

I wish to thank Yan et al. for their diligent responses to my initial review. I think the paper could make a strong contribution to ESurf with additional major revision.

My initial review was somewhat brief because the manuscript contained too many omissions to be comprehensively reviewed. My sense is that ESurf tends to limit the review of the post-discussion manuscript to a single round. In this case I urge the editors to allow one more round of review if possible given that major omissions remain.

Our Response:

We thank the reviewer's suggestion. We address the following issues one-by-one, and we also revised the manuscript systemically based on these comments.

For example, the paper still contains very limited information on calibration. I alluded to this problem by highlighting a few values (e.g.,  $h_0$ ) that were plainly incorrect in the discussion manuscript. The authors have stated that the  $h_0$  value was a typo in the initial manuscript, but I still have a limited idea of how their reported value for  $h_0$  compares against alternative values when the values of other parameters are modified to best match the calibration data. I would really like to see a figure that demonstrates the fit of the model to calibration data for the entire parameter space. The need to \*see\* the results of the calibration is particularly important when a single measurable (soil thickness in this case) is being used to calibrate a model with more than one or two parameters (7 in this case, more if one includes the 50 cm of uniform soil assumed in the initial condition which should be added to the list of unconstrained parameters requiring calibration). The model may do equally well at reproducing soil thickness data if  $h_0$  and  $P_0$  are both increased, for example, yet if the parameters are not accurate then the model may perform very poorly when applied to a different but very similar location.

Our Response:

We thank the reviewer's comments and present more results regarding the calibration for the soil thickness. We performed a *grid search* approach where the model loops across the parameter space. There are seven parameters in our hybrid mode, but due to parameter  $K_s$  (which is in charge of the overland flow erosion) is not sensitive to soil thickness, we drop this parameter for the grid search. We assign 6 equally distributed space for each parameter. We then calculate the root-minimum-square-error (RMSE) between the simulated and measured soil thickness across the site.

The distribution of the RMSE as a function of the parameters are shown in Figure S7 (each set of parameters corresponds to one RMSE; all the RMSEs are projected into each parameter space as a marginal distribution). The global minimum can be defined as the set of parameter provides the lowest RMSE. Based on Equation 7 and 8, we agree that the parameters we calibrated may not be the 'real' set of parameters. There is indeed a possibility that another set of parameters may provide equally good or better fit for the data. However, the point of this work is to find the best fit for the sampling points and apply it to other unsampled locations in the sample hillslope. In this specific study site, we find a unique set of parameters, which provide the global minimum with the grid search study. Moreover, when the threshold of RMSE decreases, the number of samples with RMSE below this threshold decrease and converges to the global minimum. Based

on these results, we conclude that the non-uniqueness is minimal and we can identify the combination of the parameters that provide the global minimum.

In addition, we have applied a Bayesian approach to estimate the posterior distribution of parameters. We assume the uniform prior distributions between the range that define the parameter space. We find that the maximum a posteriori estimates (Figure S8) also correspond to the global minimum of the RMSE.

Here, we include the range of parameter space on the revised Table 1, add Figure S7 and Figure S8, revise the description of the calibration procedure into the main text, and include the Bayesian approach in the supplementary information as shown below:

*“We use the sampling data from auger and the CPT to calibrate the model parameters (Table 1) for the south-facing and north-facing hillslopes separately. The calibration is performed using a grid search approach where the model loops through the entire parameter space, with 6 evenly distributed values for each parameter. The range of each parameter is based on literature and shown in Table 1. The overland flow coefficient  $K_s$  is not included in this process because it is insensitive to soil thickness (Fig. 3). The range of the remaining 6 parameters is based on literature and shown in Table 1. We created an ensemble of the soil thickness based on  $6^6$  (=46646) combinations of parameter sets. We then calculate the root-mean-square-error (RMSE) between the simulated and measured soil thickness across the site. Each set of parameters corresponds to one RMSE, and the distribution of the RMSE as a function of the parameters are shown in Figure S7. The global minimum can be defined as the set of parameters that provides the lowest RMSE. In this specific study site, we find a unique set of parameters, which provide the global minimum with the grid search study. Moreover, when a threshold of RMSE decreases, the number of samples with RMSE below this threshold decreases and converges to the global minimum. This result indicates the absence of other potential strong minimum, and therefore, we can identify the combination of the parameters that provide the global minimum. In addition, we have applied a Bayesian approach to estimate the posterior distribution of parameters, as well as the maximum a posteriori estimates. We assume the uniform prior distributions between the range that define the parameter space. We find that the maximum a posteriori estimates (Figure S8) also correspond to the global minimum of the RMSE.”*

We also include the following section in the Supplementary Information:

#### **Posterior distribution of the parameters**

We used the Bayesian method to calculate the posterior distribution of the parameters in this hybrid-model (Eqn. 6-8) and estimated the maximum a posteriori (MAP).

The parameter set is defined as a vector  $\theta$ , including six parameters (except for  $K_s$  in Eqn. 3b). We define the prior distribution for  $\theta$ ,  $p(\theta)$ , which is the product of independent uniform distributions between the predefined range parameter for each parameter. We define the measured soil thickness data vector  $\mathbf{z}$  (which is a  $m$ -vector with  $m$  field sampling points:  $\{z_1, z_2, \dots, z_m\}$ ) as the sum of predicted soil thickness and error vector  $\epsilon$ :

$$\mathbf{z} = \mathbf{y} + \boldsymbol{\varepsilon} \quad (S1)$$

where  $\mathbf{y}$  is the predicted soil thickness based on the model and parameter sets:

$$\mathbf{y} = f(\boldsymbol{\theta}). \quad (S2)$$

We assume that each element in the error vector follows a normal distribution with a standard deviation  $\sigma$ . We estimate the standard deviation of the measurement error based on the discrepancy between the auger and CPT measurements.

Using the Bayes' rule, we define the posterior distribution as:

$$p(\boldsymbol{\theta}|\mathbf{z}) \propto p(\mathbf{z}|\boldsymbol{\theta})p(\boldsymbol{\theta}) \quad (S3)$$

Since the error vector is normal, the likelihood  $p(\mathbf{z}|\boldsymbol{\theta})$  is a normal distribution.

To compute the posterior distribution, we use the sampling-resampling scheme (Smith and Gelfand, 1992). We sample the parameter set at each grid following the grid search  $\{\boldsymbol{\theta}^{(1)}, \boldsymbol{\theta}^{(2)}, \dots, \boldsymbol{\theta}^{(N)}\}$ , where  $N = 6^6 = 46656$  and predict the soil thickness at the measurement locations  $\{\mathbf{y}^{(1)}, \mathbf{y}^{(2)}, \dots, \mathbf{y}^{(N)}\}$  where  $\mathbf{y}^{(i)} = f(\boldsymbol{\theta}^{(i)})$ . Note that each vector represents the collection of soil thickness at the measured locations  $\mathbf{y}^{(i)} = \{y_1^{(i)}, \dots, y_m^{(i)}\}$ .

The posterior distribution is derived as:

$$p(\boldsymbol{\theta} = \boldsymbol{\theta}^{(i)}|\mathbf{z}) = \frac{l^{(i)}}{\sum_{j=1}^N l^{(j)}} \quad (S4)$$

Where the likelihood  $l^{(i)}$  is defined as the normal distribution with the mean  $\mathbf{y}^{(i)}$  and the standard deviation  $\sigma$  as:

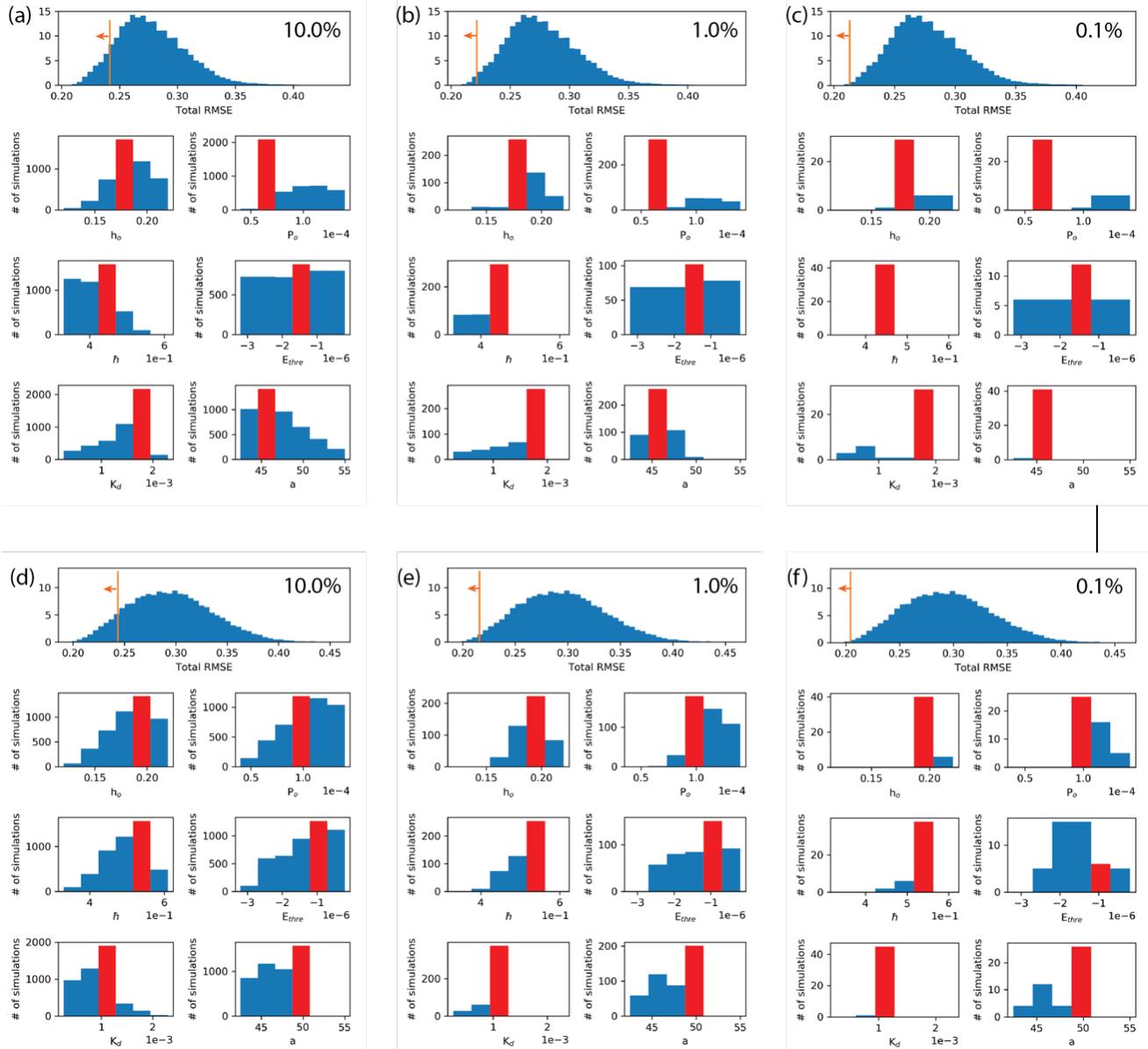
$$l^{(i)} = N(\mathbf{z} - \mathbf{y}^{(i)}, \sigma)$$

$$l^{(i)} = \prod_i^m N(z_i - y_1^{(i)}, \sigma) \quad (S5)$$

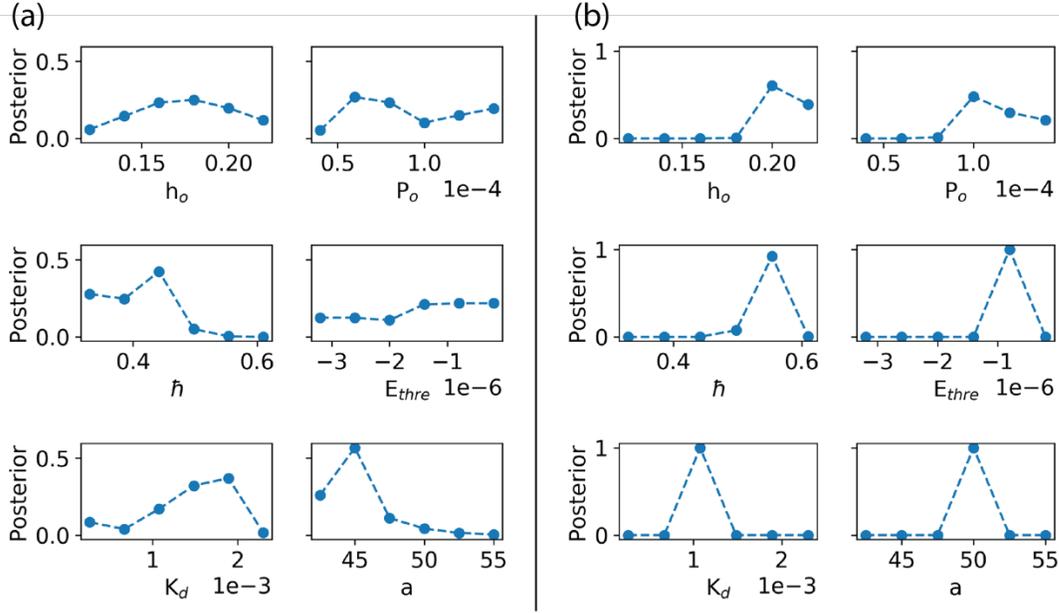
Note that  $C$  is the normalization factor. The standard deviation of the error is equal to the standard deviation of the difference. We then obtained the MAP estimate as the parameter set that provides the maximum posterior probability. The marginal distribution of each parameter can be calculated by the summation of the posterior distributions given each fixed parameter value ( Fig. S8).

## References

A. F. M. Smith and A. E. Gelfand, Bayesian statistics without tears: a sampling-resampling perspective, Am. Statistician, 46 (1992), pp. 84–88.



**Figure S7: Probability density function and histogram plots from a series of simulations of a grid search. The root-mean-square error (RMSE) between the simulated and measured soil thickness. (a-c) The RMSE for the south-facing hillslope and the corresponding histogram plots which show the distribution of each parameter that corresponds to 10.0%, 1.0%, and 0.1% of the smallest RMSE values, respectively; (d-f) The RMSE of north-facing hillslope and the corresponding histogram plots which show the distribution of each parameter that corresponds to 10.0%, 1.0%, and 0.1% of the smallest RMSE values, respectively. The red color bar represents the parameter that provides the global minimum between simulation and field measurement.**



**Figure S8: Marginal posterior probabilities of six parameters. (a) The posterior probabilities of the six parameters for the south-facing hillslope; (b) The posterior probabilities of the six parameters for the north-facing hillslope.**

**Table 1: Parameters used for fitting models of north-facing and south-facing, respectively**

symbol	Physical meaning	unit	Value range	North-facing	South-facing
$h_o$	Normalized soil depth	m	0.15–0.5 <sup>a,b,c,d</sup>	0.20	0.18
$P_o$	Potential weathering rate or the maximum bedrock weathering rate	m/yr	$5.0 \times 10^{-5} - 8.6 \times 10^{-5}$ <sup>a,b,c,d</sup>	$1.0 \times 10^{-4}$	$6.0 \times 10^{-5}$
$\bar{h}$	Spatially mean soil thickness among depositional area	m	0.2–0.8 <sup>e</sup>	0.55	0.44
$a$	The slope of the linear relationship between curvature and soil thickness among depositional area	m <sup>2</sup>	10–60 <sup>e</sup>	50	45
$K_d$	Topography diffusion coefficient, which is controlled by vegetation cover, grain size, animal disturbance, etc.	m <sup>2</sup> /yr	$1.0 \times 10^{-2} - 1.0 \times 10^{-2}$ <sup>d,f</sup>	$1.18 \times 10^{-3}$	$1.8 \times 10^{-3}$
$K_s$	Soil erodibility by overland flow erosion, which is controlled by overland flow rate, soil cohesivity, grain size, etc.	m/yr	$1.0 \times 10^{-2} - 1.0 \times 10^{-2}$ <sup>g</sup>	$1 \times 10^{-5}$	$2 \times 10^{-5}$
$E_{thre}$	The threshold value of annual soil thickness erosion rate that determines which model to use -- mass	m/yr	$1.0 \times 10^{-4} - 1.1 \times 10^{-4}$	$2 \times 10^{-6}$	$1.4 \times 10^{-6}$

	conservation method or Patton's method				
--	--	--	--	--	--

\*: we defined  $h$  as the distance along the norm direction to the land surface, which give  $e^{-h\cos\theta/h_0}$ , where  $\theta$  is the slope of the land surface in degree (Pelletier and Rasmussen, 2009). In this case,  $h_0$  is adjusted to include  $\cos\theta$  when referring to other literatures.

a: 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.

b: 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.

c: 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

d: 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.

e: Patton, N. R., Lohse, K. A., Seyfried, M. S., Godsey, S. E. and Parsons, S. B.: Topographic controls of soil organic carbon on soil-mantled landscapes, *Sci. Rep.*, 9(1), 6390, doi:10.1038/s41598-019-42556-5, 2019.

f: Fernandes, N. F. and Dietrich, W. E.: Hillslope evolution by diffusive processes: The timescale for equilibrium adjustments, *Water Resour. Res.*, 33(6), 1307–1318, doi:10.1029/97WR00534, 1997.

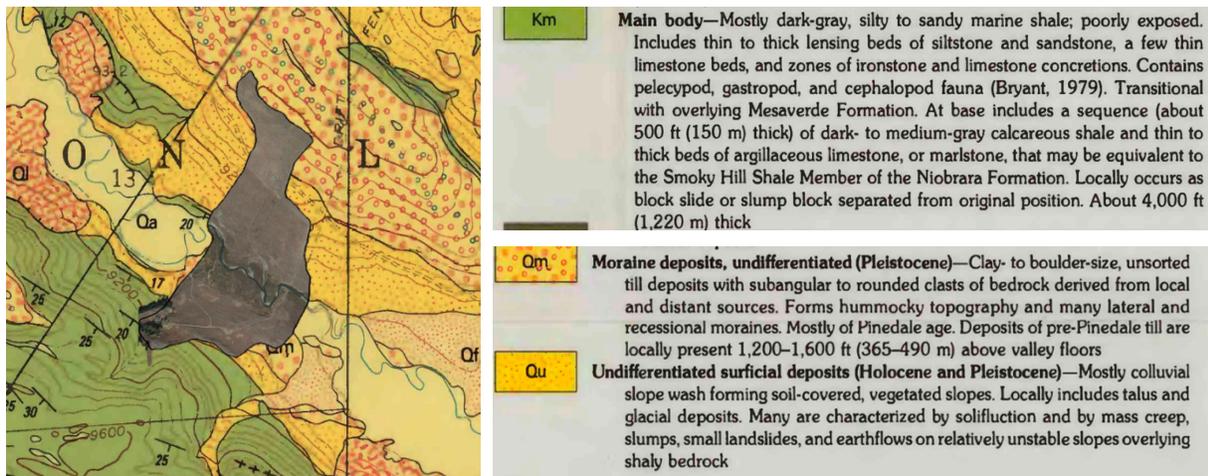
g: Kilinc, M. Y. and Richardson, E. V.: Mechanics of soil erosion from overland flow generated by simulated rainfall, *Hydrol. Pap.*, 63, 1973.

The geologic map was helpful to see that moraine deposits are not widespread in the study area. I trust the authors that most of the boreholes did not show glacial deposits but I must admit I would have liked to have seen some of these data. I think the paper would benefit from the inclusion of a brief conceptual model. It seems that authors assume that the study site was glaciated c. 18 ka and that such glaciation scoured the valley to bedrock and that glacial retreat started the “clock” for soil thickening by in situ soil production and colluvial deposition involving soil derived from upslope. This should be stated.

#### Our Response:

We thank the reviewer’s suggestion and provide a brief explanation for a conceptual model. We also include the surface geology map in the supplementary information as shown below to represent the spatial distribution of parent material and glacial deposits. We include a reference (below) to describe the bedrock and weathering condition in our study site. We include the following texts in the revised manuscript

*“Glacial deposits are mapped at many locations throughout the watershed (Gaskill, 1991; Fig. S3), but they are rather isolated and have a limited spatial extent, including in the area analyzed in this study. A former study at the same site analyzed 40 hand-augured soil cores and showed progressive changes in color and texture among soil, weathered zone, and unweathered bedrock with depth (Wan et al., 2019). Further, among total five wells drilled at the site, none of them reports the presence of glacial deposit (Tokunaga et al., 2019; Wan et al., 2019). It is likely that the glacial legacy scoured the valley to bedrock and the glacial retreat reset the ‘clock’ for soil formation mostly by in situ bedrock weathering and minor by colluvial deposition.”*



**Figure S3: A geological map of parent materials and deposits.**

Wan, J., Tokunaga, T. K., Williams, K. H., Dong, W., Brown, W., Henderson, A. N., Newman, A. W. and Hubbard, S. S.: Predicting sedimentary bedrock subsurface weathering fronts and weathering rates, *Sci. Rep.*, 9(1), 17198, doi:10.1038/s41598-019-53205-2, 2019.

Tokunaga, T. K., Wan, J., Williams, K. H., Brown, W., Henderson, A., Kim, Y., Tran, A. P., Conrad, M. E., Bill, M., Carroll, R. W. H., Dong, W., Xu, Z., Lavy, A., Gilbert, B., Romero, S., Christensen, J. N., Faybishenko, B., Arora, B., Siirila-Woodburn, E. R., Versteeg, R., Raberg, J. H., Peterson, J. E., Hubbard, S. S., Billings, S. A., Richter, D. de B., Ziegler, S. E., Prestegaard, K. and Wade, A. M.: Depth- and Time-Resolved Distributions of Snowmelt-Driven Hillslope Subsurface Flow and Transport and Their Contributions to Surface Waters, *Water Resour. Res.*, 55(11), 9474–9499, doi:https://doi.org/10.1029/2019WR025093, 2019.

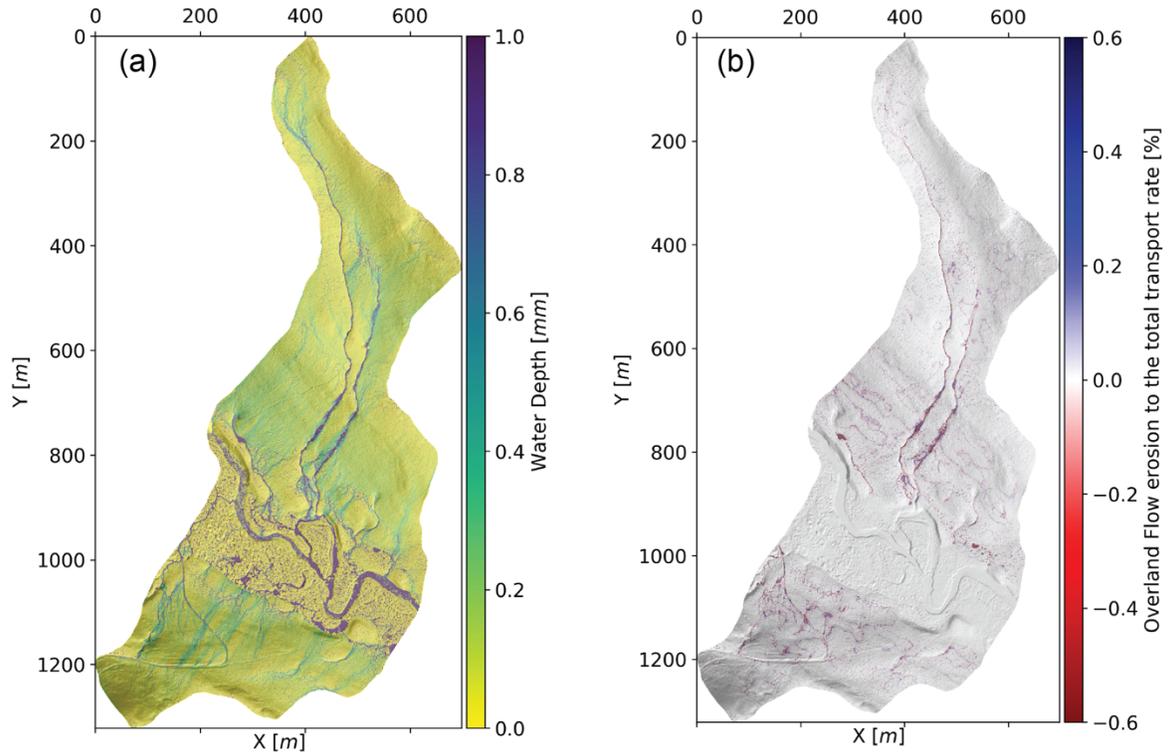
The overland flow component of the model (in terms of predicted flow depths and the erosion caused by such flow) cannot be evaluated because no results are presented on this aspect of the model. The reader needs to see a map of the overland flow and of the erosion attributable to this transport process. I have no idea how much transport is attributable to colluvial versus fluvial transport and whether the spatial patterns produced by each are realistic. This is a major omission that, again, prevents a comprehensive review of the manuscript.

*Our Response:*

We thank the reviewer’s suggestion and agree that a map is needed to present the overland and the corresponding erosion contribution to the total transport rate. We include the map of the fraction of the overland flow erosion to the total soil transport. And include the following texts:

*“The contribution of overland flow-driven soil transport (erosion and deposition) to the soil thickness formation is minor compared to the diffusion-driven soil transport. The water depth reaches steady-state after about six days with a constant rainfall intensity (i.e., 363 mm/yr) in the study area (Figure S11a). We apply this spatial map of the water depth to drive the soil transport from overland flow (Eqn. 3). The soil transport rate from overland flow mostly happens in the water pathways, which have no ponding after snow melting and storm events. The ratio of the soil transport rate from overland flow to the total soil transport rate is mostly less than 1% in the two hillslopes (Fig. S11b). This*

*minor impact from overland flow also explains why the parameter  $K_d$  is not sensitively in this hybrid model.”*



**Figure S10: (a) water depth of overland flow at a steady-state, which occurs after 6 days with constant rainfall = 363 mm/yr; (b) The rate of the overland flow erosion rate to the total soil transport rate. The overland flow erosion mechanism is from Equation 3.**