Our response:

OK, deleted.

Line 64: This is an important part but an even more critical portion that is missing is that high-resolution elevation maps are not always available.
Our response:

We agree with this point, but this statement is not an issue for this study site, and this work focuses on lidar DEM.

Line 71: How are you sure? Have you tried running this model for other locations?
Our response:

The soil mass conservation equation is a process-based model, and has been applied in many other studies as stated in the Introduction and Methodology part. The imperial approach is based on Patton's work, and has also been proved to be applicable to various sites.

Line 72: In your methodology headers, could you rename them to reflect what you put here? This will make it easier for the reader to navigate through the paper.
Our response:

We thank the reviewer's suggestion. The sentence has been revised as:

> *"In the methodology section, we introduce our hybrid modeling approach and relevant concepts such as curvature calculation with different DEM smoothing methods, sensitivity analysis of model parameters, and a machine-learning approach as a comparison with the hybrid model."*

Line 81: One thing that is not clear to me, are you predicting soil thickness on just the hillslopes or on the flood plain /river bed? If it is the latter, it will be difficult to differentiate between old mobile alluvium and mobile regolith without digging a soil pit to see it. Also the Patton et al. (2018) method did not try to calculate soil thickness in areas effected by streams or in colluvium where defining soil thickness is difficult. Please be explicit where you are doing your model.
Our response:

We agree with the reviewer, and this study focus on hillslopes only. The floodplains are not included as shown in Figure 7 and 8. Figure 4 has been revised to eliminate the floodplain area for consistency.

I am still having a difficult time seeing how/when these equations (both conservation of mass and empirical equations) are being utilized in your model. It would be nice to see (maybe in the supplemental information) the work flow of your model.
Our response:

We appreciate this suggestion. A workflow has been included in the supplementary material, and the figure has been shown above.

Page 4, Line 2: This critical gradient was used in Oregon but is it appropriate here? Should this be a site specific number based on the properties of the soil profile/hillslope? How did you end up deciding if this is the correct value? Did you run a sensitivity analysis on this?
Our response:

The critical gradient, Sc, is a site-specific number indeed. However, we do not have such value in this study site.

However, Sc should be a very large number in that 1.25 is appropriate for mountainous area because this value represents the fact that if the slope is very small, then the diffusion flux still has a linear relationship with slope, but if the topographic slope is very steep, then the diffusion flux is faster. The critical gradient usually play a role near the ridges, however, in this study site, it is very far away from a watershed ridge, so the value, as long as it is reasonably high, should have little impact in the results.

Page 5, line 36: Patton et al, 2018.
<u>*Our response:*</u>

OK, revised. Sorry for this typo.

Line 41: Are you calculating curvature in ArcGIS? If so you might want to consider dividing your curvature values by -100 such that negative values represent convex and positive values represent concave areas. Arc use the Zevenberger and Thorne (1987) and Moore et al. (1991) equations which determines curvature as the change in the slope (in percent rather than the actual gradient) in all direction, plus it reverses the sign (i.e. using the negative curvature convention). Please provide the method.
<u>*Our response:*</u>

We did not use ArcGIS to calculate curvature. Instead, I derive numerical solutions following the mathematical definition of curvature by ourselves. My polynomial fitting method is similar to what ArcGIS does for curvature, but my algorithm (written in Python) allows user to specify the averaging window. (https://desktop.arcgis.com/en/arcmap/10.3/tools/spatial-analyst-toolbox/how-curvature-works.htm). $\nabla^2\eta = \nabla \cdot \nabla\eta = \frac{\partial^2\eta}{\partial x^2} + \frac{\partial^2\eta}{\partial y^2}$, to solve this equation numerically, we discretize the equation by using Taylor series expansion and get: $\frac{\partial^2\eta}{\partial x^2} + \frac{\partial^2\eta}{\partial y^2} = \frac{\eta_{i+1,j} - 2\eta_{i,j} + \eta_{i-1,j}}{\Delta x^2} + \frac{\eta_{i,j+1} - 2\eta_{i,j} + \eta_{i,j-1}}{\Delta y^2} + O(\Delta y^2) + O(\Delta x^2)$. More details of the numerical solution can be found at:

> Yan, Q., Le, P. V. V, Woo, D. K., Hou, T., Filley, T. and Kumar, P.: Three-Dimensional Modeling of the Coevolution of Landscape and Soil Organic Carbon, *Water Resour. Res*., 55(2), 1218–1241, doi:10.1029/2018WR023634, 2019.

If this represents curvature than this is not the same symbol you use in line 54- (∇ · ∇η)
<u>*Our response:*</u>

Thanks for this suggestion. We replaced $\nabla^2\eta$ with $\nabla \cdot \nabla\eta$ to keep the consistency in the text.

Line 42: Make sure you define *a*. This was originally calculated at a 5 m grid. The units are m^2 please edit Table 1.
<u>*Our response:*</u>

OK, the units are revised. We revised the sentence in the updated manuscript:

> *"a is a constant value which is determined by having a negative linear relationship with the standard deviation of curvature."*

Line 43-44: Was your values of *a* similar to the value when you plot curvature against soil thickness (i.e. the slope of that relationship)? It would be worth checking that your modeled and measured *a* values converged to the same values.
<u>*Our response:*</u>

We take a as the slope of that relationship. It is the same thing. We did not calculate a based on Patton's method because in Patton's method, a is calculated at a 5 m resolution DEM. In this work, our curvature is calculated by having DEM smoothed over time. We stated in the text that:

"*In our model, we take a as an independent parameter instead of being calculated based on curvature, which adds one more degree of freedom to the model."*

Line 46-48: In essence our approach was founded in numerical equations like the ones you listed above. The contribution of our study is that these predictions could be systematically extrapolated from convexed into concaved areas which were difficult to do prior. Though negative values are initially obtained in our model, these values are reclassified as 0 m before the final product.

Our response:

We thank the reviewer's clarification. In our hybrid model, there is no need to reclassify the negative values.

Line 51: **Define**
Our response:

OK. We have responded this comment above. Here we show the revised text again:

"where the threshold, E_{thre}, is a condition for the soil erosion rate, and therefore is an equal or less than zero value. In most of areas, divergent topography corresponds to erosional areas and vice versa for depositional areas. However, there would still be convergent but erosional areas where overland flow erosion is stronger than diffusive deposition. In other words, areas where $\nabla \cdot q_d + \nabla \cdot q_s < E_{thre}$, it must be a convergent area, but if it is a convergent area, it is unnecessary to meet $\nabla \cdot q_d + \nabla \cdot q_s < E_{thre}$."

Line 52: The equation makes perfect sense but how are you defining these areas on the map? Please explain. Why not define them with Curvature (C) (ie. C < 0 - conservation of mass ; C > 0 - Patton's method)?
Our response:

We calculate soil transport rate at each 2-D grid cell spatially by using $\nabla \cdot q_d + \nabla \cdot q_s$. Then, we apply the criteria to each grid cell based on the equation: $\{\nabla \cdot q_d + \nabla \cdot q_s > E_{thre}, mass\ conservation\ \nabla \cdot q_d + \nabla \cdot q_s < E_{thre}, Patton's\ method$. We apply different method based on the criteria to each grid cell to calculate soil thickness. The detailed explination of the equation $\nabla \cdot q_d + \nabla \cdot q_s$ can be found at the following citation:

Yan, Q., Le, P. V. V, Woo, D. K., Hou, T., Filley, T. and Kumar, P.: Three-Dimensional Modeling of the Coevolution of Landscape and Soil Organic Carbon, *Water Resour. Res*., 55(2), 1218–1241, doi:10.1029/2018WR023634, 2019.

The reason we do not use curvature value to choose each method is as stated above:

"However, there would still be convergent but erosional areas where overland flow erosion is stronger than diffusive deposition. In other words, areas where $\nabla \cdot q_d + \nabla \cdot q_s < E_{thre}$, it must be a convergent area, but if it is a convergent area, it is unnecessary to meet $\nabla \cdot q_d + \nabla \cdot q_s < E_{thre}$."

Where is equation 9?
Our response:

It is a typo. Equation 10 should have been equation 9. We have fixed this in the revised manuscript.

Line 53: Bring in Lidar section here
Our response:

OK, added 'LiDAR DEM' in the text.

Line 54: Which one (EQ. 8, 9, or 10)?
Our response:

OK, we added (Eq. 8) in the text.

Line 55: Seems circular and/or redundant.
Our response:

OK. We have deleted:

“,and soil thickness is sensitive to the local topographic curvature (Pelletier and Rasmussen, 2009)

Line 60: Has this been done in the past?
Our response:

To the author’s knowledge, it is new and original.

Line 61: Please make clear that you are smoothing the original elevation data then calculating curvature after the DEM has been either smoothed or resampled.
Our response:

We add the following sentence in the revised manuscript:

“*The smoothed DEM is for calculating curvature used in the imperial method only, and the rest of all other calculations still use the original lidar DEM as the input.″*

Line 61: How are you calculating curvature?
Our response:

We have addressed the question above, and we repeat below:

I derive numerical solutions following the mathematical definition of curvature by myself. My polynomial fitting method is similar to what ArcGIS does for curvature, but my algorithm (written in Python) allows user to specify the averaging window. (https://desktop.arcgis.com/en/arcmap/10.3/tools/spatial-analyst-toolbox/how-curvature-works.htm).

$\nabla^2\eta = \nabla \cdot \nabla\eta = \frac{\partial^2\eta}{\partial x^2} + \frac{\partial^2\eta}{\partial y^2}$, to solve this equation numerically, we discretize the equation by using Taylor series expansion and get: $\frac{\partial^2\eta}{\partial x^2} + \frac{\partial^2\eta}{\partial y^2} = \frac{\eta_{i+1,j} - 2\eta_{i,j} + \eta_{i-1,j}}{\Delta x^2} + \frac{\eta_{i,j+1} - 2\eta_{i,j} + \eta_{i,j-1}}{\Delta y^2} + O(\Delta y^2) + O(\Delta x^2)$. More details of the numerical solution can be found at:

Yan, Q., Le, P. V. V, Woo, D. K., Hou, T., Filley, T. and Kumar, P.: Three-Dimensional Modeling of the Coevolution of Landscape and Soil Organic Carbon, *Water Resour. Res*., 55(2), 1218–1241, doi:10.1029/2018WR023634, 2019.

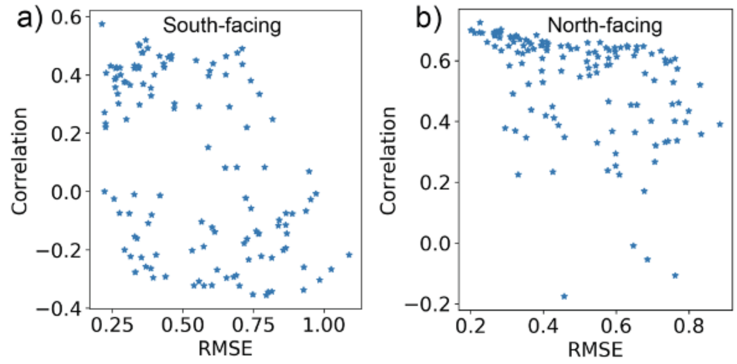
Line 68-70: Is this not assuming linear sediment transport? You mention above that non-linear sediment transport (EQ 2) is the correct mechanism driving landscape change. Wouldn't this mean this equation is inappropriate?
Our response:

As we respond above, the non-linear sediment transport law is for the simulation. We use the linear equation as a way to smooth the DEM. The reason we choose the linear equation here is that in this equation, the elevation changing rate is the product of curvature and a constant coefficient. The goal of using the linear equation is to calculate curvature, which is used for the empirical method only.

Line 72: Did you calculate these values and their associated uncertainty?
Our response:

As answered above. We calculate the values by calibrating with the sampling points.

Fig S7.



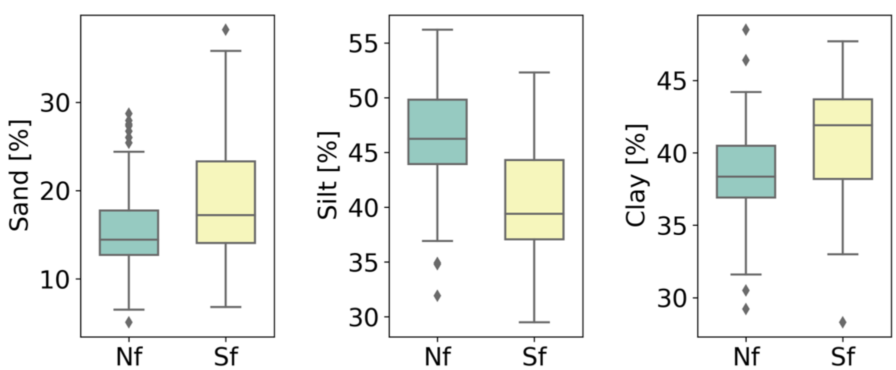
Page 8, Line 14: The only soil characteristic I see in your manuscript is soil thickness. Are there any differences in the soil types and/or the physical and chemical characteristics between aspects? If so, it would be worth mentioning in your discussion on soil thickness, transport and production.

<u>Our response:</u>

We thank the reviewer's suggestion. We replaced 'characteristics' to 'thicknesses' in the main text.

We have soil texture information along with the soil thickness data. The surface soil texture values added in the supplementary information (Fig. S8). We added the following phase into the Section 4.5 in the main text:

"*Moreover, … and more sand material from the soil texture (Fig. S8) …….*"



Line 15: Are your models appropriate here?

<u>Our response:</u>

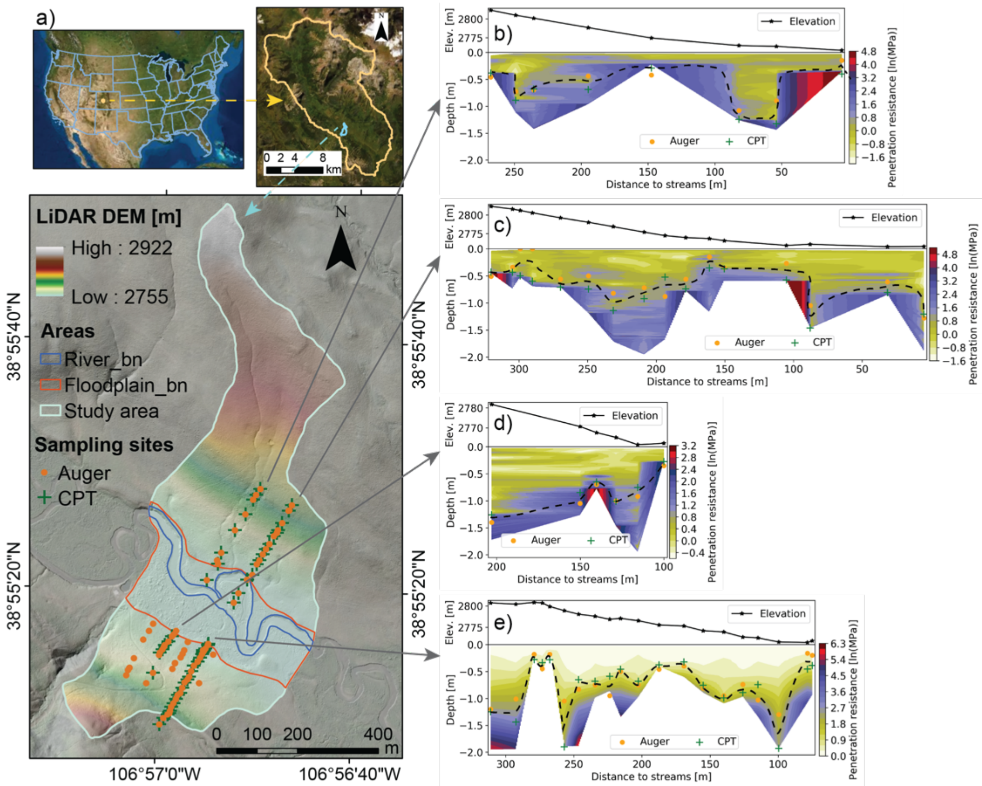
As responded earlier, our model is not appropriate for calculating soil thickness within floodplains because the formation of floodplains is mostly controlled by fluvial processes which are associated with channel migrating and flooding events. We have revised the texts properly.

Line 37-38 In Figure 1b-e could you provide a dashed line of the "preferred" soil depth along the transect?

<u>Our response:</u>

Thanks for this suggestion. The dashed lines have been added into the figure. We also include the following sentence in the caption:

"The dashed lines are estimated soil thickness by the average of the auger and CPT measurements."

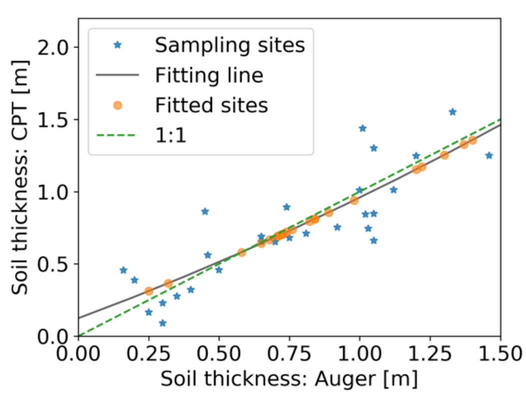


Line 38-40: Could you provide a slope of the line (such that a slope of 1 is a perfect agreement between Auger and CPT methods) and some stats on that relationship?

Our response:

OK. We have included the 1:1 ratio and provides the RMSE and correlation between the two methods in the captions.

"Correlation = 0.86, Root-minimum-square-error = 0.20 m."



Line 43: Can panels B-E have their x-axis' lined up with similar interval size and length? Also the color ramp for the Penetration resistance should be the same for each panel (ie. -1.6 to 6.3)

Our response:

We thank the reviewer's suggestion. The x-axis has been replotted as shown above. Regarding the color ramp, the goal is to separate between soil layer and bedrock/weathering layer. Each cross-section is independent, and the best way to separate soil layer from bedrock/weathering layer is to plot the maximum range of each cross-section instead of use an uniform maximum-to-minimum range for all panels.

Line 43: Consider using a black and white dot for the auger and CPT measurements. It is sometimes difficult to see in panels b-e. Alternatively, you have the locations in panel a for the CPT and auger measurements, maybe just provide a dashed line for the "preferred" soil thickness depth.
<u>*Our response:*</u>

We agree with reviewer and have provided 'preferred' dash line for soil thickness depth as shown above,

Line 44-45: The extrapolation method in panel e seems different from that in panels b-d. Panel e appears to be more simplistic and easier to follow. Could you explain the extrapolation approach and provide your preferred method?

<u>*Our response:*</u>

We use the same method, kriging, as the interpolation, not extrapolation, of the CPT data. The reason that panel e seems different from the rest of panels is because of its own geophysics property, which is out of the authors's control. We add the following phase in the caption:

> '*by using a kriging method.'*

Line 51-54: Seems out of place and peripheral to your study.
<u>*Our response:*</u>

This is important here because our assumption is that the dynamics of soil thickness is at a steady state. This assumption supports the legitimate of mass conservation method (Equation 7).

Line 51-54: It still is uncertain to me if you used the linear or non-linear transport equations. Also, what time steps did you calculate this at?
<u>*Our response:*</u>

We use the non-linear transport equation for the soil formation. The time step is one year. We include the information in the caption of Figure S3. The goal is to show if the soil thickness can reach a steady state indeed or not.

Line 73-77: Has this approach worked in the past? Could you provide some examples where this was the best method? It just seems like its adding extra unnecessary complexity and potentially adding in more error.
<u>*Our response:*</u>

We have respond this question above. To the authors' knowledge, the smoothing over time approach is new and original in this study. We believe that smoothing the DEM reduce the error and reduce the complexity of the topography for calculating curvature.

Line 95, Figure 3: Not clear how to interpret this figure. Should I look at the slope or how the well the regression fits? Please clarify.
<u>*Our response:*</u>
We thank the question, and we have responded this question above.

Figure 4b1: Please provide r2, p-value, and RMSE for both panels b and c.
We thank the reviewer for this suggestion. We include he following in the caption:

> *"The correlation, root-minimum-square-error, and p-value are 0.71m 0.18 m, and $4.2*10^{-4}$; and 0.77, 0.19 m, and $2.32*10^{-10}$ for south-facing and north-facing hillslopes, respectively."*

Figure 4: Did you exclude all the samples in the flood plain? Your sample locations don't match Figure 1 (i.e., Figure 1 has sample locations in the flood plain).
<u>*Our response:*</u>

Yes, we exclude all samples in the floodplain. As stated above, we believe the model is not appropriate to capture the processes for soil layer formation in fluvial environment. We include the following in the caption:

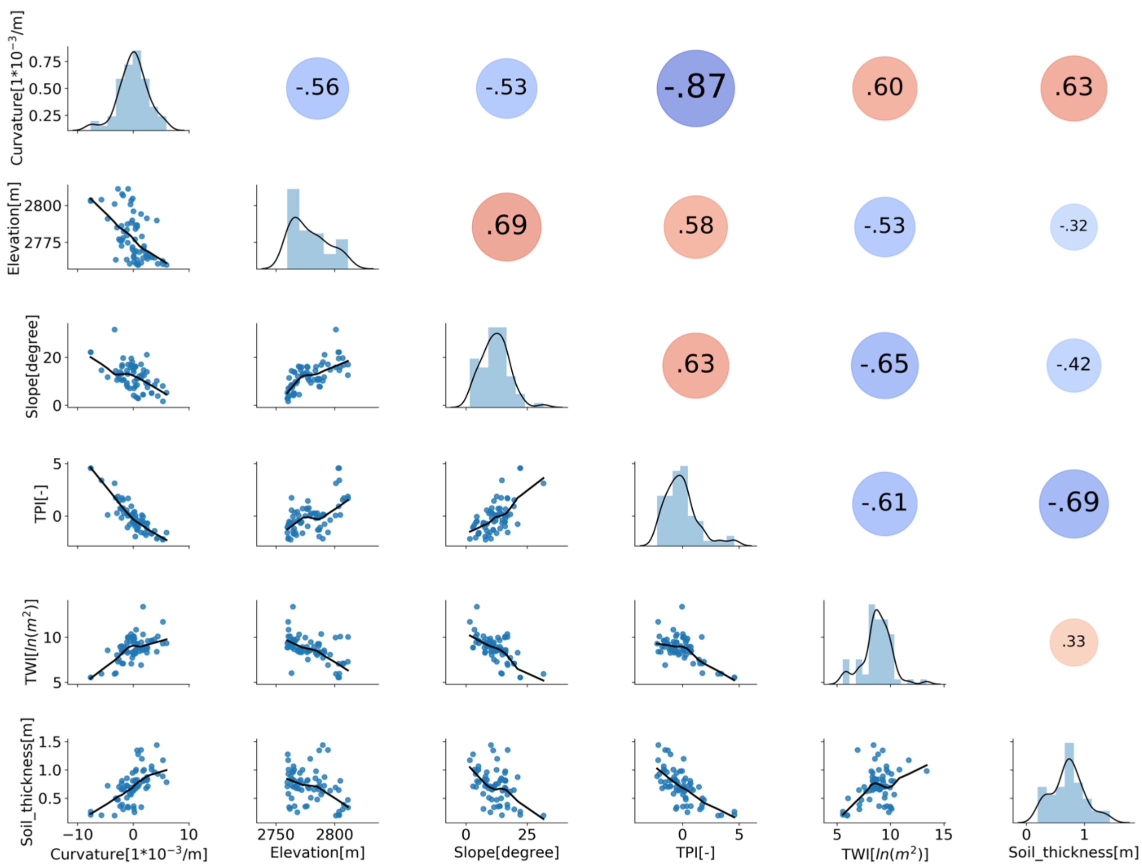
> *"Note that the sampling points in the floodplain zone are excluded because our hybrid model aims to predict the soil thickness in hillslopes."*

Figure 4: Have you considered comparing soil thickness predictions between your hybrid approach and the Patton et al (2018) approach? Highlighting the pros and cons would be good. I realize your model provides much higher resolution than our approach (5m compared to 0.5m) but it would be worth looking into.
<u>Our response:</u>

OK, we included the comparison in the supplementary information, and added the discussion part. We have responded to this question in the Minor Comments section.

Figure 5: Relabel the axis so that they are easier to follow.
<u>Our response:</u>

OK, revised.



Line 64: Is this represented by negative numbers? Please clarify.
<u>Our response:</u>

Yes. We add the following sentence in the caption for Fig. 7 revised manuscript:

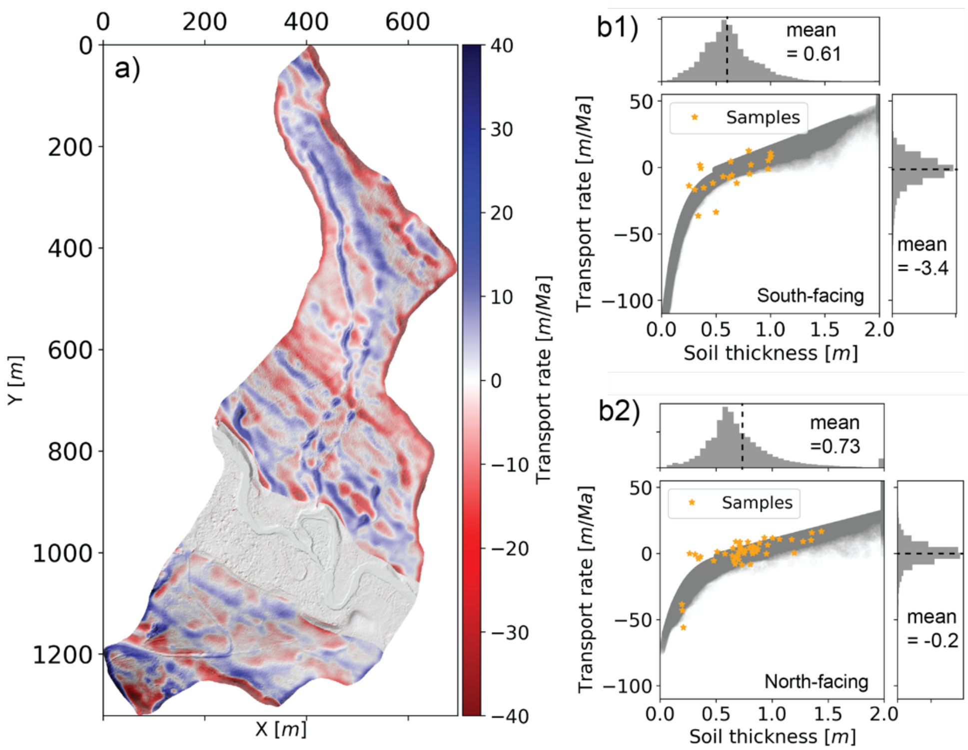
> *"Positive values of transport rate represent deposition, and negative values represent erosion."*

Figure 7: Units.
Shouldn't the x-axis be soil thickness because this is what you measured and the y-axis be transport rate? Does negative transport rates mean deposition and positive mean erosion?
<u>Our response:</u>

The unit of mean value is from the corresponding PDF plot, and the unit of the PDF lot is given in the x-axis. We do not provide the unit here to reduce the redundant information, and also due to the limitation of space in the PDF plots.

As explained above, we add the following sentence in the caption for Fig. 7 revised manuscript:

> *"Positive values of transport rate represent deposition, and negative values represent erosion"*



Line 79: What was your preferred values for bulk densities of the soil and parent material? Also what values did you select for your empirical constants? How did you select them and did they appear reasonable compared to other studies.

<u>*Our response:*</u>

We have bulk density from lab analysis of soil samples. The mean value of the bulk density among our sampling site is 0.948 (g/cm^3). The parent material is shale, as we stated in the section 3, "*the material beneath the soil layer is weathered shale"*.

We provide the bulk density and bedrock density information in Equation 5 in the revised manuscript:

> *"The mean value of the soil bulk density at sampling sites is 0.948 [g/cm3], and the bedrock bulk density for weathered shale from our deep samples is estimated to be 1.26 [g/cm3]."*

The values are obtained by calibrating the model with sampled data. Specifically, we assigned a set of values for each parameter, and run the hybrid model. We compare the results with sample data at the sampling locations. The RMSE between modeled and observed values are provided in Figure S4. In Section 4.3, we stated that:

> *"We use the sampling data from both auger and the CPT to calibrate the seven parameters (Table 1) for the south-facing and north-facing hillslopes separately (Fig. S4). The calibration follows the same sets of parameters generated in the Morris OAT method by increasing each parameter linearly from the minimum value to the maximum in a constant interval."*

We add the following sentences in the Section 4.3:

> *"We compare the soil thickness values between modeling and sample results at sampling locations. The minimum RMSEs between model and sample results of the south-facing and north-facing hillslopes are 0.20 m and 0.195 m, respectively; and the corresponding values of the seven parameters are shown in*

Table 1."

Figure 8: Shouldn't the x-axis be soil thickness because this is what you measured and the y-axis be soil production rate?
<u>*Our response:*</u>

OK, we flipped x- and y-axis in the revised manuscript.

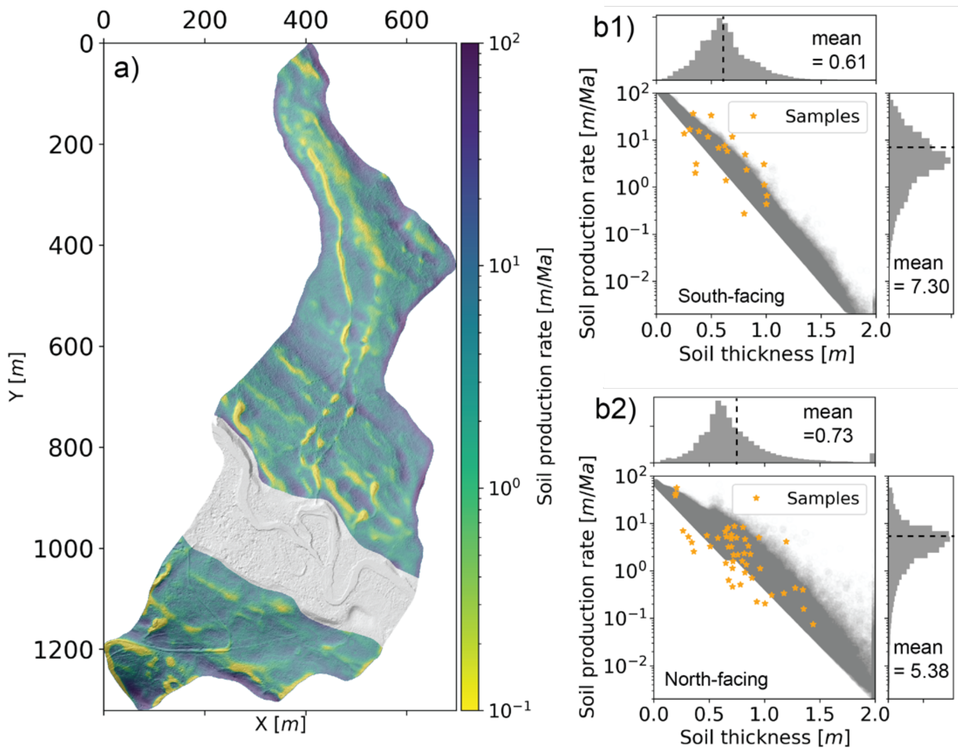


Figure 8: Can you put it in m/Ma or something similar? I'm unsure what you mean by LN and your negative values. Please explain. m/yr doesn't seem right. Past studies have soil production rates are in m/Ma
<u>*Our response:*</u>

OK, the unit has been given as m/Ma. Ln is the natural logarithm. We have used semilogarithmic plot for soil production rate.

Figure 8: All roads or human disturbed areas should be removed post processing (and possibly a buffering area of 5-10m) before the distributions on soil production and transport rates are evaluated.
<u>*Our response:*</u>

We appreciate the reviewer's concern. Our study site is a nature forest preserved area. There is no roads or any other infrastructures in our study site.

Page 18, line 9: (1) To what? (2) In your hybrid model or calculating curvature?
<u>*Our response:*</u>

We revised the sentence in the updated manuscript below:

> *"We found that smoothing lidar DEM over time has a higher efficiency than smoothing it over space to obtain the optimal topographic curvature values, which provides the least error between the modelling results and sampling soil thickness."*

Line 20: This should be moved to the discussion.
<u>*Our response:*</u>

OK, this sentence has been moved to Section 4.5.

Line 21: True but consider acknowledging that your model uses numerical equations that are founded on hillslope transport (crest to footslopes) not fluvial dominated areas (toe-slopes, flood plains).
<u>*Our response:*</u>

OK, we revised the sentences as following in the updated manuscript:

"*it would fail in alluvial depositional sites (i.e., floodplains)*…", and "…*soil thickness prediction in hillslopes.*"

Line 26-27: Are you sure? Have you done any test? Might be worth looking at some other sites to check that your model is getting reasonable soil thicknesses outside your area.
<u>*Our response:*</u>

We agree that we did not any test for other sites. As we answered above, cross-site comparison would be a future work. In this study, we mainly focus on the method and tested the methodology in two mountainous hillslopes. We revise the sentence in the following:

> *"Although the example applications in this paper are at two hillslopes, this hybrid model framework should have little limitation to analyze soil-mantled mountainous hillslopes after calibration with sampling dataset."*

Line 30: Did you calculate these values in your model or did these values come from other studies?
<u>*Our response:*</u>

We have answered this question above and repeat below:

> Specifically, we assigned a set of values for each parameter, and run the hybrid model. We compare the results with sample data at the sampling locations. The RMSE between modeled and observed values are provided in Figure S4. In Section 4.3, we stated that:
>
> > *"We use the sampling data from both auger and the CPT to calibrate the seven parameters (Table 1) for the south-facing and north-facing hillslopes separately (Fig. S4). The calibration follows the same sets of parameters generated in the Morris OAT method by increasing each parameter linearly from the minimum value to the maximum in a constant interval."*
>
> We add the following sentences in the Section 4.3:
>
> > *"We compare the soil thickness values between modeling and sample results at sampling locations. The minimum RMSEs between model and sample results of the south-facing and north-facing hillslopes are 0.20 m and 0.195 m, respectively; and the corresponding values of the seven parameters are shown in Table 1"*