Effects of complex terrain on the shortwave radiative balance : A sub–grid scale parameterization for the GFDL Land Model version 4.2

Enrico Zorzetto[1], Sergey Malyshev[2], Nathaniel Chaney[3], David Paynter[2], Raymond Menzel[2, 4], and Elena Shevliakova[2]

[1]Program in Atmospheric and Oceanic Sciences, Princeton University, Princeton, NJ, USA
[2]NOAA OAR Geophysical Fluid Dynamics Laboratory, Princeton, NJ, USA
[3]Department of Civil and Environmental Engineering, Duke University, Durham, NC, USA
[4]University Corporation for Atmospheric Research, Princeton, NJ, USA

Correspondence: Enrico Zorzetto (ez6263@princeton.edu)

Abstract.

Parameterizing incident solar radiation over complex topography regions in Earth System Models (ESMs) remains a challenging task. In ESMs, downward solar radiative fluxes at the surface are typically computed using plane parallel radiative transfer schemes, which do not explicitly account for the effects of a three-dimensional topography, such as shading and re-

5 flections. To improve the representation of these processes, we introduce and test a parameterization of radiation-topography interactions tailored to the Geophysical Fluid Dynamics Laboratory (GFDL) ESM land model. The approach presented here builds on an existing correction scheme for direct, diffuse and reflected solar irradiance terms over three-dimensional terrain. Here we combine this correction with a novel hierarchical multivariate clustering algorithm which explicitly describes the spatially varying downward irradiance over mountainous terrain. Based on a high-resolution digital elevation model, this combined

10 method first defines a set of sub–grid land units ("tiles") by clustering together sites characterized by similar terrain-radiation interactions (e.g., areas with similar slope orientation, terrain and sky view factors). Then, based on terrain parameters characteristic for each tile, correction terms are computed to account for the effects of local 3-D topography on shortwave radiation over each land unit. We develop and test this procedure based on a set of Monte Carlo ray tracing simulations approximating the true radiative transfer process over three dimensional topography. Domains located in three distinct geographic regions

15 (Alps, Andes, and Himalaya) are included in this study to allow for independent testing of the methodology over surfaces with differing topographic features. We find that accounting for the sub–grid spatial variability of solar irradiance originating from interactions with complex topography is important as these effects lead to significant local differences with respect to the plane-parallel case, as well as with respect to grid–cell scale average topographic corrections. Finally, we quantify the importance of the topographic correction for a varying number of terrain clusters and for different radiation terms (direct, diffuse,

20 and reflected radiative fluxes) in order to inform the application of this methodology in different ESMs with varying sub-grid tile structure.

1 Introduction

The presence of three-dimensional topography exerts an important control on the amount of solar radiation received by land. Over complex terrain, the incoming solar beam does not only undergo scattering and absorption within the atmospheric column, but is further modulated by the relative orientation of land surfaces, as well as by the potential shading and reflection effects by neighbouring slopes (Sirguey, 2009; Lenot et al., 2009; Lamare et al., 2020; Picard et al., 2020). The effect of surface roughness was recently shown to have important effects over snow covered surfaces, leading to a net decrease in surface reflectivity (Larue et al., 2020).

Together, these effects lead to a spatially heterogeneous distribution of the radiative fluxes received by the surface. In turn, this heterogeneity can have important consequences for the local energy and water balance, and interact with other spatially–varying land processes such as evaporation (Brutsaert, 2013), snow melting (McCabe and Clark, 2005; Bales et al., 2006; Sirguey et al., 2009) and vegetation dynamics (Granger and Schulze, 1977; Gu et al., 2002).

Representing these processes at increasingly fine scales is the goal of state-of-the-art land components of Earth System Models (ESMs) (i.e., land models). However, global circulation models (GCMs) routinely compute shortwave radiative fluxes based on plane-parallel (PP) radiative transfer schemes which do not account for the effect of topography. This discrepancy poses a challenge for adequately capturing sub–grid scale processes in land models.

Several models have been proposed to account for the interaction of downward solar irradiance with complex topography, accounting for slope orientation and shading effects (Isard, 1986; Hay and McKay, 1985; Duguay, 1993) as well as for the effect of surrounding slopes (Dozier, 1980; Dubayah et al., 1990; Dozier and Frew, 1990).

A recently developed radiation parameterization was developed to predict radiative fluxes over mountainous terrain via multiple linear regression (Chen et al., 2006; Lee et al., 2011). This approach (henceforth termed WLH) links flux corrections over mountains to a set of grid-cell average terrain variables which summarize the three-dimensional nature of the land surface and are used as predictors for shortwave fluxes. The WLH parameterization for shortwave radiation over mountains has been implemented in Global Climate Models (GCMs) and the Weather Research and Forecasting (WRF) model; it has been extensively tested over the Tibetan plateau and the Western United States (Liou et al., 2007; Essery and Marks, 2007; Gu et al., 2012; Lee et al., 2013, 2015, 2019).

However, it is expected that sub–grid variability of these topographic effects may play a relevant role given the heterogeneous nature of surface reflectivity and topographic features at scales smaller than the typical GCM grid cell. The importance of accounting for sub grid–scale topography when correcting shortwave radiative fluxes over mountains was recently pointed out by an application of WLH to the DOE E3SM Exascale Earth System Model, varying model resolution over a range of scales relevant for land processes (Hao et al., 2021).

This problem is especially relevant since in recent years the development of land models has increasingly been focusing on the description of sug–grid variability of terrain properties (Tesfa and Leung, 2017; Chaney et al., 2018). For example, in the

current iteration of the GFDL land model, this objective is achieved by summarizing grid cell heterogeneity in sets of land units (termed "tiles") characterized by approximately homogeneous physical features (Shevliakova et al., 2009; Milly et al., 2014; Zhao et al., 2018; Dunne et al., 2020), including elevation, land cover, soil properties, and other environmental variables (Chaney et al., 2018).

In the GFDL land model such a sub-grid representation has not been yet tailored to describing the interaction of shortwave radiation with topography. The objective of this work is bridging this discrepancy and developing a sub–grid parameterization for the effects of radiation over complex topography in the GFDL land model.

In this paper, the WLH approach is extended in order to account for the sub-grid heterogeneity of terrain features. For this purpose, we employ a hierarchical multivariate clustering approach (Chaney et al., 2016, 2018) to partition land domains in a set of clusters termed *tiles*. Tiles are here defined as land units characterized by homogeneous topographic effects with respect to downward shortwave radiation. Thus, the terrain variables used for clustering land surfaces encode the physical mechanisms determining the spatial variability of radiative fluxes, such as shading and reflection from nearby slopes.

This clustering approach provides a parsimonious way to include high-resolution terrain information in global ESM simulations while limiting the number of sub–grid element employed. For each terrain tile, characterized by homogeneous terrain properties, here we develop an average correction to the downward solar fluxes to account for the effects of local topography. This approach thus bridges the gap between the scale at which radiation and other physical processes are represented in the GFDL ESM, ans thus allows to study how the sub-grid heterogeneity of these processes impacts the long term evolution of the coupled physical system. In the following, we present this methodology and test it over three mountainous sites located in different geographic regions (Alps, Andes, Himalayas), showing how model resolution and number of tiles impact the performance of the methodology. Given the wide range of spatial scale involved in the description of sub–grid topography, we evaluate the performance of the parameterization across scales, focusing in particular on the possible nonlinear dependence of incident radiation on topographic features, and validating the model over independent sites.

Due to the complex interactions involved, topographic effects on shortwave radiation are generally studied based on Monte Carlo ray tracing techniques, which approximate the three dimensional radiative transfer process by tracking the fate of a large number of photons (Miesch et al., 1999; Chen et al., 2006; Mayer et al., 2010). Here we employ a Monte Carlo ray-tracing technique for describing the true physics of radiation–topography interactions. These high-resolution simulations are used to calibrate the predictive models for topographic effects over each tile, and to validate the proposed parameterization.

The paper is organized as follows: we first review the existing parameterization for radiation over rugged terrain (Lee et al., 2011), and tailor it to our problem. This in turn requires training a model to predict topography-driven corrections for radiative fluxes based on terrain properties. Finally, we present the clustering algorithm used to divide the domain study in tiles, so as to compute local flux corrections over homogeneous regions.

The approach is then validated using different domains for independent training and testing of the methodology (Sec. 3.2), exploring the effects of terrain resolution and possible consequences of nonlinear radiation-topography interactions. Finally, we explore how different tiling structures with increasing resolution improve the representation of the spatially varying radi-

90 ation fields over mountains (Sec. 3.4). We close by discussing assumptions and limitations of the proposed methodology, and suggesting future developments.

2   Methods

In order to properly account for the effects of land heterogeneity on the radiative transfer in ESMs, key variables must be obtained from high-resolution terrain data sets and properly summarized in order to capture their fine-scale effects on shortwave radiation fluxes. To this end, here we start by defining the radiative and terrain variables used to predict radiation over 3D

95 topography. Then, we describe (i) the Monte Carlo ray tracing algorithm used for training and testing the predictive model, (ii) the terrain clustering algorithm used to classify land units based on the local topographic effects on radiation fluxes, and (iii) the predictive models used to link terrain properties to radiative fluxes in each land cluster. Together, these three steps provide a framework for computing fine-scale corrections to the shortwave radiation received by sub–grid land units in the GFDL land model.

100   2.1   Characterizing shortwave radiation over mountainous terrain

A parameterization explicitly accounting for the effects of 3D topography on the shortwave radiation budget was proposed by Lee et al. (2011) based on the results of Monte Carlo photon tracing simulations (Chen et al., 2006). Here we adopt a similar approach and, following the formalism introduced by Lee et al. (2011), classify the shortwave radiation incident at a target point at the surface into 5 distinct components: The direct and diffuse downward solar fluxes (F_{dir} and F_{dif}), and their terrain-

105 reflected counterparts (F_{rdir}, F_{rdif}), which represent respectively direct beam or diffuse photons reaching the target site after a single reflection from neighbouring terrain. Finally, a coupled flux component (F_{coup}) consists of photons first reflected by the surface, and then either back-scattered by the atmosphere or reflected multiple times by the surface before reaching the target site. We note that for a flat surface $F_{rdir}, F_{rdif} = 0$ while $F_{dir}, F_{dif}, F_{coup} \neq 0$ in general. We note that in the GFDL land model, diffuse radiation received by the (flat) surface corresponds here to the sum of F_{dif} and F_{coup}. Based on this formalism,

110 the normalized flux differences between the traditional plane-parallel (PP) case and the topography-aware case (3D) are the quantities object of our analysis which can be used to correct the shortwave radiative balance in ESMs. Following Lee et al. (2011), these quantities are expressed as

$$f_{dir} = \frac{F_{dir}^{(3D)} - F_{dir}^{(PP)}}{F_{dir}^{(PP)}}; \quad f_{dif} = \frac{F_{dif}^{(3D)} - F_{dif}^{(PP)}}{F_{dif}^{(PP)}}; \quad f_{rdir} = \frac{F_{rdir}^{(3D)}}{F_{dir}^{(PP)}}; \quad f_{rdif} = \frac{F_{rdif}^{(3D)}}{F_{dif}^{(PP)}}; \quad f_{coup} = \frac{F_{coup}^{(3D)} - F_{coup}^{(PP)}}{F_{coup}^{(PP)}} \qquad (1)$$

where the reflected-direct and -diffuse components are normalized with respect to the corresponding non-reflected flux

115 component since they are equal to zero in the plane-parallel case (Lee et al., 2011). A schematic representation of these flux components is reported in Figure 1. Predicting these five f_i terms over land tiles representing heterogeneous terrain properties is the objective here. To this end, a predictive model linking the f_i's to tile terrain properties is necessary. To train such a model we use ray tracing simulations which are discussed next.
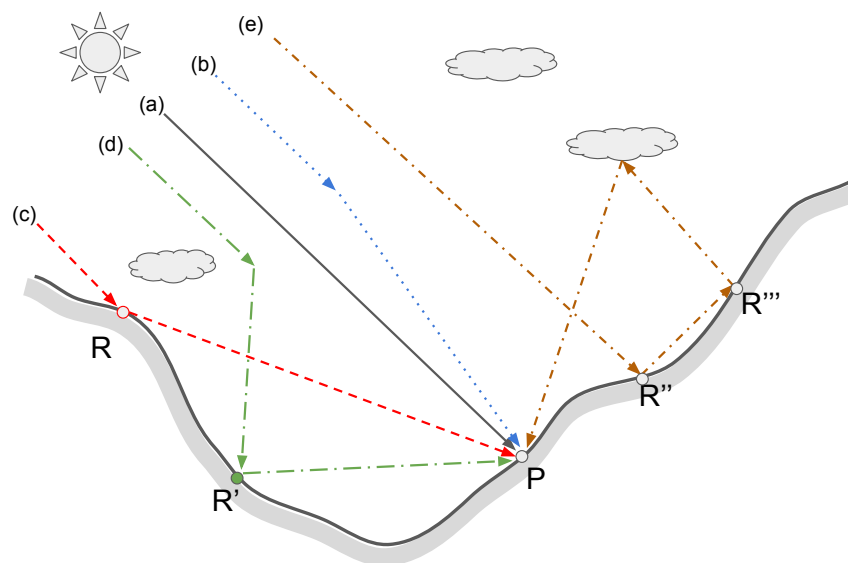
Figure 1. Conceptual representation of the five flux components used to characterize the nature of downward shortwave fluxes over rugged terrain, following the formalism used by Chen et al. (2006) and Lee et al. (2011). The figure includes the direct radiation flux (a), and diffuse radiation (b), consisting of photons which are absorbed at the surface target P after undergoing atmospheric scattering. Direct-reflected (c), and diffuse-reflected (d) fluxes represent photons which are reflected once at the surface. Finally, the coupled (e) flux component include light undergoing multiple reflections at the surface, or reflection at the surface and then atmospheric scattering. For all components, the figure shows paths incident at a point P at the surface.

2.2 A Monte Carlo Ray tracing algorithm

120 In order to develop a predictive model for the correction terms and to test the proposed methodology, we employed a 3D Monte Carlo (MC) algorithm (sometimes referred to as photon tracing or ray tracing algorithm). The MC scheme explicitly describes the interaction of downwelling shortwave radiation with a 3D surface corresponding to a region characterized by complex topography. The algorithm was implemented in a software package which is made available in the online supplementary material. The MC method has been widely used to study radiation interaction with 3D surfaces chen2006application,

125 lee2011parameterization, mayer2009radiative, mayer2010validating, villefranque2019path, larue2020snow. The MC model adopted here broadly follows previous models developed by Chen et al. (2006) and Mayer (2009).

In our MC algorithm, photons are randomly released at the top of atmosphere (TOA) and travel in a direction determined by the Sun's zenith θ_0 and azimuth ϕ_0 angles. After a path of random length, which depends on the optical properties of the medium, the photon encounters scattering or absorption based on the single scattering albedo properties of the atmospheric

130 constituents (Fu and Liou, 1992; Liou, 2002). In the present simulation, we used for each site optical properties computed from the RRTMGP radiation code (Pincus et al., 2019) using the GFDL AM4 model (Zhao et al., 2018). We note that the Monte
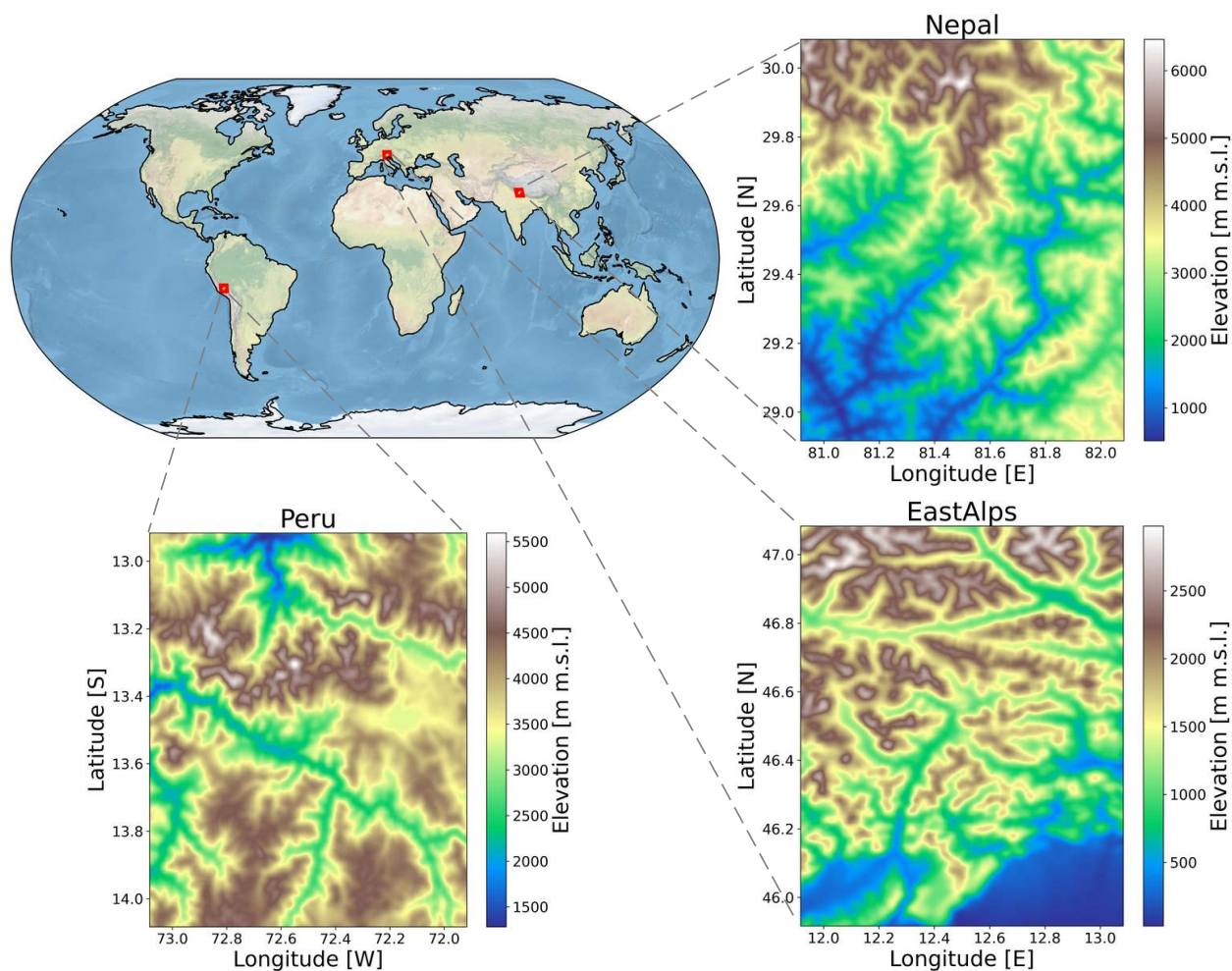
Figure 2. Location of the three land domains selected for this study.

Carlo model is run offline, prescribing the atmospheric profile and solar position in each model run. The atmospheric column used in the model is composed of 34 layers characterized by optical properties which encode the absorption and Rayleigh scattering of photons by gas molecules. We limit our analysis to aerosol-free and clear sky conditions, similar to previous studies (Chen et al., 2006; Lee et al., 2011). Therefore, atmospheric properties are completely determined by optical depths and single scattering albedo at each level (Liou, 2002). The lower boundary of the simulation domain is derived from high-resolution (90 m) terrain model derived from the Shuttle Radar Topography Mission (SRTM) dataset (Farr et al., 2007). SRTM elevation fields are used to construct a three–dimensional surface characterized by triangular mesh elements characterized by uniform albedo and Lambertian reflection. In the following we refer to "pixel" as the image elements of the high resolution
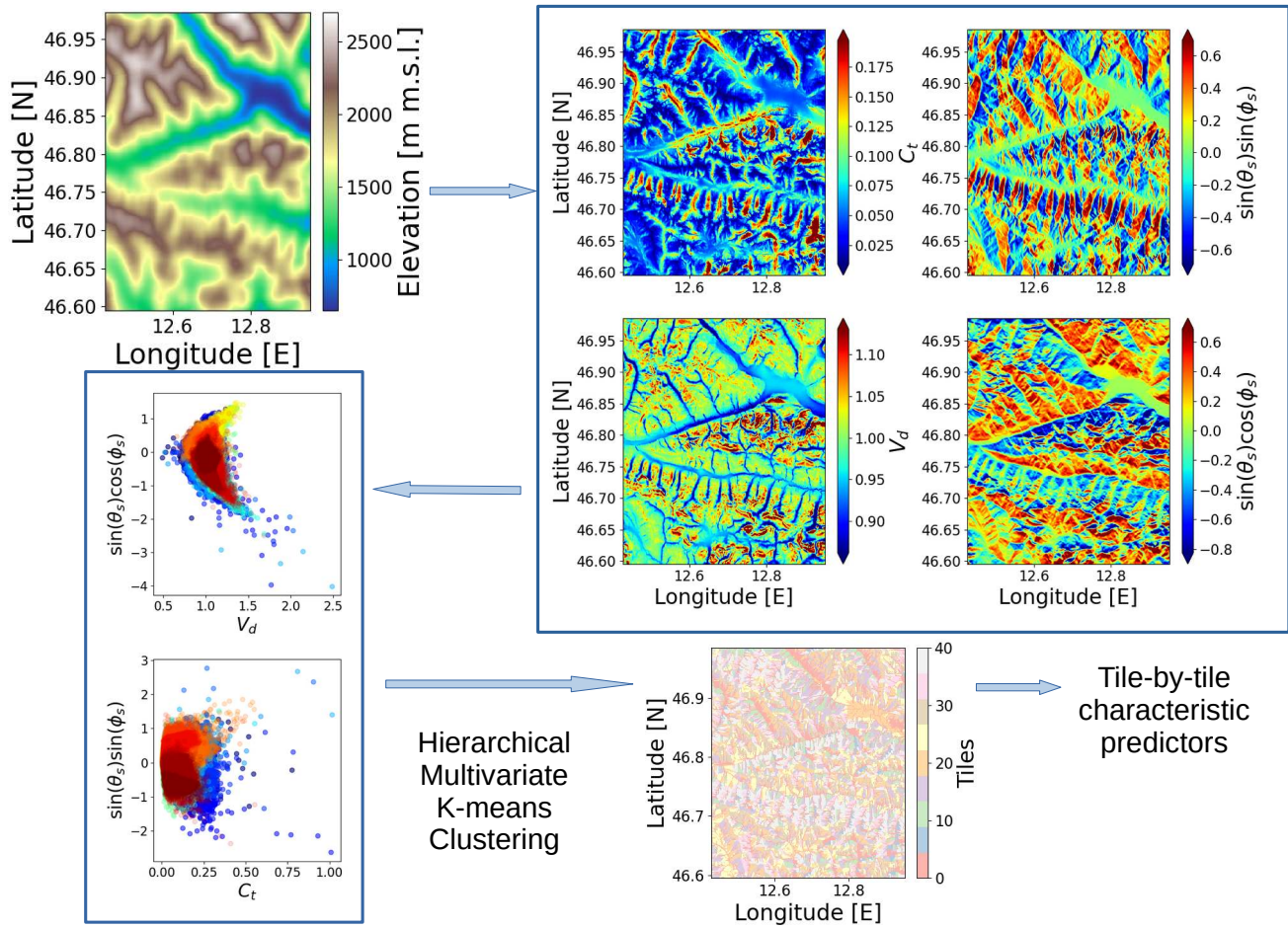
Figure 3. Schematic representation of the land clustering workflow. Land elevation maps are used to compute the variables of interest (sky and terrain view, hillslope orientation) which are then used in the hierarchical clustering step, which yields a map of homogeneous land units used to parameterize radiation-topography interactions.

input digital elevation maps, to clarify the difference with land model grid cell and land model sub-grid units, termed "tiles" in the following.

In each MC simulation, photon are traced from the TOA until they are absorbed or leave the simulation domain. For each photon, interactions with atmospheric constituents (scattering or absorption) and with land surface elements (absorption or reflection) are used to characterize the nature of radiation incident over surface elements. Tracking the path of each photon, the downward irradiance is then decomposed into the 5 flux terms introduced in Section 2.1. If E_0 is the radiation incident at the TOA with a cosine of the zenith angle $\mu_0 = \cos\theta_0$, then the irradiance received by a grid cell of the terrain surface model is given by (Mayer, 2009)

$$E_{k,l} = E_0 \cos\theta_0 \frac{1}{N} \frac{A}{A_{k,l}} \sum_{i=1}^{N_{k,l}^{(s)}} w_i \qquad (2)$$

where w_i is the energy of the i-th incident photon, N is the total number of photons traced in the simulation over a domain with
150   area A, and $N_{k,l}^{(s)}$ is the number of photons absorbed by the surface within grid cell (k,l) with area $A_{k,l}$. Photon are released
with unit energy at the TOA, and lose a fraction of this energy continuously through absorption in each atmospheric layer
(Mayer, 2009), or through absorption at the ground.

For two independent domains located in the Alps and Peru (2), the MC calculations were repeated for 6 solar zenith angles
$(\cos\theta_0 \in \{0.1, 0.25, 0.40.55, 0.7, 0.85\})$, 4 solar azimuth angles $(\phi_0 \in \{0, \pi/2, \pi, 3\pi/2\})$, and a uniform surface reflectivity
155   value set to $\alpha_s = 0.3$.

2.3 Predicting radiative fluxes over complex terrain

Over mountainous regions, the complexity of topography and surface properties determines significant departures in the inci-
dent radiation fluxes compared to their areal-average properties. In order to develop a simple parameterization to explain the
magnitudes of these departures over mountains areas, we need predictor variables encoding the interaction between downward
160   radiation and topographic features. To this purpose, we define here a set of relevant variables following previous work by
Chen et al. (2006) and Lee et al. (2011). The terrain variables used to predict downward fluxes are i) the sky view factors V_d,
which represent the fraction of the sky dome visible from a target site, ii) the terrain configuration C_t, representing the area of
surrounding terrain directly visible; iii) the solar incident angle μ_i, i.e., the angle between the direct solar beam and the normal
to the surface. These terrain variables are derived from the Shuttle Radar Topography Mission (SRTM, (Farr et al., 2007))
165   high-resolution (90m) terrain information.

In order to compute sky view factor, we use the rapid procedure proposed by Dozier and Frew (1990), whereby the unob-
structed fraction of sky hemisphere is approximated as

$$V_d \simeq \frac{1}{2\pi} \int_0^{2\pi} \left[cos\theta_s sin^2 H_\phi + sin\theta_s cos(\phi - \phi_s)(H_\phi - sinH_\phi cosH_\phi) \right] d\phi \qquad (3)$$

for a point with slope θ_s, aspect ϕ_s and horizon angle H_ϕ (i.e., angular distance between zenith and local horizon along the
170   generic azimuth direction ϕ). The terrain configuration, which quantifies the surface of surrounding slopes in direct sight, can
then be obtained as $C_t \simeq (1 + cos\theta_s)/2 - V_d$.

While fields of V_d and C_t can be computed offline for any given elevation map, the sun's incidence angle on a surface does
depend on the sun's position through the local zenith and azimuth angles (θ_0, ϕ_0)

$$\mu_i/cos\theta_s = cos\theta_0 + \sin\theta_s \tan\theta_0 cos(\phi_s - \phi_0) = \mu_0 + \sin\theta_0 (S_{sa} \sin\phi_0 + C_{sa} \cos\phi_0) \qquad (4)$$

175 where topographic information is encoded in the two terms $S_{sa} = \sin\theta_s \sin\phi_s / cos\theta_s$ and $C_{sa} = \sin\theta_s \cos\phi_s / cos\theta_s$. Additionally, for parameterizing the effect of topography on diffuse radiation we use a standardized elevation $h_n = (h - \mu_h)\sigma_h$ obtained normalizing elevation based on its grid cell average μ_h and standard deviation σ_h.

For the purpose of parameterizing solar fluxes, we divide these terrain variables by the local terrain slope obtaining the normalized variables $\tilde{\mu}_i = \mu_i / \cos\theta_s$, $\tilde{V}_d = V_d / \cos\theta_s$, and $\tilde{C}_t = C_t / \cos\theta_s$ as recommended by (Lee et al., 2013). In previous
180 work (Lee et al., 2011) these terrain predictors are averaged over an area representative of an entire model grid-cell and used as predictors to derive average correction terms f_i for the 5 flux components introduced in Section 2.1. Here, high resolution fields of these predictor variables are first computed based on the original 90m digital elevation maps. These fields are then employed to inform the partitioning of the land domain into a set of tiles for which we expect topography to have a similar effect on radiation. This step is achieved using a hierarchical clustering methodology described next. Based on this partitioning,
185 relations to predict the topographic effects on radiation (the f_i's) will be applied to each tile independently.

2.4 Hierarchical clustering of terrain properties

In order to capture the spatial variability of radiative fluxes, here we employ a hierarchical multivariate clustering approach (HMC) which was recently introduced to study the role of heterogeneity in hydrological and land models (Chaney et al., 2016, 2018). Here we tailor HMC to the case of shortwave radiative fluxes by performing the land clustering based on terrain
190 properties (namely $\tilde{\mu}_i$, $\tilde{V}_d$, and $\tilde{C}_t$) which are known to modulate the downwelling radiation over mountains as discussed in the previous section.

The land fraction of the study sites, which are chosen to represents in size a typical ESM grid cell, is first divided in a maximum of three components: soil, glacier, lake. The soil fraction is then subdivided into a set of *tiles* characterized by homogeneous terrain properties relevant for capturing the effects of topography on radiative transfer. Additionally, lake and
195 glacier areas, where present, are treated as individual separate tiles. Here the land clustering is based on 4 variables: The normalized sky view factor, terrain configuration, S_{sa}, and C_{sa}. Note that these variables are independent of the sun's position. Once the direction of the incoming beam is given (ϕ_0 and $\cos\theta_0$), average values of S_{sa} and C_{sa} over any given tile uniquely identify the solar incident angle for each point on the land surface by means of eq. 6.

In a first step, based on this set of variables, the land component of the domain is first divided in characteristic hillslope
200 elements. These are obtained by first delineating catchments based on a threshold area of $1 \times 10^5 \, \mathrm{m}^2$ and by dividing each basin in up to three contiguous hillslope elements, corresponding to left side, right side, and headwaters. Then, hillslope elements are aggregated in k *"characteristic hillslopes"* via k-means clustering (MacQueen et al., 1967) in the 4-dimensional space of the variables $\tilde{V}_d$, $\tilde{C}_t$, S_{sa}, and C_{sa}; this enables to obtain land units characterized by similar radiation-topography interaction. Then, each land unit so obtained is further partitioned into p sub-units by a second application of the k-means
205 clustering based on the 4 variables $\tilde{V}_d$, $\tilde{C}_t$, S_{sa}, and C_{sa}. A conceptual summary of this procedure is described in Figure 3. Therefore, in the current configuration the number of tiles (n_t) used to describe land heterogeneity within a single land model grid cell vary between kp and $kp + 2$ depending on the presence of lake and glacier units. The reason for the first step (subdivision of land in characteristic hillslopes) originates from the desire of partitioning land in hydrologically coherent units.

In the context of shortwave radiation received by land, it is important to understand what is the effect and the benefits of this multilevel clustering. Therefore, we will test the methodology for different tile configurations over the same domains. First we will vary only the number of characteristic hillslopes obtained with the first stage of the clustering approach (i.e., varying k keeping p fixed), and then repeat the calculations fixing k and varying the number p of sub-units within each characteristic hillslope.

Therefore, based on this procedure, the land model grid cells are subdivided into a number of tiles n_t. A maximum of two additional tiles can be included representing lake and a glacier areas, if these present in a given grid cell. Glacier boundaries are determined using the GLIMS database (Raup et al., 2007). Each tile is characterized by statistically homogeneous values of the variables of interest for 3D radiative transfer. Over each tile $t = 1, \ldots, n_t$ we then compute the average value of each predictor variable. For a generic variable $\Theta \in \left\{ \tilde{V}_d, \tilde{C}_t, S_{sa}, C_{sa} \right\}$ we have

$$\langle \Theta \rangle_t = \frac{\sum_{i=1}^{n_p} A_i \Theta_i \mathbb{1}_{t,i}}{\sum_{i=1}^{n_p} A_i \mathbb{1}_{t,i}} \tag{5}$$

with the $< \cdot >_t$ operator representing the operation of average over the points classified as part of tile t, $\mathbb{1}_t$ the is the indicator function selecting tile t ($\mathbb{1}_{t,i} = 1$ if pixel i of the high resolution terrain map belongs to tile t, $\mathbb{1}_{t,i} = 0$ otherwise). A_i and Θ_i are area and the value of the property Θ computed for the i-th pixel in a high resolution terrain map with n_p pixels. For the solar incidence angle, once the solar position is known we can write its average value over a tile t as

$$\langle \tilde{\mu}_i \rangle_t = \mu_0 + \sin \theta_0 \left(\langle S_{sa} \rangle_t \sin \phi_0 + \langle C_{sa} \rangle_t \cos \phi_0 \right) \tag{6}$$

Therefore, the average values of these quantities over each tile represent the 'characteristic' value of the properties over that tile and can be used to summarize the effect of multi-scale radiative transfer over mountainous regions.

2.5 A predictive model for the flux correction terms

The final step needed in order to parameterize the effect of complex topography on incident shortwave radiation is specifying a predictive model to link the terrain variables defined in Section 2.3 to the 5 flux components (Sec. 2.1). This step will enable to extend the results of costly ray tracing simulations at the global scale, and provide predictive equations which can be directly applied to high resolution terrain maps as well as to the local averages from the tiling structure defined in Sec. 2.4.

Here we focus on two different approaches, namely a multiple linear regression model (MLR, based on the previous WLH parameterization) and on an alternative approach based on random forest regression (RFR). We explore these two distinct families of statistical models in order to evaluate the potential role on nonlinearity for different radiation flux terms, keeping in mind that MLR, the simpler statistical model, is in general to be preferred due to the greater ease of interpretation as well as the reduced computational expense when employed in Earth System Models.

MC photon tracing simulations are performed at the native 90 m resolution of terrain products available. In the following analysis, results are coarsened at a range of spatial scales up to 10 km in order to test the robustness of the proposed parameterization to the spatial scale of the terrain variable used as predictors, and explore potential dependencies of radiation-topography

interaction on the spatial scale considered. In each case, both MC simulation results and terrain properties are averaged at the same spatial resolution in order to train and test the predictive model at that specific spatial scale. MC simulations corresponding to the 3D case (i.e., over a three dimensional surface) and over a flat surface (plane parallel case, PP) were subtracted in order to compare the "true" flux deviations as defined in eq. (1). Additionally, given the finite extension of the domain used in the MC simulations, spurious effects may be present due to the periodic boundary conditions imposed, especially for the lower

245   values of solar zenith angle. To mitigate this issue, only the central part of each simulation domain (of size $\sim 1° \times 1°$) is used to train the predictive model; A fraction of size 0.2 times the linear dimension of the domain in discarded at each boundary.

The MC simulations were repeated for varying zenith angles (from $\mu_0 = 0.1$ to $\mu_0 = 0.85$) and 4 azimuth angles. The results for each azimuth are pulled together so that a single predictive model for each zenith angle is derived. For application to ESMs, the results should be interpolated when predictions for a generic zenith angle are needed.

250   After training each model for each solar zenith angle and flux component, we obtain a set of predictive equations linking the f_i's (i.e., the normalized deviation between 3D and PP case) to the terrain predictors at the same spatial scale:

$$f_{dir} = g_{dir,m,\cos\theta_0,\Delta x}\left(\langle \tilde{V}_d \rangle_{\Delta x}, \langle \tilde{\mu}_i \rangle_{\Delta x}\right) \tag{7}$$

$$f_{dif} = g_{dif,m,\cos\theta_0,\Delta x}\left(\langle \tilde{V}_d \rangle_{\Delta x}, \langle \tilde{\mu}_i \rangle_{\Delta x}, \langle h_n \rangle_{\Delta x}\right) \tag{8}$$

$$f_{rdir} = g_{rdir,m,\cos\theta_0,\Delta x}\left(\langle \tilde{V}_d \rangle_{\Delta x}, \langle \tilde{C}_t \rangle_{\Delta x}, \langle \tilde{\mu}_i \rangle_{\Delta x}\right) \tag{9}$$

255    $$f_{rdif} = g_{rdif,m,\cos\theta_0,\Delta x}\left(\langle \tilde{V}_d \rangle_{\Delta x}, \langle \tilde{C}_t \rangle_{\Delta x}\right) \tag{10}$$

$$f_{coup} = g_{coup,m,\cos\theta_0,\Delta x}\left(\langle \tilde{V}_d \rangle_{\Delta x}, \langle \tilde{C}_t \rangle_{\Delta x}, \langle \tilde{\mu}_i \rangle_{\Delta x}\right) \tag{11}$$

where g represent the parametric relations linking flux correction terms to terrain variables for different predictive models ($m = MLR$ or $m = RFR$). These equations are derived for each flux component, solar zenith angle $\cos\theta_0$ and spatial scale Δx. We note that these equations were derived for a single surface reflectivity value. The direct and diffuse flux components

260   are independent on albedo; Reflected fluxes are linearly dependent on albedo so that predicted value can be rescaled by the surface albedo. Finally, g_{coup} is nonlinearly dependent on the surface albedo so that predictions for different albedo values can e.g. be obtained by interpolation (Lee et al., 2011). Once the f_i predictions are available, the dimensional value of shortwave fluxes over rugged terrain can finally be obtained from eqns. (1). As noted in (Lee et al., 2015), simply applying a correction to the downward radiation received by land will not ensure energy conservation, as expected in general as some of the 3D

265   effects parameterized here would lead to energy fluxes between neighbouring land model grid cells. To resolve this issue, the procedure proposed by (Lee et al., 2015) can be used, where an albedo correction is devised to ensure energy conservation.

3 Results and Discussion

3.1 Spatial distribution of solar irradiance

A representative output obtained from Monte Carlo ray tracing simulations for the Eastern Alps domain is portrayed in Figure 4,
270 featuring the spatial distribution of differences between 3D topography and PP case for a given incoming solar beam direction,
while fields of the corresponding terrain variables for the same domain are reported in Figure 5.

The direct flux appears prominently modulated by the presence of topography, with the distribution of shaded slopes following that of solar incident angles compute based on the solar angles for the current simulation (5). In the case of the diffuse
flux, differences between the 3D and PP simulations are less apparent, while a similar behavior is observed for the reflected
275 flux components.

From visual inspection of these simulation results, it is apparent that the spatial variation in the direct flux are primarily
controlled by the solar incident angle. The diffuse flux, on the other hand, shows a spatial variability primarily consistent
with that of the sky view factor, as expected based on previous studies (Chen et al., 2006). Reflected and coupled flux spatial
variability appear more complex, potentially controlled by the terrain configuration. This exploratory analysis reveals that, with
280 the sole exception of the direct flux component, variations in downward fluxes are arguably not explained by a single linear
relation, but can in general involve multiple terrain predictors and potentially nonlinear effects. This hypothesis will be tested
in the next section by examining the skill of a linear and nonlinear model in describing these relationships.

For the direct flux, histograms of the normalized differences between PP and 3D simulations are featured in Figure 6, where
the simulation results have been aggregated at a range of increasing spatial averaging scales. For low solar zenith angles, the
285 frequency distribution of f_{dir} tends to have an atom in at -1, lower boundary value corresponding to the case of complete
shade. This is clearly a limitation for the linear model approach, since this behaviour imposes a nonlinear relation between f_{dir}
and terrain variables (solar incident angle and sky view). However, averaging the results at increasing spatial scales we see this
effect progressively diminishes, as the probability of a complete shade decreases. For both the domains examined in Figure 6,
this behavior is similar and similarly decreases with the cosine of the solar zenith angle. We note that at spatial scales larger
290 than 5 and 10 km the effect disappears. We note that it is a this range of scales that previous parameterization of 3D radiation
over mountains were trained (Lee et al., 2011).

3.2 Model evaluation: sensitivity to the spatial scale and role of nonlinear effects

To quantify the sensitivity of the proposed methodology on the resolution of terrain data, we aggregate the MC-simulated
radiation fields and corresponding terrain variables fields at increasing spatial scales ($L_s \in \{0.5, 1, 2, 3, 5, 10\}$ km) and in each
295 case fit the predictive models as described in Sec. 2.5, comparing the relative performance of multiple linear regressor and
random forest predictions.

As shown in Figure 7, the role of nonlinearity is relevant for the direct flux for low solar angles, and for the diffuse and
coupled fluxes. Reflected components are quite linear, as shown by the small difference between random forest and linear
regressors.
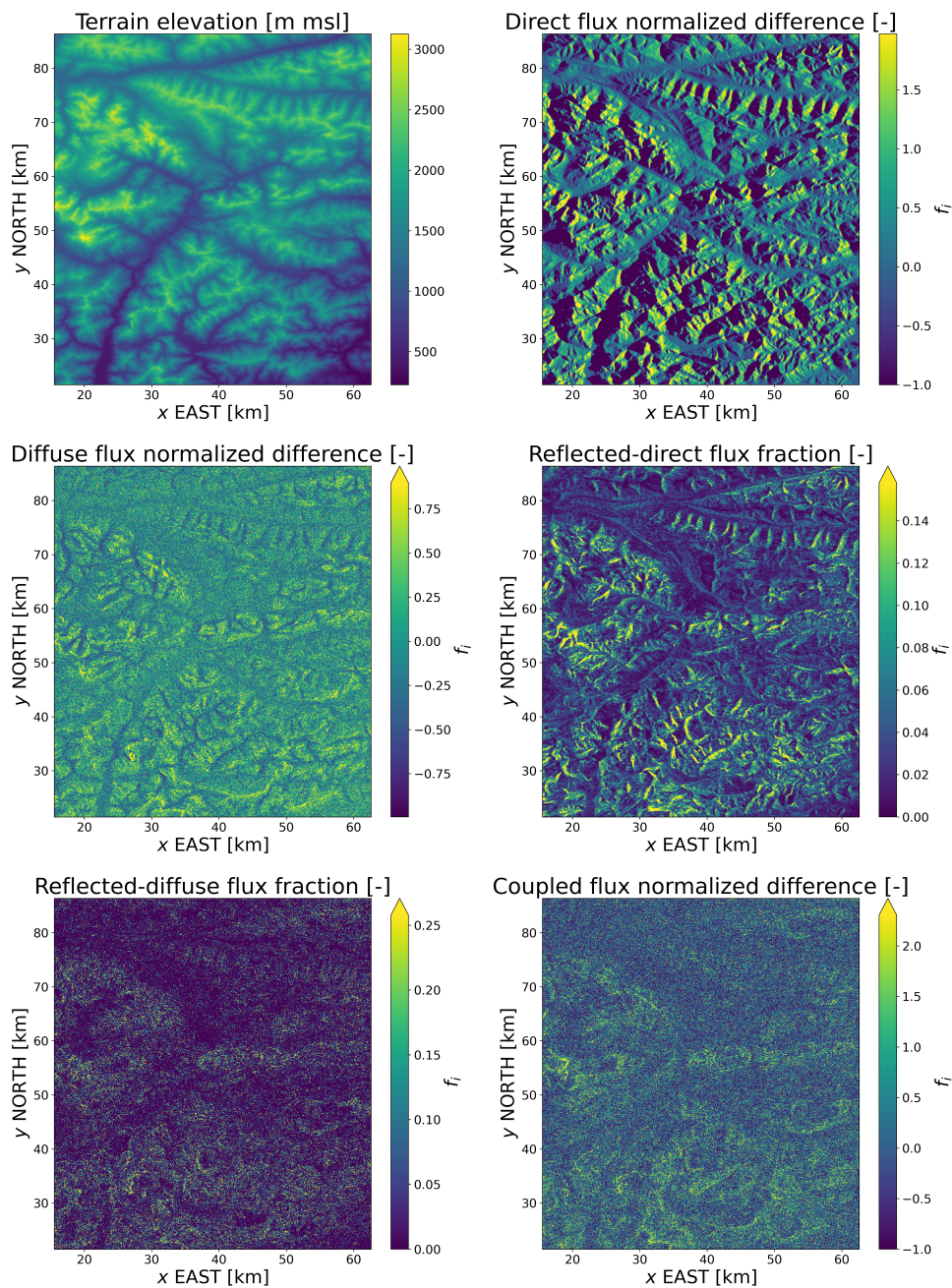
Figure 4. Normalized differences in downward fluxes obtained running Monte Carlo simulations using the real topography and a flat domain, respectively ($\mu_0 = 0.40$, $\phi_0 = \pi/2$).
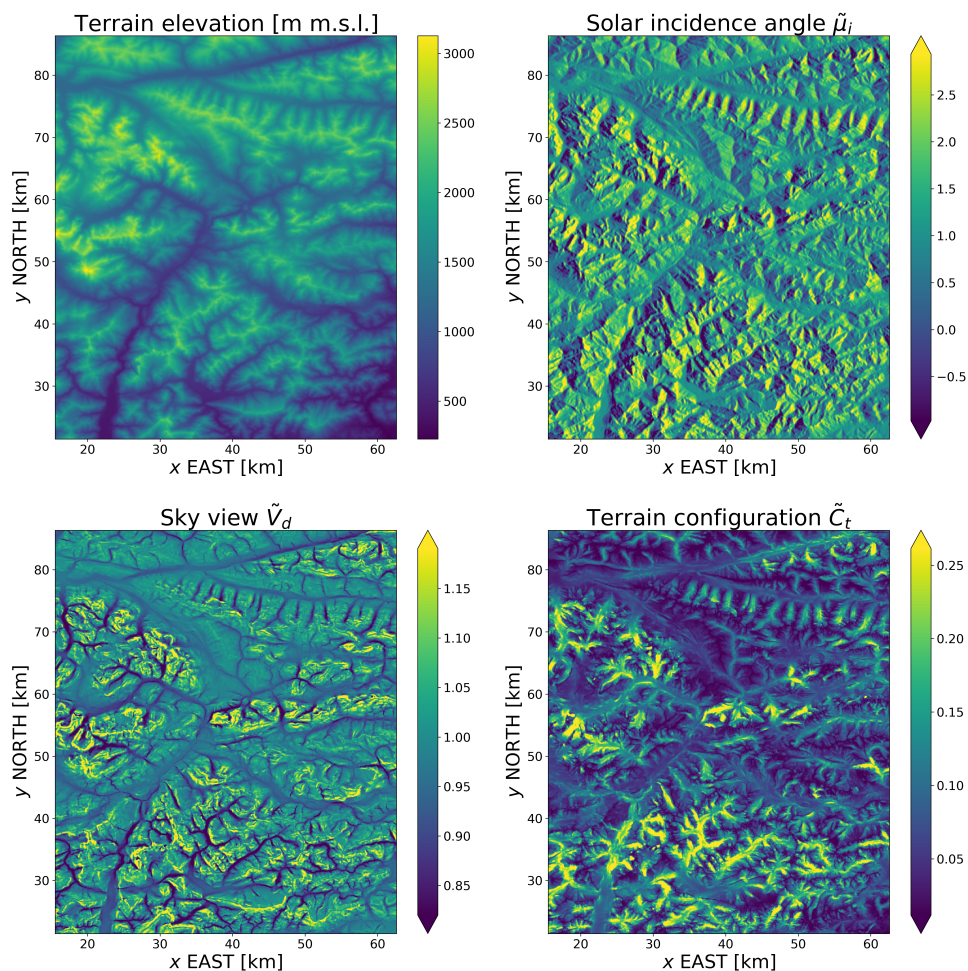
Figure 5. Terrain variables computed for the same domain shown in the previous figure (EastAlps). Solar incidence angle was computed for $\mu_0 = 0.40$, $\phi_0 = \pi/2$.

The predictive model for the direct flux shows high values of coefficients of determination (R^2), with similar performance for RFR and MLR, indicating that a linear model is well suited to describe this quantity. The only discrepancy is observed for very low solar angles ($\mu_0 = 0.1$), case in which RFR outperforms the MLR. We believe this is primarily due to the effect of completely shaded areas in the domain, which are characterized by sharp transitions better described by ensemble of decision tree due to the nonlinear behavior. We note that this effect is relevant only for comparatively small spatial scales ($L_s < 3km$) and low solar angles. In other cases, a MLR model describes the direct flux with great accuracy, case in which both models exhibit lower performance. Moreover, model performance at the lowet solar angles is mitigated by the fact that these conditions
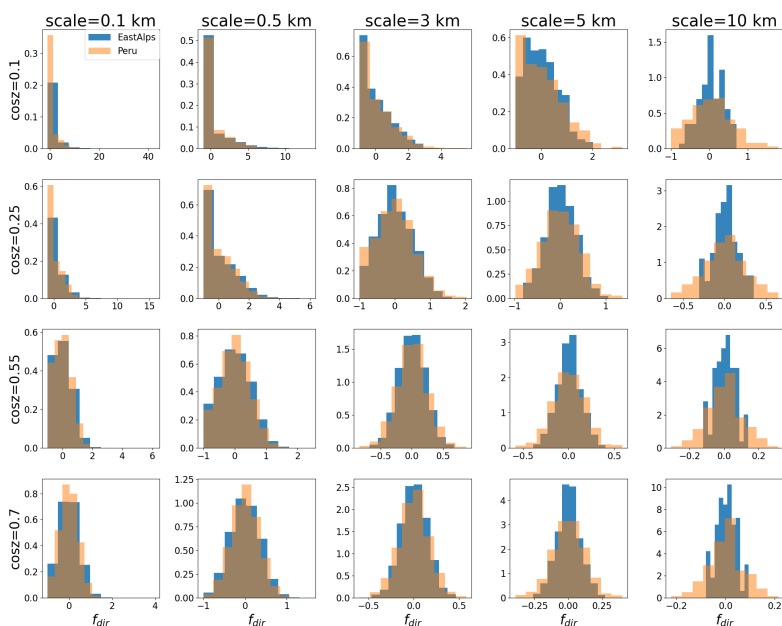
Figure 6. Histograms of the direct radiation correction f_{dir} between the 3D and PP cases, for varying zenith angle and spatial averaging scale, for the Peru and EastAlps domain.

(e.g., dusk and down) generally account for a small fraction of the irradiance received by land over most geographic locations and time of the year.

Reflected-direct and Reflected-diffuse fluxes also exhibit a clear linear dependence on the terrain predictors, with MLR and RFR having similar R^2 values at all averaging scales and solar angles, with the only difference between the two approaches again appearing for the reflected-direct flux for very small solar angles.

On the other hand, appreciable differences between MLR and RFR are observed for diffuse and coupled fluxes. In these cases, it appears that diffuse radiation is better described by a nonlinear model, as is the case for irradiance originating from multiple reflections at the ground and atmospheric scattering (f_{coup}). In these cases, consistent with previous findings by (Lee et al., 2011), the predictive ability of linear models is lower.

In almost all cases analyzed here, the performance of the models tested over the training site (EastAlps) or over an independent location (Peru) are similar, which indicates that our results are not location dependent or a result of overfitting. This is an important check especially for the RFR method employed here.

Based on these results, we recommend the adoption of the linear models at least for direct and reflected fluxes, given the good performance and model simplicity. Applications of RFR are in principle possible in ESMs and has been shown here to have good predictive performance for this specific problem. However, this comes at the cost of a lower interpretability and based on the present analysis here RFR is not the model of choice, given the limited increase in predictive skill with respect to MLR especially when tested in cross validation.
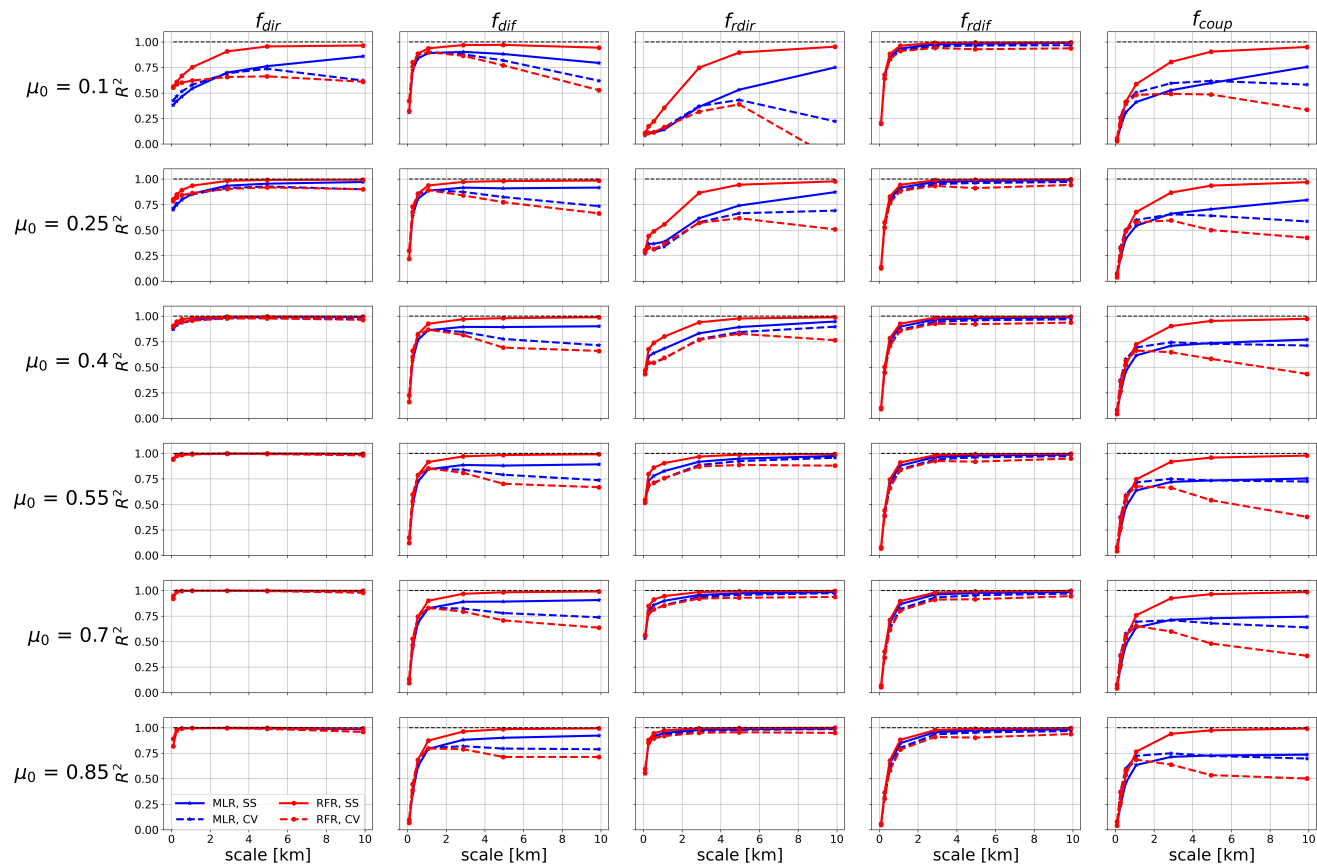
Figure 7. Coefficient of determination R^2 for the two different predictive models: Multiple Linear Regression (MLR, blue lines), and Random Forest Regression (RFR, blue lines). Results are reported for varying averaging scales and cosines of the solar zenith angle ($/mu_0$). The models were trained 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).

3.3 Nonlinearity and effect of averaging for the direct flux

325   The direct flux component is characterized by a nonlinearity in the case of completely shaded areas.

This behavior should be taken into account in our model, as averaging terrain properties over tiles with varying characteristic size and spatial configuration would lead in general to changes in the average predicted f_{dir} for a given tile if model predictions are averaged over areas which, in the high resolution true field, include partial shades. One possible way to capture this behavior is to predict f_{dir} first over an entire grid cell to obtain its average value, and then impose that tile-by-tile predictions must match

330   the average value of the direct flux correction over the entire grid cell ($\langle \tilde{f}_{dir} \rangle$). This can be achieved through the following

transformation for a generic tile $i = 1, ... n_t$. We define the corrected value $\tilde{f}_{dir}^{(i)}$ as

$$\tilde{f}_{dir}^{(i)} = \langle f_{dir} \rangle + \left(\langle \tilde{f}_{dir} \rangle - \tilde{f}_{dir}^{(min)} \right) \frac{f_{dir}^{(i)} - \langle f_{dir} \rangle}{\langle f_{dir} \rangle - f_{dir}^{(min)}} \tag{12}$$

This correction ensures that the $\tilde{f}_{dir}$ values predicted by the model for each tile conserve the grid cell average value, by correcting the original value $\langle f_{dir} \rangle = \sum_{i=1}^{n_t} p_i f_{dir}^{(i)}$, with p_i the fractional area of the grid cell assigned to tile i. This transfor-

335   mation also preserves the minimum value over the grid cells, so that $\tilde{f}_{dir}^{(min)} = f_{dir}^{(min)} = min_{i=1}^{n_t} f_{dir}^{(i)}$. For other flux variable this correction is not necessary if a linear model is used for predicting their average values over tiles, as done here.

Figure 8 shows the frequency distribution of corrections for the various flux terms varying the number of tiles. In particular, for the direct flux, the distribution of tile-by-tile estimates is shown to converge to the histograms of full high resolution results for $n_t = 250$ tiles. Thus, application of the correction given in eq. (12) allows to preserve the grid cell average correction while

340   adequately represeting the sub-grid scale variability of f_{dir}.

3.4 Model sensitivity to the number of tiles used

The predictive model can be used to produce tile-by-tile estimate of flux differences and compare results with those predicted for the original high resolution terrain map. We repeat this analysis for different configurations of the hierarchical clustering scheme, so as to test the sensitivity of the results to the number of tiles used to characterize domain heterogeneity. We expect a

345   larger number of tiles to lead to a better representation of the flux component over rugged terrain. However, a too large number of tiles would not be feasible for running ESMs over large domains or the entire globe.

An example of the spatial distribution of tiles obtained by applying the HMC algorithm is reported in Figure 9 for the EastAlps domain, where the land domain is partitioned into 5 and 25 clusters, respectively. Two additional tiles are used to represent lakes and glaciers, which are present over this domain, albeit accounting for a small fraction of the surface. Once the

350   tiled grid is defined using the HMC method, local averages of terrain parameters can be computed directly over each land unit. For example, the lower panels in Figure 9 show the spatial distributions of the terrain view factor $\tilde{C}_t$) averaged over each tile, showing that larger number of tiles greatly improve the representation of the spatial variability of topography.

A natural test for the ability of the tiled grid of reproducing the actual spatial distribution of solar radiation can be performed as follows: We test here the results for three HMC configurations obtained by varying the number of characteristic hillslopes

355   (k) as well as the number of land units within each characteristic hillslope (p). For illustration purposes, we consider two case: a fixed value of $k = 5$ and a varying p, and the opposite (varying k, setting $p = 5$). This experiments leads to a set of grid configurations with a number of tiles per grid cell varying from 5 to 1000, with different weight given to the first level (partitioning of land in hillslopes) and the second levels, in which each characteristic hillslope is further subdivided in p homogeneous land units contributing to the overall number of tiles n_t. This experiment thus elucidates the relative performance

360   of the two different levels of the hierarchical clustering approach in capturing the spatial heterogeneity of the domain.

Figures 10 and 11 show how increasing the number of land tiles improves the description of the direct flux and diffuse flux components, respectively, over the EastAlps domain. Note that the same disaggregation of the domain in tiles is used for

predicting the distribution of both variables. Even in the case of a fairly low number of tiles (e.g., 5 tiles in panels 10B and 11B) the sub–grid structure is able to capture the main feature of both direct and diffuse radiation fields, even though their variations

365 are known to be controlled by different terrain properties (primarily aspect and sky view, respectively, as can be seen by the spatial distribution of direct and diffuse fluxes in the high resolution results). For both f_{dir} and f_{dif}, the tile-by-tile predictions appear to converge to the original high resolution field when increasing the number of tiles (results for 25 and 250 tiles are shown in lower panels of Figures 10 and 11). Note that this result is quite significant, since for the domain examined here the number of points need to obtain the high resolution field without using a clustering approach would be of the order of 10^6 (for

370 a $1° \times 1°$ grid cell at the native 90m resolution of SRTM data).

To obtain a more quantitative description of the convergence to the high-resolution fields, we show how spatial statistics of the radiation fields (spatial standards deviation σ_x, skewness γ_x and kurtosis ξ_x) vary with increasing tile count with respect to the same statistics computed for the reference high resolution field.

For all the terrain predictors (Figure 12) we find that increasing the number of tiles, the spatial variability of terrain predictors

375 converges to that of original high-resolution fields, as expected. However, this convergence appears to be faster for the solar incidence angle when compared to $\tilde{V}_d$ and $\tilde{C}_t$. Higher order statistics of the spatial fields (skewness and kurtosis) also tend to converge to the high resolution field values, although with a larger variability. While spatial standard deviation is generally used as a metric for assessing spatial variability, examining skewness and kurtosis helps to make sure the entire distribution of tile-by-tile results is converges to the high resolution benchmark, since they better capture asymmetry and extremes in tile

380 values. However, we note that these metrics are not very meaningful for the smallest number of tiles shown in Figure 12) due to the small sample size. However, the fact that for a large enough number of tiles (e.g., $n_t > 20$) these metric appear to converge to the true values increase our confidence that the hierarchical clustering scheme provides a good description of the topography heterogeneity.

For the flux variables (Figure 13) we find a similar behavior, with convergence of the spatial standard deviation being

385 generally faster that that higher order statistics. In this case, the convergence is faster for the direct flux while slower for all other flux component, as expected since F_{dir} is primarily controlled by $\tilde{\mu}_i$ while the other flux components are show relevant dependence on either sky view factor or terrain configurations.

To further analyze the configuration of the tiling structure used, we also tested different tiling configurations obtained by fixing the number of characteristic hillslopes ($k = 5$) and varying the number of lower-level land units in each hillslope (p),

390 or conversely varying k with $p = 5$ fixed. We find that generally convergence is faster using a larger p, i.e., dividing each characteristic slope in a larger number of tiles as opposed to increasing the number of characteristic hillslopes. This is not surprising. However, differences are generally small, and therefore the model proposed appears flexible and can in principle be applied with tiling predefined in order to accommodate for other physical processes as well.
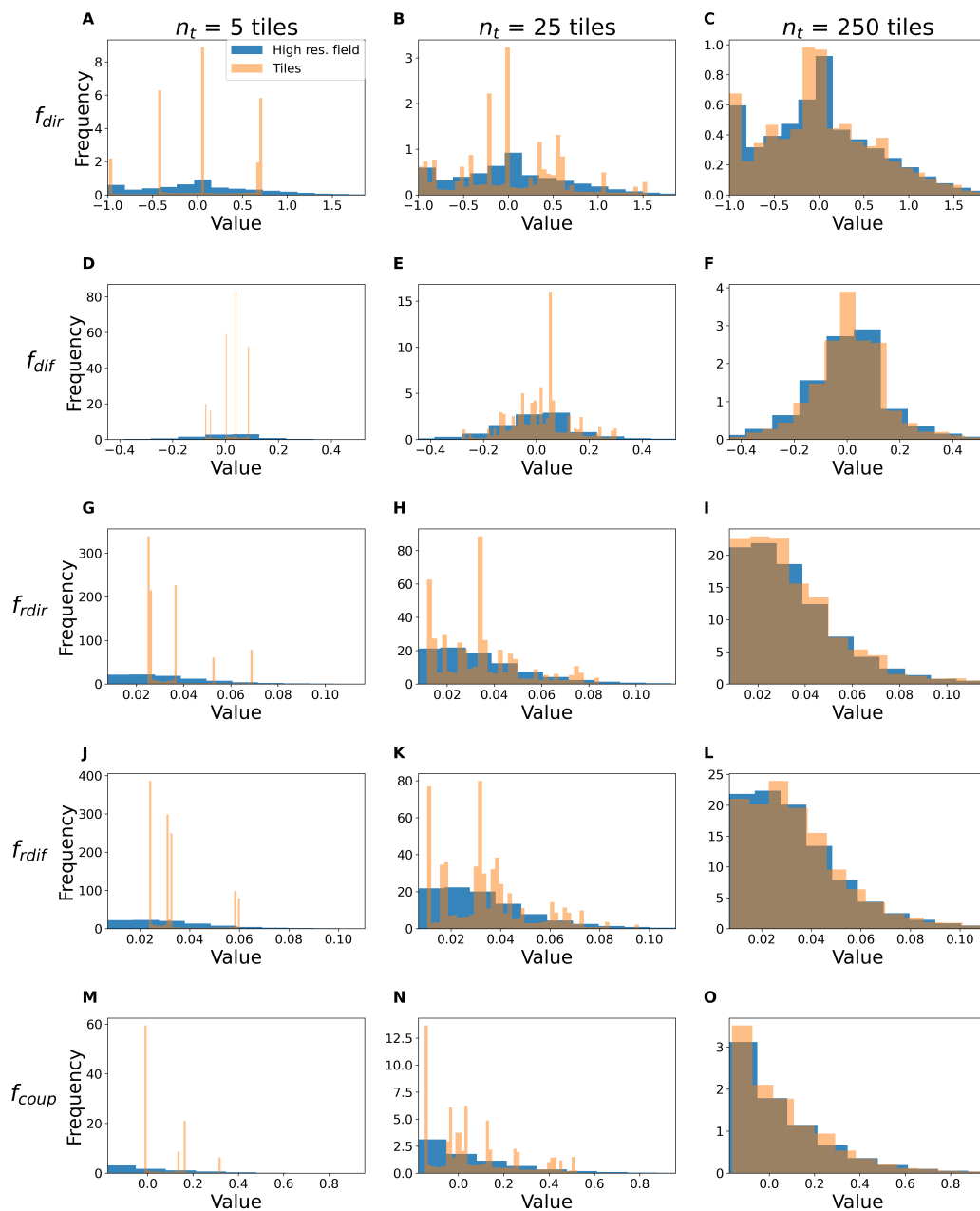
Figure 8. Histograms of the distribution of 3D - PP incident radiation differences for varying number of tiles. Average 3D correction over each tile (orange) are compared with the frequency distribution of the high resolution field (blue histograms). Results are shown for the EastAlps domain, for a given solar angle ($\mu_0 = 0.4$)
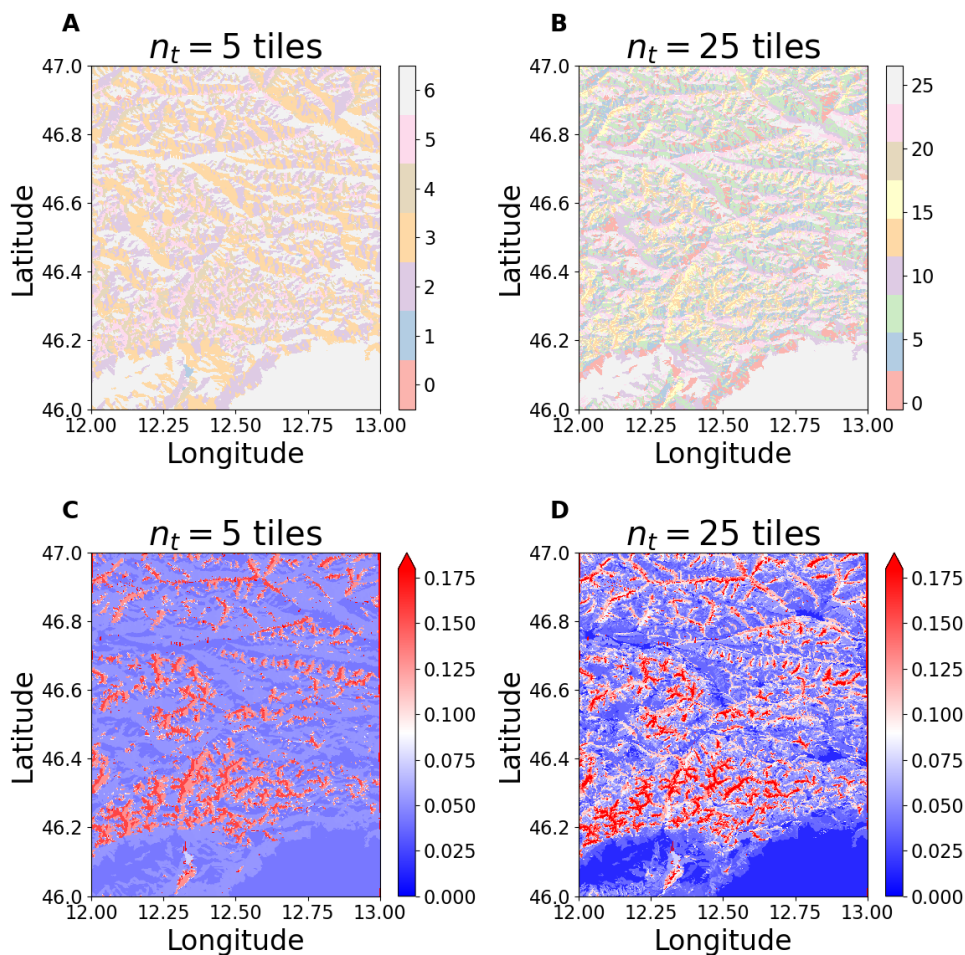
Figure 9. Effect of a different number of tiles for modelling the sub-grid variability of the direct flux over the study site. Top panels: spatial distribution of the tiles representing homogeneous land units. Note that for this domain in addition to the n_t tiles part of land hillslopes there are two additional tiles corresponding respectively to glaciers (tile 0) and lakes (tile 1) so that the effective number of tiles is 7 and 27 for panels (A) and (B). Bottom panel: spatial distribution of the terrain view factor $\tilde{C}_t$ local average evaluate over each tile.

4 Conclusions

395 Here we described a methodology to compute solar fluxes over mountainous terrain, accounting for the sub–grid variability of topographic properties within a characteristic ESM grid cell. Topographic parameters modulating the incident solar irradiance, combined to results from Monte Carlo radiation simulation over 3D surface, were used to train a predictive model and cluster land surface based on topography-radiation interactions. The methodology as presented here is tailored to a tiling structure scheme recently introduced in the GFDL land model LM4.2 to describe the heterogeneity of hydrological properties. However,
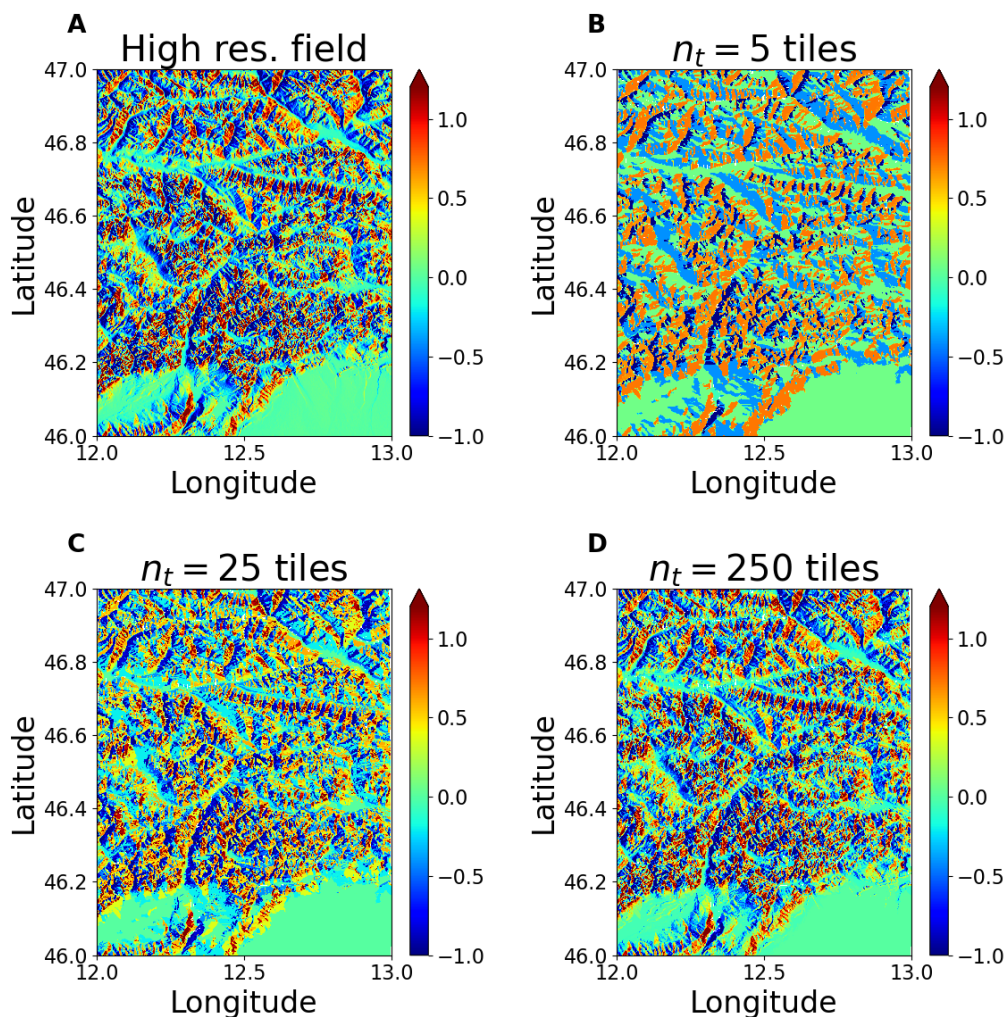
Figure 10. Normalized differences between 3D and PP direct fluxes for high resolution predictions and predictions for different numbers of tiles ($k = 5, 25, 250$), obtained setting a fixed number of hillslopes and varying the number of units for each hillslope ($k = 5$).

we believe this approach can be suitable for applications to others land surface models, as the clustering technique used here allows for a parsimonious description of the spatially varying solar fluxes. For this reason, we tested the sensitivity of the approach to the number of tiles used over independent sites characterized by complex topography. The results appear consistent over different geographical domains, and indicate that even a limited number of tiles can reproduce a significant fraction of the spatial variability observed in the high resolution fields. This result is particularly relevant when compared with standard approaches focused on increasing the land model resolution without adopting a clustering-based approach to construct a sub–grid land structure. Increasing the number of tiles does not only improve the representation of spatial variances, but
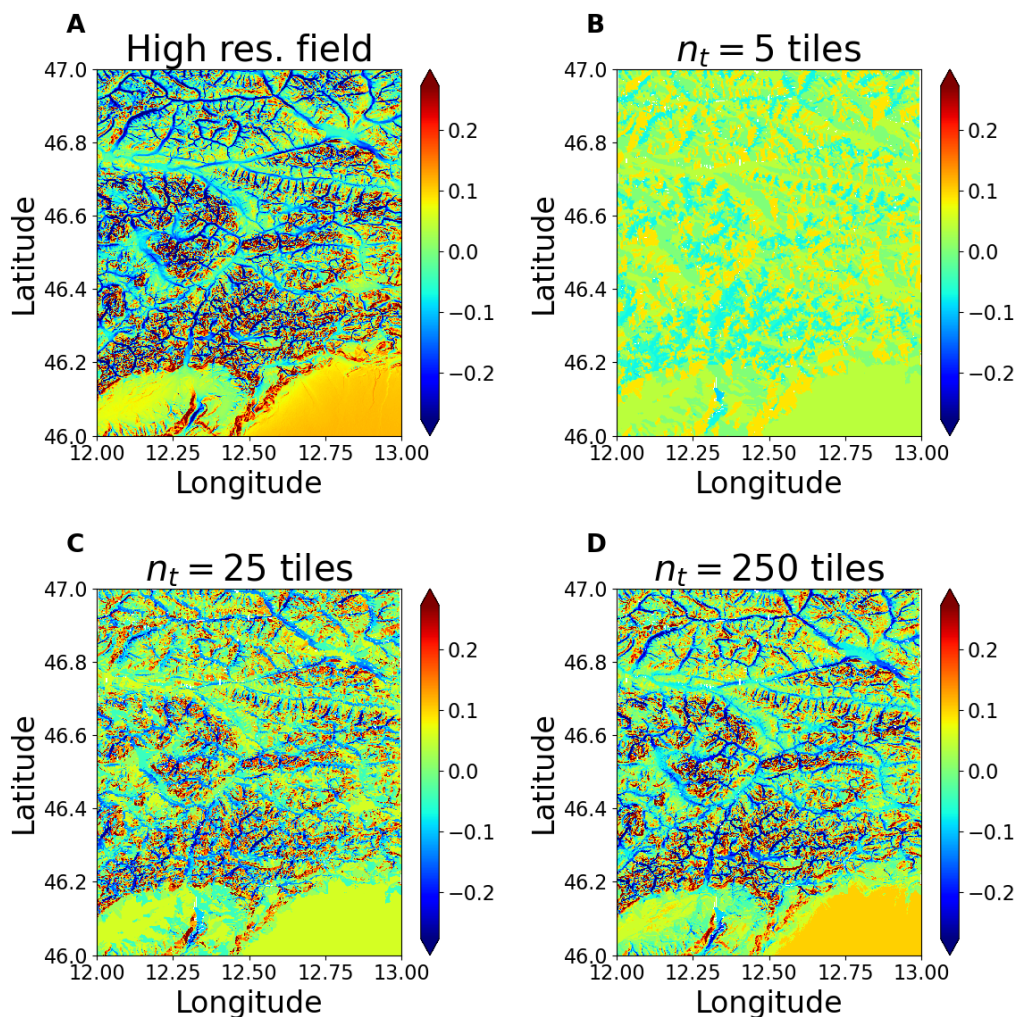
Figure 11. Normalized differences between 3D and PP diffuse fluxes for high resolution predictions and predictions for different numbers of hillslopes ($k = 5, 25, 250$) setting a fixed number of tiles / hillslope ($k = 5$).

also improves convergence of higher order statistics. However, even when using a lower number of tiles, the results remain consistent with previously (i.e., the methodology can ensure that the grid cell average correction is downward radiation is conserved) developed grid-cell average correction and thus are to be considered an improvement with respect to current plane-parallel radiative transfer.

The current methodology, as well as previous studies on the topic (e.g., (Chen et al., 2006; Lee et al., 2011)) employed Monte Carlo simulations based on clear sky conditions. Studying the effects of aerosols and cloud in the radiative transfer over complex terrain remains an open research avenue. In particular, he presence of a spatially varying cloud cover could
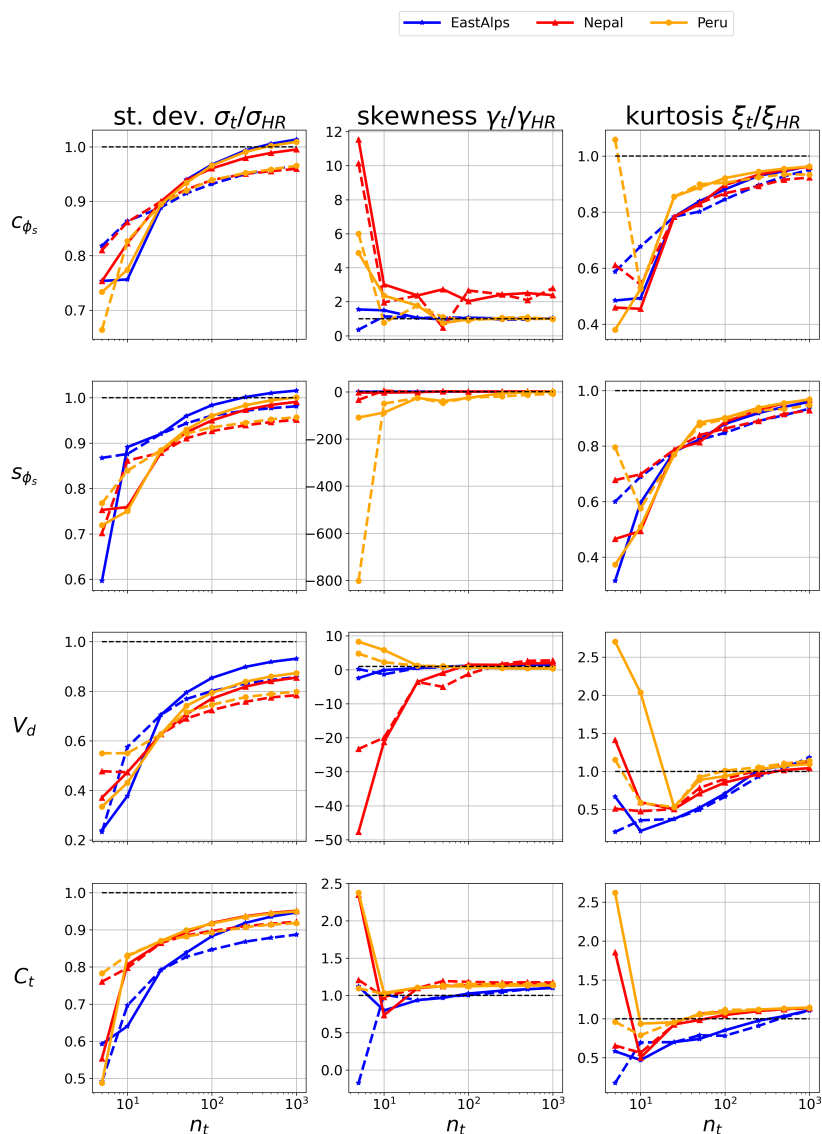
Figure 12. Convergence to the high resolution terrain properties (slope orientation, sky view and terrain view) obtained by increasing the number of tiles used in the parameterization. Two different tiling schemes are shown, respectively obtained fixing the number of hillslopes ($k = 5$) and varying that of sub-grid units p (continuous lines), or fixing $p = 5$ and varying k (dashed lines). Results are reported for terrain variables.
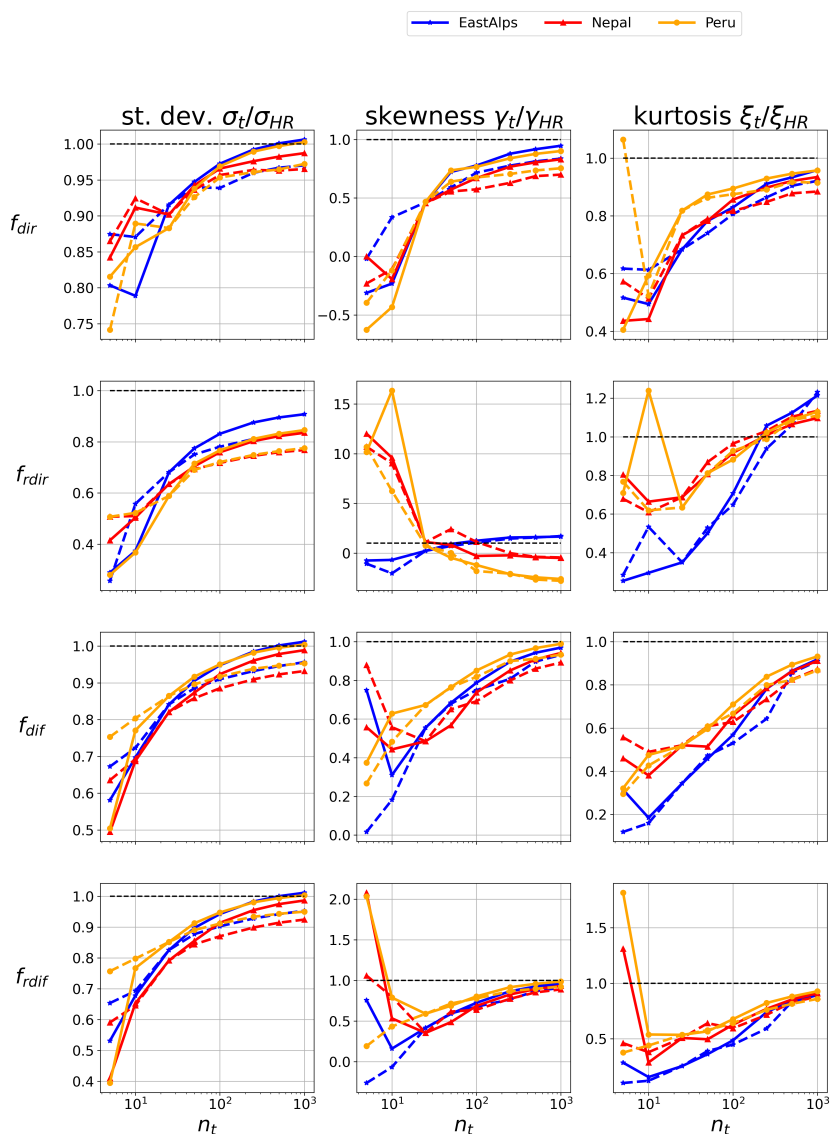
Figure 13. Convergence to the high resolution radiation fields (direct, diffuse, reflected-direct and reflected-diffuse differences between 3D and PP cases). Results are obtained by increasing the number of tiles used in the parameterization. Two different tiling schemes are shown, respectively obtained fixing the number of hillslopes ($k = 5$) and varying that of sub-grid units p (continuous lines), or fixing $p = 5$ and varying k (dashed lines). Results are reported for flux normalized differences.

profoundly influence the spatial distribution of irradiance over mountainous terrain. However, in addition to the numerical
415 challenges connected to the radiative transfer problem, including a spatially varying cloud cover would inevitably increase the
number of parameters needed to parameterize the radiation received by the surface, thus posing a relevant parameterization
challenge.

Studying the relationship between terrain predictors and irradiance differences between 3D and PP cases allowed to quantify
the importance of nonlinear effects and the relative skill of linear models and random forest predictors in capturing these
420 relationships. We found that nonlinear effects are relevant primarily at the finer spatial scales, and decrease drastically with
spatial averages at increasing spatial scales. This results is consistent with previous investigations at coarser spatial scales based
on linear models.

Based on our simulation study over a set mountainous domains, we quantified the difference in downward fluxes originating
from topographic effect with respect with that obtained from traditional plane-parallel radiative transfer schemes for varying
425 number of sub–grid tiles. Our results support the implementation of this methodology in the GFDL ESMs, which will be
pursued as the next step to evaluate the effects of this correction of water and energy fluxes at the surface.

Code and data availability. The code used in this projects are made available during the review process in a zenodo repository with DOI
10.5281/zenodo.7110618.

The preprocessing software used to construct the land database for the GFDL land model v4.2 is included as supplementary material.
430 A Model dataset necessary to run the analysis is available in a zenodo repository with DOI 10.5281/zenodo.6975857

Author contributions. All authors contributed to research design. E.Z. developed software, performed the simulations study, and drafted a
first version of this manuscript. All authors contributed to the manuscript.

Competing interests. The authors declare no competing interests

Acknowledgements. The authors acknowledge funding from the NOAA Climate Program Office (CPO) grant number NA18OAR4320123
435 "3D-Land Energy and Moisture Exchanges: Harnessing High Resolution Terrestrial Information to Refine Atmosphere-to-Land interactions
in Earth System Models". Funding support for this work was also provided through Interagency Agreement Number 80HQTR21T0015 from
NASA Earth Science Division's High Mountain Asia program. Dr. Lucas Harris and Dr. Linjiong Zhou are acknowledged for their comments
and suggestions on an internal review of the manuscript.

References

440 Bales, R. C., Molotch, N. P., Painter, T. H., Dettinger, M. D., Rice, R., and Dozier, J.: Mountain hydrology of the western United States, Water Resources Research, 42, 2006.

Brutsaert, W.: Evaporation into the atmosphere: theory, history and applications, vol. 1, Springer Science & Business Media, 2013.

Chaney, N. W., Metcalfe, P., and Wood, E. F.: HydroBlocks: a field-scale resolving land surface model for application over continental extents, Hydrological Processes, 30, 3543–3559, 2016.

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

Chen, Y., Hall, A., and Liou, K.: Application of three-dimensional solar radiative transfer to mountains, Journal of Geophysical Research: Atmospheres, 111, 2006.

Dozier, J.: A clear-sky spectral solar radiation model for snow-covered mountainous terrain, Water Resources Research, 16, 709–718, 1980.

450 Dozier, J. and Frew, J.: Rapid calculation of terrain parameters for radiation modeling from digital elevation data, IEEE Transactions on Geoscience and Remote Sensing, 28, 963–969, 1990.

Dubayah, R., Dozier, J., and Davis, F. W.: Topographic distribution of clear-sky radiation over the Konza Prairie, Kansas, Water Resources Research, 26, 679–690, 1990.

Duguay, C. R.: Radiation modeling in mountainous terrain review and status, Mountain research and development, pp. 339–357, 1993.

455 Dunne, J., Horowitz, L., Adcroft, A., Ginoux, P., Held, I., John, J., Krasting, J., Malyshev, S., Naik, V., Paulot, F., et al.: The GFDL Earth System Model version 4.1 (GFDL-ESM 4.1): Overall coupled model description and simulation characteristics, Journal of Advances in Modeling Earth Systems, 12, e2019MS002 015, 2020.

Essery, R. and Marks, D.: Scaling and parametrization of clear-sky solar radiation over complex topography, Journal of Geophysical Research: Atmospheres, 112, 2007.

460 Farr, T. G., Rosen, P. A., Caro, E., Crippen, R., Duren, R., Hensley, S., Kobrick, M., Paller, M., Rodriguez, E., Roth, L., et al.: The shuttle radar topography mission, Reviews of geophysics, 45, 2007.

Fu, Q. and Liou, K.: On the correlated k-distribution method for radiative transfer in nonhomogeneous atmospheres, Journal of Atmospheric Sciences, 49, 2139–2156, 1992.

Granger, J. and Schulze, R.: Incoming solar radiation patterns and vegetation response: examples from the Natal Drakensberg, Vegetatio, 35,
465 47–54, 1977.

Gu, L., Baldocchi, D., Verma, S. B., Black, T., Vesala, T., Falge, E. M., and Dowty, P. R.: Advantages of diffuse radiation for terrestrial ecosystem productivity, Journal of Geophysical Research: Atmospheres, 107, ACL–2, 2002.

Gu, Y., Liou, K., Lee, W.-L., and Leung, L.: Simulating 3-D radiative transfer effects over the Sierra Nevada Mountains using WRF, Atmospheric Chemistry and Physics, 12, 9965, 2012.

470 Hao, D., Bisht, G., Gu, Y., Lee, W.-L., Liou, K.-N., and Leung, L. R.: A Parameterization of Sub-grid Topographical Effects on Solar Radiation in the E3SM Land Model (Version 1.0): Implementation and Evaluation Over the Tibetan Plateau, Geoscientific Model Development Discussions, pp. 1–23, 2021.

Hay, J. E. and McKay, D. C.: Estimating solar irradiance on inclined surfaces: a review and assessment of methodologies, International Journal of Solar Energy, 3, 203–240, 1985.

475 Isard, S.: Evaluation of models for predicting insolation on slopes within the Colorado alpine tundra, Solar Energy, 36, 559–564, 1986.

Lamare, M., Dumont, M., Picard, G., Larue, F., Tuzet, F., Delcourt, C., and Arnaud, L.: Simulating optical top-of-atmosphere radiance satellite images over snow-covered rugged terrain, The Cryosphere, 14, 3995–4020, 2020.

Larue, F., Picard, G., Arnaud, L., Ollivier, I., Delcourt, C., Lamare, M., Tuzet, F., Revuelto, J., and Dumont, M.: Snow albedo sensitivity to macroscopic surface roughness using a new ray-tracing model, The Cryosphere, 14, 1651–1672, 2020.

480   Lee, W.-L., Liou, K., and Hall, A.: Parameterization of solar fluxes over mountain surfaces for application to climate models, Journal of Geophysical Research: Atmospheres, 116, 2011.

Lee, W.-L., Liou, K., and Wang, C.-c.: Impact of 3-D topography on surface radiation budget over the Tibetan Plateau, Theoretical and applied climatology, 113, 95–103, 2013.

Lee, W.-L., Gu, Y., Liou, K., Leung, L., and Hsu, H.-H.: A global model simulation for 3-D radiative transfer impact on surface hydrology

485   over the Sierra Nevada and Rocky Mountains., Atmospheric Chemistry and Physics, 15, 2015.

Lee, W.-L., Liou, K.-N., Wang, C.-c., Gu, Y., Hsu, H.-H., and Li, J.-L. F.: Impact of 3-D Radiation-Topography Interactions on Surface Temperature and Energy Budget Over the Tibetan Plateau in Winter, Journal of Geophysical Research: Atmospheres, 124, 1537–1549, 2019.

Lenot, X., Achard, V., and Poutier, L.: SIERRA: A new approach to atmospheric and topographic corrections for hyperspectral imagery,

490   Remote Sensing of Environment, 113, 1664–1677, 2009.

Liou, K., Lee, W.-L., and Hall, A.: Radiative transfer in mountains: Application to the Tibetan Plateau, Geophysical Research Letters, 34, 2007.

Liou, K.-N.: An introduction to atmospheric radiation, Elsevier, 2002.

MacQueen, J. et al.: Some methods for classification and analysis of multivariate observations, in: Proceedings of the fifth Berkeley sympo-

495   sium on mathematical statistics and probability, vol. 1, pp. 281–297, Oakland, CA, USA, 1967.

Mayer, B.: Radiative transfer in the cloudy atmosphere, in: EPJ Web of Conferences, vol. 1, pp. 75–99, EDP Sciences, 2009.

Mayer, B., Hoch, S., and Whiteman, C.: Validating the MYSTIC three-dimensional radiative transfer model with observations from the complex topography of Arizona's Meteor Crater, Atmospheric Chemistry and Physics, 10, 8685–8696, 2010.

McCabe, G. J. and Clark, M. P.: Trends and variability in snowmelt runoff in the western United States, Journal of Hydrometeorology, 6,

500   476–482, 2005.

Miesch, C., Briottet, X., Kerr, Y. H., and Cabot, F.: Monte Carlo approach for solving the radiative transfer equation over mountainous and heterogeneous areas, Applied Optics, 38, 7419–7430, 1999.

Milly, P. C., Malyshev, S. L., Shevliakova, E., Dunne, K. A., Findell, K. L., Gleeson, T., Liang, Z., Phillipps, P., Stouffer, R. J., and Swenson, S.: An enhanced model of land water and energy for global hydrologic and earth-system studies, Journal of Hydrometeorology, 15,

505   1739–1761, 2014.

Picard, G., Dumont, M., Lamare, M., Tuzet, F., Larue, F., Pirazzini, R., and Arnaud, L.: Spectral albedo measurements over snow-covered slopes: theory and slope effect corrections, The Cryosphere, 14, 1497–1517, 2020.

Pincus, R., Mlawer, E. J., and Delamere, J. S.: Balancing accuracy, efficiency, and flexibility in radiation calculations for dynamical models, Journal of Advances in Modeling Earth Systems, 11, 3074–3089, 2019.

510   Raup, B., Racoviteanu, A., Khalsa, S. J. S., Helm, C., Armstrong, R., and Arnaud, Y.: The GLIMS geospatial glacier database: a new tool for studying glacier change, Global and Planetary Change, 56, 101–110, 2007.

Shevliakova, E., Pacala, S. W., Malyshev, S., Hurtt, G. C., Milly, P., Caspersen, J. P., Sentman, L. T., Fisk, J. P., Wirth, C., and Crevoisier, C.: Carbon cycling under 300 years of land use change: Importance of the secondary vegetation sink, Global Biogeochemical Cycles, 23, 2009.

515 Sirguey, P.: Simple correction of multiple reflection effects in rugged terrain, International Journal of Remote Sensing, 30, 1075–1081, 2009.

Sirguey, P., Mathieu, R., and Arnaud, Y.: Subpixel monitoring of the seasonal snow cover with MODIS at 250 m spatial resolution in the Southern Alps of New Zealand: Methodology and accuracy assessment, Remote Sensing of Environment, 113, 160–181, 2009.

Tesfa, T. K. and Leung, L.-Y. R.: Exploring new topography-based subgrid spatial structures for improving land surface modeling, Geoscientific Model Development, 10, 873–888, 2017.

520 Zhao, M., Golaz, J.-C., Held, I., Guo, H., Balaji, V., Benson, R., Chen, J.-H., Chen, X., Donner, L., Dunne, J., et al.: The GFDL global atmosphere and land model AM4. 0/LM4. 0: 2. Model description, sensitivity studies, and tuning strategies, Journal of Advances in Modeling Earth Systems, 10, 735–769, 2018.