Manuscript MS No.: egusphere-2022-770

Response to reviewers

We thank the two anonymous referees for the attention devoted to our work. We report in the following a point-by-point response to the comments from the reviewers. Our own comments (in black unformatted text) are reported following referee comment *(in blue).*

On behalf of all authors,

Enrico Zorzetto

Referee #1

Topography controls many land surface processes. This manuscript combined an existing parameterization for solar radiation over complex terrain with a novel hierarchical multivariate clustering algorithm in GFDL. This work is very interesting and promising for applying in land surface models. However, how the authors considered the land cover types with different albedo values and energy balance is not clear; the performance of the proposed tile-level methods against the original grid-cell level methods for calculating regional average values is unknown; and more details in the physical explanations of some equations needs to be clarified,. Besides, how will the authors combine their tile separating and the existing tile schemes in GFDL? Please see below for my specific comments.

We thank this reviewer for their comments. See the point-by-point response below.

Major comments:

1. *Line108-109: the authors stated that in GFDL, the diffuse radiation received by the (flat) surface corresponds here to the sum of Fdif and Fcoup. If so, how did the authors calculate Fdif and Fcoup in Eq.1 for GFDL?*

   This is not a new issue for 3D radiation studies as most atmospheric models provide as a boundary condition to the land model Fdif as the sum of these two terms. The recommended way to obtain this quantity is to compute from the atmospheric model an additional estimate of Fdif by imposing a completely black surface (so that Fcoup=0 as there is no surface reflection) and then pass both values to the land model, so that Fcoup can be obtained from their difference. We will explicitly discuss this approach in the revised manuscript version.

In this work, the proposed parametrization for these two radiation flux terms was compared with Monte Carlo simulations in which they could be calculated exactly by tagging each photon after an atmospheric scattering event or after one or more than one reflection at the ground, so that the results shown in the paper are consistent in how Fdif is defined, so that Fdif shown and defined in the paper does never include Fcoup.

2. *Eq. 1 and line 260: how did the authors consider the land surface with different albedo (e.g., snow and vegetation)? Different land cover types may have different albedo and thus different reflected energy from adjacent terrain.*

In the analysis performed in this paper the surface albedo is assumed to be uniform over each domain. While this is not a realistic assumption, we believe at this stage this was a necessary one. Assuming uniform albedo allowed us to isolate the effects of topography, summarized in the set of predictors used in the predictive model developed here. Using non uniform albedo would considerably increase the complexity of the problem. First, a spatially variable albedo would have required a much more extensive set of Monte Carlo experiments, necessary to sample a wide enough range of surface conditions (i.e., combinations of local topographic features and surface reflectivity values). Second, in developing the statistical model linking terrain predictors to 3D radiation effects, the presence of spatially varying albedo would have greatly increased the dimensionality of the problem and the number of predictors needed. For example, in addition to the terrain view factor at each point, some additional measure accounting for the reflectivity of adjacent slopes visible from a target point would be needed. While this can be done for a point, defining such measure for our land tiles would have been quite complex, since they have variable configuration and geometry. Moreover, such a measure would need to be time-varying based on the condition of the surface (e.g., presence of vegetation, snow cover etc., would need to be considered at each model time step to compute the topographic correction) On the other hand, including surface albedo as a clustering variable to create land tile is not particularly difficult. We also note that in the current model formulation, the topographic corrections for reflected fluxes are indeed albedo-dependent, so that when applied to the surface are scaled by the albedo (However, as noted above, in the present study the albedo value is spatially uniform).

We will add these considerations to the revised manuscript. We believe this is an important point worth of future investigation, but difficult to pursue for the reasons stated here.

3. *Eq. 2: here the irradiance Ek,l is based on horizontal plane or the inclined plane of pixel k,l? Please give more physical explanations for equation 2.*

In the revised manuscript we will further clarify the meaning of eq. (2), reported here:

$$E_{k,l} = E_0 \cos\theta_0 \frac{1}{N} \frac{A}{A_{k,l}} \sum_{i=1}^{N_{k,l}^{(s)}} w_i$$

This equation is used to compute the energy $E_{k,l}$ received by each cell of the land surface model, identified by the indices k,l. While a 3D mesh is used in the Monte Carlo simulations, the area of the cell $A_{k,l}$is defined as the area of the cell on the horizontal plane, for consistency with the definition of area in the land model where the method will be applied, while A is the horizontal area of the entire domain and N the total number of incident photons. The equation computes the radiation received by a single land surface cell as a fraction of the radiation flux at the top of the surface E_0 by summing the "energy packets" w_i of the photons absorbed over that area. Since the interactions of each photon are tagged (e.g., atmospheric scattering and / or previous reflections at the surface) the radiation received can be classified in one of the 5 flux components as defined in the paper. We will clarify the description of the equation to include these details.

4. *Line 264-266: please give more details about the energy conservation and albedo modification.*

We agree with the comment and will add some additional comments on energy conservation and albedo modifications. We did not add much detail on this in our original submission because the approach is described in Lee et al., 2015, but we agree that this is relevant for applications. The reason for this correction is that, by correcting fluxes received by land due to topography, energy is not necessarily conserved within a single model column. This is for example the case due to non-local effects: some model grid cells may in general receive overall more or less radiation due to their average topographic properties. This is in general accompanied by changes in the radiation received by neighboring grid cells. Moreover, in the case of reflections, the surface receives in general more energy with respect to the case of a flat surface. Properly accounting for these local and non-local effects is challenging in current ESMs, in which each land grid cell is directly coupled with the atmosphere but not directly with nearby model columns. A way to ensure that 3D radiation effects can be accounted for was proposed by Lee et al., 2015: In this approach, an effective "3D albedo" is computed for each land model grid cell, such that a land grid—cell characterized by this “3D albedo” and forced by plane-parallel radiation (PP) absorbs the same amount of radiation as in the case of a surface characterized by actual land albedo (PP), forced by the 3D-corrected downward radiation fluxes. By returning the 3D albedo (which effectively represents the reflectivity of a “rough” land surface) to the atmosphere, energy is conserved while accounting for the 3D topographic correction.

5. *Line 273-274: will their difference be larger for cloudy condition?*

We note that the presence of clouds adds considerable complexity to the problem of radiation-topography interactions, and has not been considered in our work. Previous studies, in particular Lee et al., 2011, also based their work on clear-sky condition in order to make the problem more manageable. The main limitation would be the number of Monte Carlo simulations to run to sample different atmospheric conditions, and the increasing number of parameters in the statistical model used to estimate 3D topographic radiation corrections. However, given the importance of cloud cover we will add the following comments when discussing the limitations of our work.

It is quite possible that in cloudy conditions the differences between 3D and PP radiation fields will be different than in the case of clear sky analyzed here. However, we would expect the largest differences to arise in case of non-homogeneous cloud cover over the domain. This is a very difficult problem to model, as the number of configurations of 3D clouds and topography would be difficult to manage.

Even In the case of homogeneous cloud cover, we would expect a change of the relative magnitude of the 5 flux components, although the effect on 3D corrections proposed here remains unknown. We welcome future work focusing on this important issue.

6. *Figure 5: why can sky view factors be larger to 1?*

The reason is that what we are plotting here is the ratio of the sky view factor (a number between 0 and 1) to the cosine of the local terrain slope (smaller or equal to one), so that their ratio can be larger than unity. This quantity is defined in the paper using a tilde and is used as a predictor in the statistical model. However, we will add a comment on this point to make this definition clearer for the reader.

7. Eq. 12: is this method only empirical?

The equation is not empirical. The correction in eq. (12) was obtained by imposing the grid--cell average of the quantity of interest is conserved within the specified domain, while at the same time preserving its physical lower bound. We will clarify this in the manuscript.

8. How about the performance of the proposed tile-level methods against the original grid-cell level methods for calculating regional average values?

   We agree that comparing the performance of the sub—grid parameterization with the regional average value is important. We will add a comparison between grid—average model and sub—grid estimates in the revised manuscript. However, we note that the analysis reported in Figure 7 already goes in this direction, as it tests the predictive model against "ground truth" Monte Carlo simulations for increasing values of spatial averaging scale. We note here that at small scales, performance increases with averaging scale. Therefore, it is easier for the model to capture the average effect over at a large enough scale. At large enough scale (say, above 10km) the approach should effectively become equivalent to a grid—cell average prediction, depending on the grid cell scale.

   On the other hand, we note here that the main advantage of capturing the sub—grid distribution of irradiance can be appreciated only when examining the effects on a model run. This Is true especially for variables that are nonlinearly related to shortwave radiation, such as e.g., snow cover and land surface temperature. In this case, we believe that using a refined sub-grid distribution of solar radiation, while keeping its average value constant, can lead to non-zero grid average effects. This should be the object of a follow-up study, running a land model with both sub—grid and grid average models.

   We will add these considerations to the revised version of the manuscript.

9. The authors presents the results based on the corrected factors in Eq.1. However, they may be not easy to understand. How about presenting some results about the radiation fluxes directly, which will be more clearer for the readers?

   We agree with this comment and will add these results in the revised version of the manuscript. We have shown dimensionless quantities since the model produces corrections in this form. However, dimensional corrections can be directly obtained from these and we agree that including them in the presentation would improve the physical insight from our results.

10. The authors proposed tile-level topographic correction methods for solar radiation over complex terrain. However, current sub-grid tile schemes in GFDL consider different soil and vegetation, and topographic characteristics for simulating water and carbon cycles. How did the authors merge their clustering methods for radiation and the existing scheme in GFDL for other processes?

    The approach proposed in our manuscript was developed keeping in mind the necessity of describing other physical processes at the sub—grid scale. In particular, the fact that the clustering is hierarchical is not strictly necessary for the purpose of solar radiation-topography interactions. A single--level terrain clustering would suffice for this purpose. However, the multi-level clustering used here accounts for the need for other land processes. For example, the outer level clustering (i.e., the partition of the domain in k characteristic hillslopes) is done to obtain hydrologically coherent units (for an example of their application to study soil moisture, we refer the reader to Chaney et al., 2018). We retain this flexible sub—grid structure, and in our sensitivity study (Figure 13 in the manuscript) we compare different specifications of k and p (number of characteristic hillslopes, and number of inner clusters within each hillslope) to test the sensitivity of our parameterization to these changes in sub-grid structure. Finally, it is likely that other variables may need to be added to the clustering to account for sub—grid heterogeneity of other processes. For example, soil properties. This can be directly done with the framework used here, at the price of an increase in the number of tiles used. Producing an effective global--scale model grid able to meet these demands is possible but requires some tradeoffs. The analysis in this paper contribute to this effort by quantifying the number of tiles needed over mountainous terrain for the sole purpose of capturing the spatial variability in shortwave radiation.

Minor comments:

1. Line 16-18: It will be better to show some quantitative metrics rather than only descriptive expression.

    We agree with this comment. We will report in the abstract the main quantitative findings from our work: In particular, the magnitude of local topographic effects in our domain with respect to grid average estimates, and the number of sub—grid units necessary to represent sub—grid heterogeneity in downward fluxes.

2. Line 42: why did the author call this method ‘WLH’?

    In the revised version of the manuscript, this method will be called LLH following the initials of the authors of the Lee et al., 2011 paper.

3. Line 60-85: these summarize the objective and work of this paper. I suggest the authors simplify them for making them clearer.

   We agree with the suggestion and will simplify this part of the introduction providing a better summary of our objectives.

4. Line124-125: Citing the corresponding papers may be better.

   The papers cited here were not properly formatted. We will revise the formatting of this paragraph in the revised version, including the correct citations.

5. Line 93: how about vegetation with different PFTs?

   Vegetation is certainly relevant for this problem, since it significantly modulates land albedo. Vegetation was not explicitly considered in our study, except to the extent to which it contributes to the albedo of each land tile. In our work, albedo is indeed accounted for to estimate the magnitude of reflected fluxes (see our response to major comment #2). However, vegetation PFTs and land use are not used to cluster the domain in our study. For the clustering, we focused on topographic quantities derived from the digital elevation model which are known to modulate irradiance over mountains. Surface albedo, vegetation and land use are certainly relevant variables for the problem at hand, and including them in the set of variables used to cluster land is in principle possible. However, the main benefit of doing so would stem from being able explicitly track radiation reflected between pairs of tiles: i.e., explicitly considering the albedo of nearby slopes visible from a tile, instead of computing reflected fluxes base on a uniform albedo value. While appealing, this approach would greatly increase the dimensionality of the problem and lead to a much more complex parameterization for reflected fluxes. One additional complication is the potentially time varying land use and vegetation structure which is not considered in our static definition of land—sub--grid structures.

6. Line 207: kp -> k*p?

   Agree, will revise as suggested

7. *Line 413: he -> the*

   Agree, will revise as suggested

Referee #2

Summary and general comments

In this study, a parameterisation for the effects of sub-grid topography on surface shortwave radiation is presented. In a first step, the authors apply Monte Carlo ray tracing to simulate surface shortwave radiation for 3 geographic domains with complex terrain. These experiments serve as a reference to develop the (sub-grid) parametrisation. In a next step, terrain properties (μ, sky view factor and terrain configuration) are linked to modulated radiation fluxes with two statistical models – a Multiple Linear and a Random Forest Regression. Finally, sub-grid effects are considered by merging land units within a grid cell with similar terrain properties by means of hierarchical clustering.

The aim of this study is very interesting and relevant – namely improving the representation of surface shortwave radiation fluxes in an Earth System Model. Due to the plane parallel radiative transfer schemes applied in such models, surface radiation is typically simulated rather inaccurately in areas with complex terrain. The implementation of parameterisations, *particularly on a sub-grid scale, has the potential to strongly reduce such biases. The approach presented by the authors is very interesting and the manuscript is well written and structured. However, I struggled to understand certain sections in detail – for instance the hierarchical clustering section in the methods and some passages in the Results and Discussion. Furthermore, the Results and Discussion section is sometimes incomplete in my opinion and should be extended (see the following comments for more details).*

Major comments

Section about hierarchical clustering (2.4)

Until section 2.4, the methodology is very well described. However, I struggled to follow section 2.4. For instance, why do you want to partition land in hydrologically coherent units? From a "terrain-radiation-perspective" – this is not obvious. Has this approach been chosen due to an already existing tile classification in the GFDL Land Model?

This approach we followed has been selected to be compatible with the existing structure of the GFDL land model. Many physical processes other that radiation-topography interactions benefit from the sub-grid structure. However, using a different sub-grid partition for different processes does not appear to be a viable solution, due to considerable increase in model complexity and computational expense. For this reason, land processes are solved in a column for each land tile, using as boundary condition downward fluxes corrected based on 3D topography.

We report here a response to reviewer #1 which addresses this point:

> The approach proposed in our manuscript was developed keeping in mind the necessity of describing other physical processes at the sub—grid scale. In particular, the fact that the clustering is hierarchical is not strictly necessary for the purpose of solar radiation-topography interactions. A single--level terrain clustering would suffice for this purpose. However, the multi-level clustering used here accounts for the need for other land processes. For example, the outer level clustering (i.e., the partition of the domain in k characteristic hillslopes) is done to obtain hydrologically coherent units (for an example of their application to study soil moisture, we refer the reader to Chaney et al., 2018). We retain this flexible sub—grid structure, and in our sensitivity study (Figure 13 in the manuscript) we compare different specifications of k and p (number of characteristic hillslopes, and number of inner clusters within each

hillslope) to test the sensitivity of our parameterization to these changes in sub-grid structure. Finally, it is likely that other variables may need to be added to the clustering to account for sub—grid heterogeneity of other processes. For example, soil properties. This can be directly done with the framework used here, at the price of an increase in the number of tiles used. Producing an effective global--scale model grid able to meet these demands is possible but requires some tradeoffs. The analysis in this paper contribute to this effort by quantifying the number of tiles needed over mountainous terrain for the sole purpose of capturing the spatial variability in shortwave radiation.

I'm also confused why the clustering is performed twice (first in k hillslopes, then in p sub-units). I think a detailed flow diagram (e.g. with an example of the step-wise classification of sub-units of a geographic domain) would help the reader to understand these steps. Furthermore, it is also not obvious to me why lakes and glaciers represent separate classes. And are glaciers and lake classes further divided into sub-classes according to their terrain properties? Finally, some parts of section 3.4 (e.g. starting from line 354 could also be moved to the method section).

The reason for the hierarchical clustering is precisely that it must accommodate other physical processes other that radiation. For example, the subdivision in hillslopes is suitable to hydrological studies (e.g., Chaney et al., 2018). See our detailed response to the previous comment. We agree that a workflow diagram would help the reader understand our work and will include one in the revised manuscript.

In the current GFDL land model structure, lake and glacier are treated as separate land classes (solved as a separate "vertical columns"). In our work, land is subdivided in tiles using the hierarchical clustering scheme, while glacier and lake are treated each as a single tile, each characterized by average topographic properties.

However, we note that in our study domains glacier and lakes constitute a small fraction of the total area. There is no reason why glacier and lakes could not also be subdivided in multiple tiles if they occupy a relevant portion of the domain. This would be comparable to what we have done here, since the only difference for the purpose of our study would be the reflectivity of the surface. Our results here would equally apply to the case in which lake and glaciers were also partitioned in multiple clusters, since for the purposes of this study these three land classes behave the same way.

Analysis and results – improve consistency and completeness

- I'm missing the third domain (Nepal) in Fig. 7. I guess you used one domain to train the model and the other two domains for cross-validation – right?

Due to the considerable computational expense of ray tracing simulations, we performed these analyses for two domains only (for a large number of solar angles) and used them for training and testing. Three domains are then used to construct clustering and evaluate cluster-by-cluster results with high-resolution results.

- I think a performance comparison of the sub-grid to a grid-scale parameterisation would be very interesting to show. With this, you could emphasize the additional benefit of the sub-grid scale scheme.

  We agree that this comparison would be useful, as also pointed out by referee #1. See our comment below:

We agree that comparing the performance of the sub—grid parameterization with the regional average value is important. We will add a comparison between grid—average model and sub—grid estimates in the revised manuscript. However, we note that the analysis reported in Figure 7 already goes in this direction, as it tests the predictive model against "ground truth" Monte Carlo simulations for increasing values of spatial averaging scale. We note here that at small scales, performance increases with averaging scale. Therefore, it is easier for the model to capture the average effect over at a large enough scale. At large enough scale (say, above 10km) the approach should effectively become equivalent to a grid—cell average prediction, depending on the grid cell scale.

On the other hand, we note here that the main advantage of capturing the sub—grid distribution of irradiance can be appreciated only when examining the effects on a model run. This Is true especially for variables that are nonlinearly related to shortwave radiation, such as e.g., snow cover and land surface temperature. In this case, we believe that using a refined sub-grid distribution of solar radiation, while keeping its average value constant, can lead to non-zero grid average effects. This should be the object of a follow-up study, running a land model with both sub—grid and grid average models.

We will add these considerations to the revised version of the manuscript.

- The discussion of certain findings should be extended. From the results, it seems that a tile number of ~100 captures the sub-grid characteristics already very well. Do you agree? And would such a number be feasible in an online ESM simulation?

  We agree with the referee that this is a very important point, since one of the main purposes of the clustering technique used here is to make this problem tractable at the global scale.

  The fact that about 100 clusters capture a significant fraction of the spatial distribution of irradiance is a very encouraging result in our opinion. To date, it is unfeasible to run a global model with 100 tiles / grid cell. An average number of 5 to 10 tiles is possible, however. We note that the domains selected for this study are characterized by very complex terrain, but not all land areas are. Therefore, we would argue that a grid setup could be constructed ranging from say 20 to 50 tiles in high mountains areas to a few over flat areas would be feasible. In the revised paper, we provide an estimate and guideline for constructing such a grid. For example, this can be done by increasing linearly the number of tiles from a small number (say, 3) to a large number (say, 20) based on the local elevation standard deviation. Moreover, we note that reproducing the entire spatial variance of solar irradiance (the values to which plots in Figure 13 converge) is certainly a bold objective for a global scale model, and a more limited number of tiles would still be appealing when compared to using simple grid—cell average values.

Minor comments

Content-related (text)

Line 42: what does the abbreviation “WLH” stand for?

The abbreviation was revised to LLH, to indicate the initials of the authors of the study referenced here.

L139: “uniform” albedo -> how realistic is this assumption?

See also related comment to referee #1:

In the analysis performed in this paper the surface albedo is assumed to be uniform over each domain. While this is not a realistic assumption, we believe at this stage this was a necessary one. Assuming uniform albedo allowed us to isolate the effects of topography, summarized in the set of predictors used in the predictive model developed here. Using non uniform albedo would considerably increase the complexity of the problem. First, a spatially variable albedo would have required a much more extensive set of Monte Carlo experiments, necessary to sample a wide enough range of surface conditions (i.e., combinations of local topographic features and surface reflectivity values). Second, in developing the statistical model linking terrain predictors to 3D radiation effects, the presence of spatially varying albedo would have greatly increased the dimensionality of the problem and the number of predictors needed. For example, in addition to the terrain view factor at each point, some additional measure accounting for the reflectivity of adjacent slopes visible from a target point would be needed. While this can be done for a point, defining such measure for our land tiles would have been quite complex, since they have variable configuration and geometry. Moreover, such a measure would need to be time-varying based on the condition of the surface (e.g., presence of vegetation, snow cover etc., would need to be considered at each model time step to compute the topographic correction) On the other hand, including surface albedo as a clustering variable to create land tile is not particularly difficult. We also note that in the current model formulation, the topographic corrections for reflected fluxes are indeed albedo-dependent, so that when applied to the surface are scaled by the albedo (However, as noted above, in the present study the albedo value is spatially uniform).

We will add these considerations to the revised manuscript. We believe this is an important point worth of future investigation, but difficult for the reasons stated here.

L139: I appreciate such clear definitions, it simplifies the comprehensibility of the subsequent text greatly!

Thank you for the comment!

L157: I'm not sure if I understand this sentence correctly. Do you mean that radiation fluxes significantly departure **locally** from areal-average fluxes?

Precisely. This was clarified to make it clear for the reader what we mean.

L 162: "represents the fraction of the sky dome visible from a target site" -> technically, this is incorrect. Compare e.g. with Helbig et al. (2009) (text next to Eq. 8) and Zakšek et al. (2011). The sky view factor definition of Dozier and Frew (1990) yields the fraction of hemispherical radiation received under the assumption of isotropic radiation. The same is valid for the subsequent explanation of the terrain configuration factor C_t.

We agree with the comment, the sentence was not correct. We will rephrase stating that the sky view factor used here is defined as the ratio of diffuse sky irradiance at a point to that on an unobstructed horizontal surface, in case of isotropic diffuse radiation.

L171: Could you explain why you use this terrain configuration definition and not simply $C_t = 1.0 – V_d$ (compare e.g. with Chu et al., 2021)?

Thank you for pointing out this interesting reference.

According to Dozier and Frew (1990) and e.g., Chen et al. (2006) the relation between sky view factor and terrain configuration factor can be approximated to (See eq. 9b in Dozier and Frew 1990, or eq. 6 in Chen et al, 2006):

$$C_t = \frac{1}{\pi}\int_0^{2\pi}\int_{H_\phi}^{\psi_\phi} \sin\theta[\cos\theta\cos\theta_s + \sin\theta\sin\theta_s\cos(\phi-\phi_s)]d\theta d\phi \approx \frac{1+\cos\theta_s}{2} - V_d$$

The relation $C_t = 1.0 – V_d$ can be further obtained by assuming locally flat surface $(\cos\theta_s = 1)$, but in general does not hold for a sloping surface.

L174: I would briefly introduce and explain the parameters μ_i and μ_0 here.

Agree, will revised as suggested

L194: I'm a bit confused by these lines. It seems that you perform the clustering only for soil elements (also according to line 207; kpand kp + 2) and not for glaciers and lakes. What is the reason behind this? I guess glaciated areas and lakes can also have very variable topographic parameters (like e.g. sky view factor).

Yes, that is correct. The reason for this is that in the current GFDL land model configuration, glacier and lakes are treated separately from "soil" land, and in the current work we cluster soil in hillslopes and clusters, but consider lakes and glaciers each as a single cluster characterized by their areal-average topographic properties.

However, we note that in our study domains glacier and lakes constitute a small fraction of the total area. There is no reason why glacier and lakes could not also be subdivided in multiple tiles if they occupy a relevant portion of the domain. This would be comparable to what we have done here, since the only difference for the purpose of our study would be the reflectivity of the surface. Our results here would equally apply to the case in which lake and glaciers were also partitioned in multiple clusters, since for the purposes of this study these three land classes behave the same way.

L204: It's not obvious to me why you apply the clustering a second time. Generally, to increase the comprehensibility of this section, it might be worth to extend the workflow diagram displayed in Fig. 4. One could show the classification of a certain domain (resolved for every single step).

[see also following comment, and our response to major comment #1] We agree that this point should be clarified in the text and will add a workflow diagram as suggested.

The main reason for the multi-level clustering is obtaining a flexible sub—grid structure able to be applied to other processes other that the topographic radiation correction pursued here. We discuss this in detail in the response to comment #1.

L207: I'm still a bit puzzled – what is the motivation behind categorizing land surface based on hydrological properties? I don't see the connection to topography-radiation-processes.

[*See also our response to the previous comment, and to major comment #1*]

L298: "reflected components are quite linear" -> for f_{rdir}, the deviations between MLR and RFR are quite substantial…

This is correct. We argue that this is due to the interaction of solar incident effects (relevant for direct incident light) and terrain configuration, which is relevant for reflected fluxes. We will point this out in the revised text and amend this sentence.

L 305: "case in which…" -> I don't understand this part; there is probably something missing.

We agree the sentence is unclear and will revise it as follows: *"In the case of larger spatial averaging scales and larger solar angles, the MLR describes the direct flux with great accuracy".*

L316: First of all, I'm confused about which region (East Alps vs. Peru) is the (in-)dependent domain. The caption of Fig. 7 does not agree with the statement here.

Furthermore, I'm not convinced that results from RFR are not location dependent. Looking at Fig. 7, the RFR method consistently indicates a worse performance for the cross-validation domain than the MLR method. For me, this is an indication that obtained relations from the RFR simulation are very location-dependent and not easily transferable to other terrain geometries (i.e. the model is overfitted).

As stated in the caption of Figure 7, here *"The models were trained over the Peru site and tested over the same site (SS, continuous lines) and over the independent EastAlps site for cross validation (CV, dashed lines). ".* At line 316 we meant to say that we did run both configurations (switching the training and testing sites) and obtained comparable results. We will make this clearer in the text and will include the other case in the online supplementary material for completeness.

We agree with the second part of this comment. The statistical models used "learn" from the input data and this is indeed the reason why the cross validation was performed. We think it is important to quantify this effect. Of course, for applications we are interested in extrapolating to different regions. For extrapolation, we agree that the in-sample performance is not representative. For this reason, the out-of-sample goodness of fit (which, while not being as good as the in-sample one, is not that bad) can be used to assess method performance. We also agree that RFR does appear to overfit the training data to some extent (exhibiting a larger difference between in- and out-of-sample cases with respect to MLR), and this is the reason why in the discussion we argue about the advantages of the linear model (MLR), despite some nonlinear behavior being observed for some of the flux components. Will make this tradeoff between the two models more explicit in the discussion section.

L388: It would be interesting to see the results for these tests too. Maybe you could show them in the supplementary material.

Agree. These results can be automatically generated running the code included with manuscript and will be included in a supplementary material in our revised submission.

Typos, phrasing and stylistic comments

L124: references not correctly rendered

Agree, will be revised as suggested

L153: I was a bit confused by this line, it might be better to write something like: "The MC calculations were performed for three independent domains (Nepal, Peru, East Alps)..." (if that is what you mean)

We run the Monte Carlo simulations for two domains only, as explained above due to the computational expense and the need to run for several solar angles. Three domains are used in the clustering study to get additional data points to evaluate the effect of the number of tiles, but the third domain is not used to train and test the predictive model for the radiation correction terms. We will clarify this important point to make it clear.

L157: "determines" -> "determine"

Agree, will revise as suggested

L162: "represent" -> "represents"

Agree, will revise as suggested

L166: "in order to compute **the** sky view factor"

Agree, will revise as suggested

L198: "eq. 6" -> "Eq. 6"

Agree, will revise as suggested

L215: "if these **are** present in a given grid cell."

Agree, will revise as suggested

L220: "**the** is the indicator" -> "is the indicator"

Agree, will revise as suggested

L263: "eqns. (1)" -> "Eq. (1)"

Agree, will revise as suggested

L273: "angles compute based" -> "angles compute**d** based"

Agree, will revise as suggested

L273: "simulation (5)" -> "simulation (Fig. 5)"?

Agree, will revise as suggested

L290: I would rewrite this to e.g.: "...larger than approximately 5 km the effect disappears."

Agree, will revise as suggested

L302: "case in which" -> "a case in which"

Agree, will revise as suggested

Figures and Tables

Figure 2: The colorbar labelling is erroneous – I guess it should be "Elevation [m a.s.l.]". The same is true for the upper-left panel in figure 3. Furthermore, the degree symbol is missing for the cardinal directions.

The colorbar label refers to elevation above mean sea level (m.s.l.). We do not believe it is erroneous as reported in our submission. Elevation [m a.s.l.] as suggested by the reviewer would also be correct. However, we will clarify that we mean elevation above mean sea level to avoid any confusion.

Figure 4: It seems from these panels (x/y-coordinates) that the MC model was run on a map projection. Could you specify the projection somewhere?

Agree – it is an equal area Mollweide projection. We will include this information in the revised manuscript.

Figure 7: μ_0 not correctly rendered in caption

Will be revised as suggested

New references

Chu, Q., Yan, G., Qi, J., Mu, X., Li, L., Tong, Y., et al. (2021). Quantitative analysis of terrain reflected solar radiation in snow-covered mountains: A case study in Southeastern Tibetan Plateau. Journal of Geophysical Research: Atmospheres, 126, e2020JD034294. https://doi.org/10.1029/2020JD034294

Helbig, N., Löwe, H., & Lehning, M. (2009). Radiosity Approach for the Shortwave Surface Radiation Balance in Complex Terrain, Journal of the Atmospheric Sciences, 66(9), 2900-2912. https://doi.org/10.1175/2009JAS2940.1

Zakšek, K.; Oštir, K.; Kokalj, Å½. Sky-View Factor as a Relief Visualization Technique. Remote Sens. 2011, 3, 398-415. https://doi.org/10.3390/rs3020398

References

Lee, Wei-Liang, K. N. Liou, and Alex Hall. "Parameterization of solar fluxes over mountain surfaces for application to climate models." Journal of Geophysical Research: Atmospheres116.D1 (2011).

Lee, W-L., et al. "A global model simulation for 3-D radiative transfer impact on surface hydrology over the Sierra Nevada and Rocky Mountains." Atmospheric Chemistry and Physics15.10 (2015): 5405-5413.

Dozier, Jeff, and James Frew. "Rapid calculation of terrain parameters for radiation modeling from digital elevation data." IEEE Transactions on geoscience and remote sensing 28.5 (1990): 963-969.

Chaney, N. W., Van Huijgevoort, M. H., Shevliakova, E., Malyshev, S., Milly, P. C., Gauthier, P. P., & Sulman, B. N. (2018). Harnessing big data to rethink land heterogeneity in Earth system models. Hydrology and Earth System Sciences, 22(6), 3311-3330.