AGU ADVANCING EARTH AND SPACE SCIENCE

JAMES | Journal of Advances in Modeling Earth Systems®

RESEARCH ARTICLE
10.1029/2021MS002862

Key Points:

- The Energy Exascale Earth System Model (E3SM) land model (ELM) component was used to evaluate the impacts of sub-grid topographic representations
- Sub-grid topographic representation has large impacts on surface energy balance and surface boundary conditions
- Topounit provides better performance than the default sub-grid structure in the E3SM land model

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

D. Hao,
dalei.hao@pnnl.gov

Citation:

Hao, D., Bisht, G., Huang, M., Ma, P.-L., Tesfa, T., Lee, W.-L., et al. (2022). Impacts of sub-grid topographic representations on surface energy balance and boundary conditions in the E3SM land model: A case study in Sierra Nevada. *Journal of Advances in Modeling Earth Systems*, *14*, e2021MS002862. https://doi.org/10.1029/2021MS002862

Received 4 OCT 2021
Accepted 22 MAR 2022

Author Contributions:

Conceptualization: Dalei Hao, Gautam Bisht
Data curation: Dalei Hao
Formal analysis: Dalei Hao
Funding acquisition: Gautam Bisht, L. Ruby Leung
Investigation: Dalei Hao, Gautam Bisht, L. Ruby Leung

Impacts of Sub-Grid Topographic Representations on Surface Energy Balance and Boundary Conditions in the E3SM Land Model: A Case Study in Sierra Nevada

Dalei Hao[1], Gautam Bisht[1], Meng Huang[1], Po-Lun Ma[1], Teklu Tesfa[2], Wei-Liang Lee[3], Yu Gu[4,5], and L. Ruby Leung[1]

[1]Atmospheric Sciences and Global Change Division, Pacific Northwest National Laboratory, Richland, WA, USA, [2]Earth Systems Science Division, Pacific Northwest National Laboratory, Richland, WA, USA, [3]Research Center for Environmental Changes, Academia Sinica, Taipei City, Taiwan, [4]Joint Institute for Regional Earth System Science and Engineering, University of California Los Angeles, Los Angeles, CA, USA, [5]Department of Atmospheric and Oceanic Sciences, University of California Los Angeles, Los Angeles, CA, USA

Abstract Sub-grid topographic heterogeneity has large impacts on surface energy balance and land-atmosphere interactions. However, the impacts of representing sub-grid topographic effects in land surface models (LSMs) on surface energy balance and boundary conditions remain unclear. This study analyzed and evaluated the impacts of sub-grid topographic representations on surface energy balance, turbulent heat flux, and scalar (co-)variances in the Energy Exascale Earth System Model (E3SM) land model (ELM). Three sub-grid topographic representations in ELM were compared: (a) the default sub-grid structure (D), (b) the recently developed sub-grid topographic structure (T), and (c) high spatial resolution (1KM). Additionally, two different solar radiation schemes in ELM were compared: (a) the default plane-parallel radiative transfer scheme (PP) and (b) the parameterization scheme (TOP) that accounts for sub-grid topographic effects on solar radiation. A series of offline simulations with the three grid discretization structures (D, T, and 1KM) and two schemes of solar radiation (TOP and PP) were carried out using the Sierra Nevada, California. 1KM simulations with TOP well capture the spatial heterogeneity of surface fluxes compared to Moderate Resolution Imaging Spectroradiometer remote sensing data. There are significant differences between TOP and PP in the 1-km simulated surface energy balance, but the differences in mean values and standard deviations become small when aggregated to the grid scale (i.e., 0.5°). The T configuration better mimics the 1KM simulations with TOP than the D configuration and better captures the sub-grid topographic effects on surface energy balance and boundary conditions. These results underline the importance of representing sub-grid topographic heterogeneities in LSMs and motivate future research to understand the sub-grid topographic effects on land-atmosphere interactions over mountainous areas.

Plain Language Summary Topography has significant impacts on surface energy balance and land-atmosphere interactions, which has not been well represented in land surface models (LSMs). Our study quantified the impacts of representing sub-grid topography on surface energy balance and surface boundary conditions in the Energy Exascale Earth System Model (E3SM) land model (ELM). The topounit-based sub-grid structure that parsimoniously captures sub-grid topographic heterogeneity provides a better performance than the default sub-grid structure in ELM and is computationally efficient. Our study underlines the importance of representing sub-grid topographic heterogeneities in LSMs and Earth System Models and is promising to promote future research to understand the sub-grid topographic effects on terrestrial processes and land-atmosphere interactions over complex terrain.

1. Introduction

Ubiquitous heterogeneity of the land surface plays a pivotal role in surface energy balance and land-atmosphere exchanges of momentum, heat, and water (Bou-Zeid et al., 2020; de Vrese et al., 2016; Levy et al., 2020). Sub-grid heterogeneity comprises spatial variabilities in land use/land cover (LULC) types, soil characteristics, and topography (Giorgi & Avissar, 1997). Sub-grid heterogeneity affects the surface energy partitioning, modifies the vertical structure of the planetary boundary layer (PBL), creates mesoscale atmospheric circulations, and affects cloud formation and regional climate (Brunsell et al., 2011; Chen et al., 2020; Lee, Zhang, & Klein, 2019;

Methodology: Dalei Hao, Gautam Bisht, Meng Huang, Teklu Tesfa, Wei-Liang Lee, Yu Gu, L. Ruby Leung
Project Administration: Gautam Bisht
Resources: Dalei Hao, Gautam Bisht
Software: Dalei Hao
Supervision: Gautam Bisht, L. Ruby Leung
Validation: Dalei Hao
Visualization: Dalei Hao
Writing – original draft: Dalei Hao
Writing – review & editing: Dalei Hao, Gautam Bisht, Meng Huang, Po-Lun Ma, Teklu Tesfa, Wei-Liang Lee, Yu Gu, L. Ruby Leung

Maronga & Raasch, 2013; Wu et al., 2009; Zhang et al., 2010; Zheng et al., 2021). Accounting for sub-grid heterogeneities of land surface plays a vital role in land surface modeling and land-atmosphere coupling (de Vrese et al., 2016; Fisher & Koven, 2020).

Topographic variability, an important heterogeneous feature of the land surface, has large impacts on many land surface processes. Topography fundamentally organizes water, energy, and biogeochemical processes at hillslope scales by modifying the downwelling solar radiation at the land surface and laterally transporting water from ridges to valleys (Fan et al., 2019). By geometric shadowing, topography causes a strong contrast of incoming solar radiation between sunny and shady slopes (Hao et al., 2019; Proy et al., 1989). While valleys generally receive less sky diffuse radiation than hilltops due to the obstructions from adjacent slopes, they receive more multi-scattering radiation from adjacent slopes (Dubayah, 1992). The variability in the incoming solar radiation leads to spatial variability in climate, soil, and vegetation characteristics that follow the topographic distributions (Ropars & Boudreau, 2012). For instance, north-facing and south-facing slopes support different vegetation types, densities, and species diversities (Dearborn & Danby, 2017). Representing these topographic effects on incoming solar radiation as well as the topography-dependent soil and vegetation distributions at sub-grid levels in coarse resolution global LSMs is necessary for accurately resolving terrestrial processes in regions with heterogeneous terrain.

While sub-grid heterogeneity of LULC is routinely accounted for in LSMs, the sub-grid variability due to topography has often been neglected. LSMs account for LULC heterogeneity using a computationally efficient tiling approach in which a coarse-scale grid cell is statistically or geographically divided into sub-grid cells, each of which is a single LULC type. For example, the Community Land Model Version 5.0 (CLM5.0; Lawrence et al., 2019) and the Energy Exascale Earth System Model (E3SM) land model Version 1 (ELMv1; Golaz et al., 2019) use a nested sub-grid hierarchical structure where each grid is statistically composed of multiple land units, each land unit can have multiple soil columns, and each soil column can have multiple plant functional types (PFTs). These existing sub-grid schemes neglect the joint distributions of sub-grid topography, soil, and vegetation. Sub-grid heterogeneities can also be represented by running LSMs at high spatial resolutions (∼1 km), but this method is computationally prohibitive for global Earth System Models (ESMs). Recently, high-resolution Digital Elevation Model (DEM) data have been used to add topography-based sub-grid structures in LSMs, including CLM and ELM (Chaney et al., 2016; Fiddes & Gruber, 2012; Ke et al., 2013; Tesfa & Leung, 2017). Tesfa and Leung (2017) extended the default sub-grid scheme of ELM by introducing topographic units (topounits), each associated with different topographic features. In the new sub-grid scheme, an ELM coarse-scale grid (0.5° or coarser) can have multiple topounits and each topounit can have multiple land units, soil columns, and PFTs. In addition, most existing LSMs adopt a two-stream solar radiative transfer scheme with the plane-parallel (PP) assumption that neglects the sub-grid topographic effects on solar radiation (Sellers, 1985). Recently, a scheme (TOP) to account for the sub-grid topographic effects on solar radiation has been implemented in ELM (Hao et al., 2021). The aforementioned improvements in ELM offer a great opportunity to study the sub-grid topographic effects on land surface processes.

How sub-grid topographic representations in LSMs may affect the simulated surface energy balance and surface boundary conditions is unclear. Sub-grid topography can affect the radiative transfer processes and hence influence the surface energy balance (Hao et al., 2021). The mean surface turbulent heat flux (e.g., latent heat and sensible heat fluxes) (Lyons & Halldin, 2004) and high-order scalar (co-)variances (i.e., the variances and covariances of scalar quantities, such as near-surface temperature and humidity) are used in atmospheric turbulence parameterization of ESMs for land-atmosphere coupling (Guo et al., 2015). These surface boundary conditions are crucial for simulating atmospheric flows and circulations in ESMs (Machulskaya & Mironov, 2018). Different representations of sub-grid heterogeneity in LSMs can lead to different surface turbulent heat flux and scalar (co-) variances that can in turn affect the associated PBL processes (Román-Cascón et al., 2021). ELM can be configured in three different ways to capture the sub-grid heterogeneity: (a) a coarse-resolution simulation (denoted as D) with the default sub-grid structure that excludes sub-grid topographic heterogeneity, (b) a coarse-resolution simulation (denoted as T) with topounit-based sub-grid structure that parsimoniously captures sub-grid topographic heterogeneity, and (c) a high-resolution (i.e., 1 km) simulation (denoted as 1KM) that explicitly accounts for sub-grid heterogeneity. Generally, the 1KM simulations with spatially explicit sub-grid heterogeneity can be regarded as a reference simulation (Schneider et al., 2017). The accuracy of the D- and T-based simulations to simulate the surface energy budget and surface boundary conditions over heterogeneous terrain needs to be

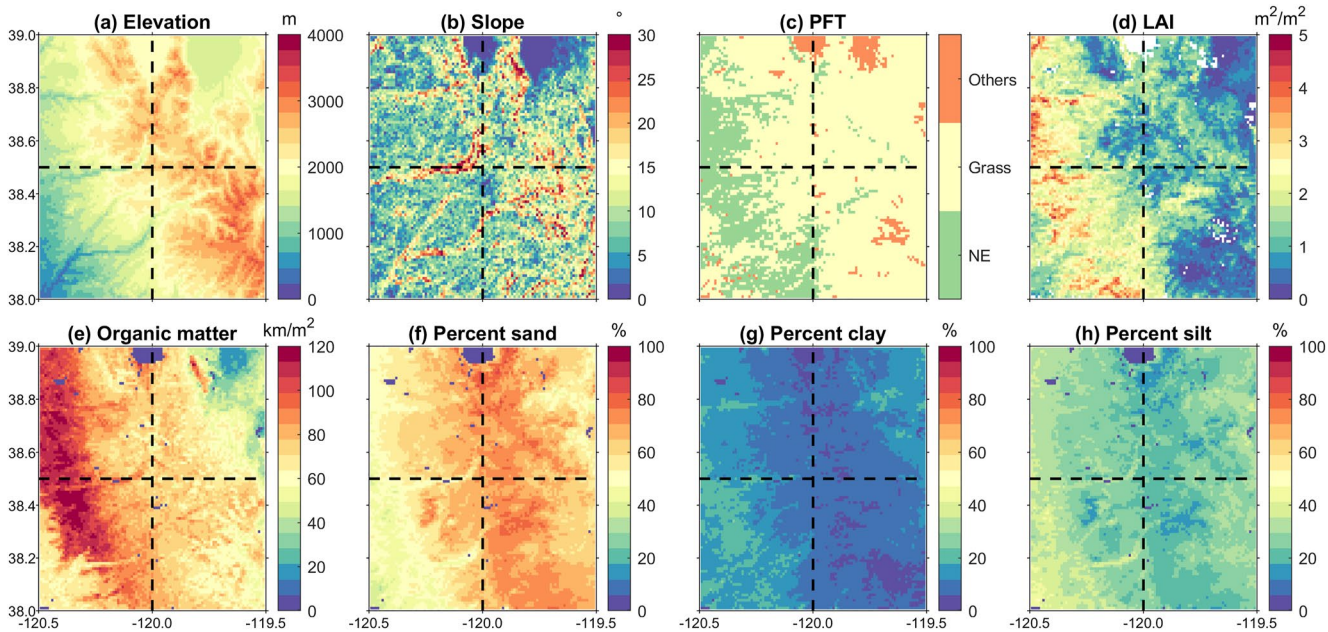

Figure 1. Spatial distributions of (a) elevation, (b) slope, (c) PFT, (d) leaf area index in August, (e) organic matter density, and (f–h) percentages of sand, clay, and silt (d and e), respectively, over the study area at a resolution of 1 km. In panel (c), NE, grass, and others represent the Needleleaf Evergreen Trees-Temperate, C3 grass, and other PFTs.

evaluated. In addition, the role of sub-grid topography within solar radiative transfer schemes (i.e., TOP and PP) on surface energy balance terms and surface boundary conditions deserves further investigations.

The objective of this study is to analyze the effects of sub-grid topographic heterogeneity on surface energy balance and surface boundary conditions for turbulent heat flux and scalar (co-)variances. Land surface parameters at a spatial resolution of 1 km were first developed based on existing high-resolution data sets for vegetation, soil, and DEM. Then, a series of offline ELM simulations with three different sub-grid topographic representations (D, T, and 1KM) and two solar radiation radiative transfer schemes (TOP and PP) were carried out over a region of heterogeneous terrain. The surface boundary conditions for the variance and co-variance quantities under different cases were next derived using the method accounting for the surface heterogeneity in Machulskaya and Mironov (2018), which is suitable for analyzing results over heterogeneous surfaces. 1KM simulations with TOP were compared to the Moderate Resolution Imaging Spectroradiometer (MODIS) remote sensing data. Finally, the effects of sub-grid topographic representations in ELM on surface energy balance and boundary conditions were analyzed and discussed.

2. Materials and Methods

2.1. Study Area

A representative 1° grid (38°–39°N and 120.5°–119.5°W) in the Sierra Nevada, California, as an important source of California's water supply with heterogeneous distribution of topography was selected as a testbed in the study. Accurate modeling of surface energy balance and surface boundary conditions over the Sierra Nevada has important implications for Earth System Modeling. The study area covers various topographic features, such as hill, ridge, valley, and saddle with an elevation ranging from 330 to 3,418 m (Figure 1a) and a slope (Figure 1b) varying from 0° (i.e., flat surface) to 41° (i.e., steep slope). The study region is mainly dominated by forests (southwest) and shrubland/grasslands (northeast) with a few small urban areas, croplands, and lakes (Figure 1c). The southwestern regions have higher leaf area index (LAI), organic matter density, clay and silt contents and lower sand contents than the northeastern regions (Figures 1d–1h).

Table 1
Specifications of High-Resolution Data Sets Used in This Study

Group	Parameter	Data source	Period	Spatial resolution (m)	Temporal resolution	References
Climate	Temperature	WorldClim V1	Climatological	1,000	Monthly	Hijmans et al. (2005)
	Precipitation	WorldClim V1	Climatological	1,000	Monthly	Hijmans et al. (2005)
Vegetation	Land cover	MODIS MCD12Q1 V6	2010	500	Yearly	Friedl et al. (2002)
	Leaf area index	MODIS MCD15A3H V6	2003–2010	500	4-day	Myneni et al. (2002)
	Stem area index	Derived from LAI	2003–2010	500	4-day	Myneni et al. (2002)
Soil	Organic matter density	Soilgrid v2	–	250	–	Poggio et al. (2021)
	Percent sand	Soilgrid v2	–	250	–	Poggio et al. (2021)
	Percent clay	Soilgrid v2	–	250	–	Poggio et al. (2021)
Topography	Elevation	SRTM DEM	–	90	–	Rabus et al. (2003)
Evaluation	Direct and diffuse albedo	MODIS MCD43A3 V6	2001–2010	500	Daily	Schaaf et al. (2002)
	Snow cover fraction	MODIS MOD10A1 V6	2001–2010	500	Daily	Hall et al. (2002)
	Latent heat flux	MODIS MOD16A2 V6	2001–2010	500	8-day	Mu et al. (2007)
	GPP	MODIS MOD17A2H V6	2001–2010	500	8-day	Running et al. (2004)

2.2. High-Resolution Land Surface Data Sets for ELM

In this study, a series of high-resolution, sub-kilometer data sets for vegetation, soil, and topography were collected to develop spatially continuous land surface parameters at 1 km for ELM (Table 1), following the methods in Ke et al. (2012). To unify the spatial resolution, the soil, topography, and LAI/SAI data sets were first aggregated to a spatial resolution of 1 km using an area-weighted average method, and the land cover data set was aggregated to 1 km using a majority resampling method. The default values were used for all other land surface parameters in ELM.

MODIS 500-m land cover and LAI data were acquired from the Google Earth Engine (Gorelick et al., 2017). Specifically, the IGBP land cover classification data in the yearly 500 m MCD12Q1 V6 land cover-type product (Friedl et al., 2002) for the year 2010 were used to determine lake, wetland, urban, and PFTs based on the methods in (Bonan et al., 2002; Ke et al., 2012). The IGBP classification scheme was first converted to ELM's PFTs that are composed of needleleaf evergreen trees, needleleaf deciduous trees, broadleaf evergreen trees, broadleaf deciduous trees, shrub, grass, and crop. The WorldClim V1 1 km monthly climatological temperature and precipitation data (Hijmans et al., 2005) were further used to classify the PFTs into tropical, temperate, and boreal climate groups based on the rules presented by Bonan et al. (2002). The fractions of C3 and C4 grasses were derived using the method in Still et al. (2003). The monthly LAI time series were derived based on the 4-day 500 m MCD15A3H V6 LAI product (Myneni et al., 2002) from 2003 to 2010. The methodology in Zeng et al. (2002) was used to calculate monthly stem area index (SAI) from the monthly LAI data.

The Soilgrid v2 data with a resolution of 250 m (Poggio et al., 2021) were used to represent high-resolution soil characteristics. This data was generated using site-level soil profile data and remote sensing-based soil covariates and machine learning methods trained based on site-level soil profile data (Hengl et al., 2017). Specifically, organic matter density, percent clay, and percent sand at seven standard depths (i.e., 0, 5, 15, 30, 60, 100, and 200 cm) were used to derive corresponding soil organic matter and structure in ELM.

The Shuttle Radar Topography Mission (SRTM) DEM data (Rabus et al., 2003) with a spatial resolution of 90 m were used to derive topography-related parameters, including mean elevation, mean slope, and standard deviation of elevation. These data were also used to derive 1 km topographic factors used in the parameterization of sub-grid topographic effects on solar radiation in ELM (see Section 2.3).

MODIS data sets were used in the evaluation of 1 km simulations (Table 1). Specifically, direct albedo (α_{dir}), diffuse albedo (α_{dif}), snow cover fraction (f_{snow}), latent heat flux (F_{lh}), and gross primary productivity (GPP) were collected. All these MODIS data sets were upscaled to 1 km using the area-weighted average methods.

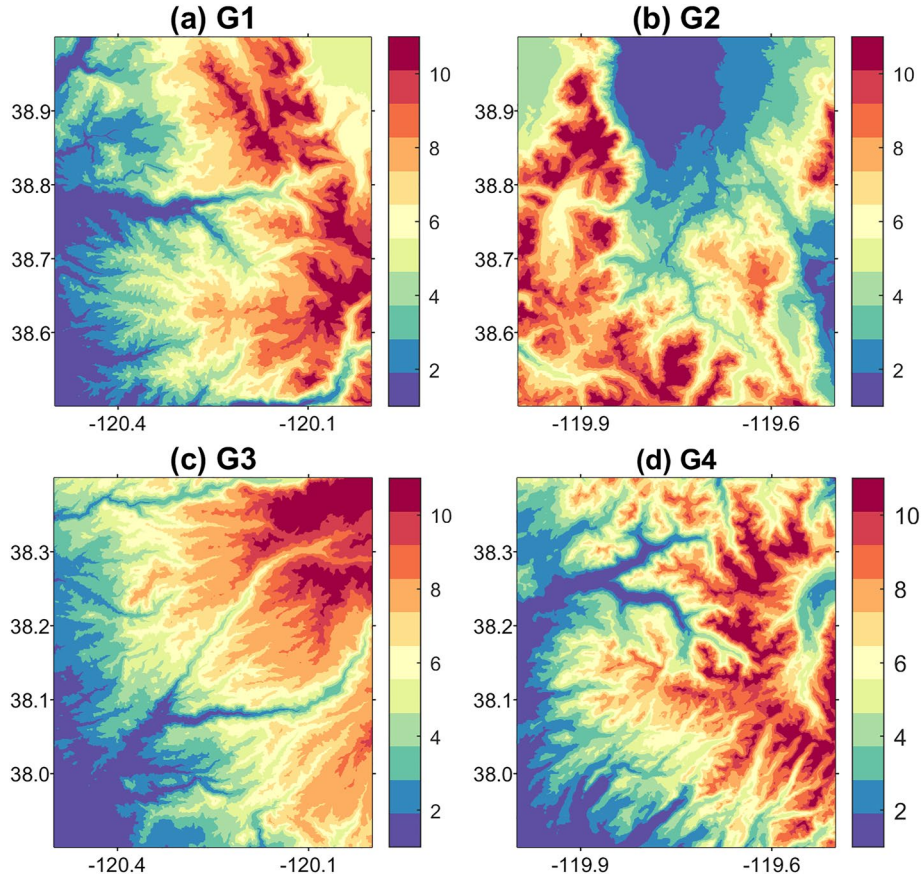

Figure 2. Spatial patterns of topounits within four different 0.5° × 0.5° grids. For each grid, the topounit code varies from 1 to 11.

2.3. Sub-Grid Topographic Improvements in ELM

E3SM, supported by the U.S. Department of Energy is an advanced ESM developed to address the grand challenge of actionable predictions of Earth system variability and change (Leung et al., 2020). ELMv1 originated from CLM4.5 and has incorporated a few new features, such as a new variably saturated flow model (Bisht et al., 2018) and an updated module for phosphorus cycle dynamics (Yang et al., 2019).

ELMv1 adopts a nested hierarchical sub-grid structure to represent surface heterogeneity. Each ELM grid cell is first divided into different landunits representing glacier, lake, wetland, urban, natural vegetation, and crops. A naturally vegetated landunit has a single soil column that is shared by multiple PFTs. Each PFT has its own vegetation characteristics (e.g., LAI/SAI). However, surface topography is not considered in this sub-grid structure.

Recently, a new sub-grid topographic structure (topounit) has been included in ELM to improve the representation of sub-grid topographic effects on land surface processes (Tesfa & Leung, 2017). Specifically, each 0.5° grid is divided into different topounits based on the topographic elevation. The study area includes four different 0.5° grids (i.e., G1, G2, G3, and G4) and each grid with distinct topographic features (i.e., elevation, slope, and aspect) includes 11 topounits and shows different topounit distributions (Figure 2). The spatial patterns of topounits in Figure 2 generally follow the spatial distribution of elevation in Figure 1a. Different topounits have different elevation ranges. For instance, the elevation of topounit-1 for G1 ranges from 1,040 to 1,713 m, and that of topounit-11 ranges from 2,303 to 2,952 m. Besides, each topounit can have its own PFT and soil characteristics.

In addition, a parameterization of sub-grid topographic effects on solar radiation has been implemented in ELM (Hao et al., 2021). This cost-effective parameterization uses multiple linear regression methods to build the relationship between several topographic factors and sub-grid topographic effects (Lee et al., 2011). The topographic

factors include the standard deviation of elevation, grid-average cosine of the local solar incident angle, sky view factor, and terrain configuration factor, which are precomputed based on the 90 m SRTM data (Table 1). This parameterization has been successfully applied in the Weather Research and Forecasting (WRF), CLM4, and ELM at different spatial scales ranging from 800 m to 200 km (Hao et al., 2021; Lee, Liou, et al., 2019; Zhao et al., 2016).

In this study, we used ELMv1 with the two abovementioned sub-grid topographic improvements to analyze the effects of sub-grid topographic representations on surface energy balance and boundary conditions.

2.4. Computation of Surface Boundary Conditions of Scalar (Co-)variances

Most of the existing ESMs couple the land and atmosphere using grid-scale mean fluxes and state variables, thus ignoring the effects of sub-grid heterogeneity. For instance, the current version of E3SM couples the land and atmosphere at the grid level (André et al., 1978), which assumes that land-atmosphere interaction is dominated by spatially homogeneous processes, so this method is hereafter referred to as the HOM method. In contrast, Machulskaya and Mironov (2018) developed a patch(tile)-based method that accounts for the effects of sub-grid heterogeneity. Hereafter, this method is referred to the HET method. Specifically, when considering the heterogeneity of the underlying surface as composed of different patches, a generic variable x can be decomposed into three components (Avissar & Chen, 1993):

$$x = \langle \bar{x}_p \rangle + \bar{x}_p'' + x_{sp} \tag{1}$$

where $\langle \bar{x}_p \rangle$ represents the average value over a grid, $\bar{x}_p$ represents the average value over the pth patch, $\bar{x}_p''$ $(= \bar{x}_p - \langle \bar{x}_p \rangle)$ represents an average fluctuation of a patch-level value away from the grid-level average value, and x_{sp} represents a sub-patch-level fluctuation, which can be estimated by the HOM method. Here, we define the sum of the patch- and grid-level fluctuations as

$$x' = \bar{x}_p'' + x_{sp} \tag{2}$$

By construction, $\langle \bar{x}_p'' \rangle = 0$ and $\bar{x_{sp}} = 0$, that is, the grid-level average value of the patch-level fluctuations and the patch-level average value of the sub-patch level fluctuations are zero (Machulskaya & Mironov, 2018). Then, the scalar variance for any quantity x (e.g., temperature and humidity) denoted by $\langle x'^2 \rangle$ can be derived as the sum of sub-patch- and patch-level variances:

$$\langle x'^2 \rangle = \left\langle \overline{\left(\bar{x}_p'' + x_{sp}\right) \cdot \left(\bar{x}_p'' + x_{sp}\right)} \right\rangle = \left\langle \bar{x}_p''^2 \right\rangle + \left\langle \overline{x_{sp}^2} \right\rangle \tag{3}$$

Similarly, the scalar covariance between two quantities, x and y (e.g., temperature and humidity), can be calculated as

$$\langle x'y' \rangle = \left\langle \overline{\left(\bar{x}_p'' + x_{sp}\right) \cdot \left(\bar{y}_p'' + y_{sp}\right)} \right\rangle = \langle \bar{x}_p'' \bar{y}_p'' \rangle + \langle \overline{x_{sp} y_{sp}} \rangle \tag{4}$$

The discrete forms of Equations 3 and 4 can be expressed as

$$\langle x'^2 \rangle = \sum_1^N f_p \cdot (\bar{x}_p - \langle \bar{x}_p \rangle)^2 + \sum_1^N f_p \cdot \overline{x_{sp}^2} \tag{5}$$

$$\langle x'y' \rangle = \sum_1^N f_p \cdot (\bar{x}_p - \langle \bar{x}_p \rangle) \cdot (\bar{y}_p - \langle \bar{y}_p \rangle) + \sum_1^N f_p \cdot \overline{x_{sp} y_{sp}} \tag{6}$$

where f_p is the area fraction of the pth patch, N is the number of all patches, $\bar{x}_p$ and $\bar{y}_p$ represent the average value for the pth patch, x_{sp} and y_{sp} represent the sub-patch fluctuations for the pth patch, and $\langle \bar{x}_p \rangle$ and $\langle \bar{y}_p \rangle$ represent the grid-level average value.

In this study, we focus on the surface boundary conditions for the mean turbulent heat flux (i.e., latent heat and sensible heat flux) and three (co-)variances that included temperature variance ($\langle \theta'^2 \rangle$), humidity variance ($\langle q'^2 \rangle$),

Table 2
Model Configurations With Different Sub-Grid Topographic Representations for Each 0.5° × 0.5° Grid

Case ID	Representations of sub-grid topography	Solar radiation parameterization	No. of grid cells	No. of elevation bands	Spatial resolution		
					Vegetation	Soil	Atmospheric forcing
D_PP	Default	PP	1	1	Sub-grid	0.5°	0.5°
D_TOP	Default	TOP	1	1	Sub-grid	0.5°	0.5°
T_PP	Topounit	PP	1	11	Sub-topounit	Topounit	0.5°
T_TOP	Topounit	TOP	1	11	Sub-topounit	Topounit	0.5°
1KM_PP	1 km	PP	50 × 50	1	1 km	1 km	0.5°
1KM_TOP	1 km	TOP	50 × 50	1	1 km	1 km	0.5°

and temperature-humidity co-variance ($\langle \theta' q' \rangle$), which are used as the lower boundary conditions for atmospheric turbulence parameterization of E3SM (Xie et al., 2018).

2.5. Experimental Design and Analysis

A series of offline ELM simulations with different configurations, as summarized in Table 2, were designed and conducted for the four 0.5° grids of the study area to investigate the effects of sub-grid topography on surface energy balance and boundary conditions. Three different sub-grid topographic representations (D, T, and 1KM) and two different schemes of sub-grid topographic effects on solar radiation (TOP and PP) were considered in the simulations. The 1 km land surface parameters generated in Section 2.2 were used to derive the 1 km ELM surface data set. Each ELM grid cell in the 1KM configuration has a single PFT and its own unique soil characteristics. The surface data set for the D and T configurations was generated using column fraction- and PFT fraction-weighted average methods, respectively, for soil and vegetation variables to upscale the 1 km surface data set. For the T configuration, column and PFT fractions for each topounit were calculated and thus vegetation and soil characteristics of each topounit are different. For the 1KM configuration, the topographic factors needed for the TOP scheme were derived from the 1 km topographic factor data set introduced in Sections 2.2 and 2.3. For the D configuration, 1 km topographic factors were averaged to grid (0.5°) level, while for the T configuration, 1 km topographic factors were averaged to topounit level.

All simulations used the prescribed satellite phenology (SP) mode, and the 0.5° × 0.5° Global Soil Wetness Project data set (GSWP3) (Dirmeyer et al., 2006) was used as the meteorological forcing data. Specifically, 3-hourly total incident solar radiation, incident longwave radiation, total precipitation, surface air pressure, air temperature, wind speed, and specific humidity in GSWP3 were used. The spatial heterogeneity of the atmospheric forcing data was not considered in the study by using the same atmospheric forcing data sets for all 1 km grids (in the 1KM configuration) and different topounits (in the T configuration). Model outputs were saved at half-hourly time step. Simulations for all model configurations were performed for 31-years from 1980 to 2010. The first 20 years were considered as model spin up and the simulation outputs from 2000 to 2010 were used in the subsequent analysis.

To further disentangle the contributions of vegetation and soil, additional topounit-based simulations with/without topounit-dependent vegetation and with/without soil variations were conducted (Table 3). Considering that the differences between TOP and PP in the simulated surface energy balance terms are small for the T configurations (see Section 3.3), only the TOP solar radiation parameterization was used in the sensitivity experiments. Specifically, these topounit-based experiments consisted of heterogeneous vegetation and soil (V + S), only heterogeneous soil (S), only heterogeneous vegetation (V), and no heterogeneity in vegetation and soil (N). Other settings were the same as the simulations in Table 2.

The surface energy balance terms were directly outputted from the aforementioned simulations and the surface boundary conditions for scalar (co-)variances were calculated at a half-hourly scale using the methods described in Section 2.4. Specifically, net solar radiation (R^s_{net}), f_{snow}, surface radiative temperature (T_s), F_{lh}, sensible heat

Table 3
Sensitivity Experiments Under the T Configurations and TOP Solar Radiation Scheme With/Without Considerations of Vegetation or Soil Heterogeneity for Each 0.5° × 0.5° Grid

Case ID	Representations of sub-grid topography	Solar radiation parameterization	No. of grid cells	No. of elevation bands	Spatial resolution		
					Vegetation	Soil	Atmospheric forcing
V + S	Topounit	TOP	1	11	Sub-topounit	Topounit	0.5°
S	Topounit	TOP	1	11	Sub-grid	Topounit	0.5°
T	Topounit	TOP	1	11	Sub-topounit	0.5°	0.5°
N	Topounit	TOP	1	11	Sub-grid	0.5°	0.5°

flux (F_{sh}), and GPP were used in the analysis. Based on the Stefan-Boltzmann law and the assumed surface emissivity of 1.0, the emitted longwave radiation was used to derive T_s. The 11-year averaged daily values from 2000 to 2010 were calculated. Then, their mean value (mean), standard deviation (std), and coefficient of variation *(cv)* were used to describe the effects of sub-grid heterogeneity on the surface energy balance terms and turbulent heat flux. For the simulations with different configurations, mean and std were derived by

$$\text{mean} = \frac{\sum_1^N w_i \cdot x_i}{\sum_1^N w_i} \tag{7}$$

$$\text{std} = \sqrt{\frac{\sum_1^N w_i (x_i - \text{mean})^2}{\sum_1^N w_i}} \tag{8}$$

where N is the number of sub-grid cells, x represents the variables of interest (e.g., latent heat and sensible heat flux), x_i is the value of the ith sub-grid cell, and w_i is the area fraction of type i. For the 1KM configuration, $w_i = 1.0$. The cv is calculated as

$$cv = \frac{\text{std}}{\text{mean}} \tag{9}$$

1KM_TOP was compared to MODIS data to evaluate the performance of 1KM_TOP in capturing the spatially heterogeneous pattern of land surface fluxes. α_{dir}, α_{dif}, f_{snow}, F_{lh}, and GPP (listed in Table 1) were used in the comparison. All the ELM simulated and MODIS data from 2001 to 2010 were averaged to the annual scale except that f_{snow} was aggregated to winter. Considered that MODIS α_{dir} and α_{dif} represent the surface albedo at local solar noon, the ELM-simulated α_{dir} and α_{dif} at local solar noon were extracted to compute the annual averaged values.

In addition, 1KM_TOP was used as a reference in the analysis and the correlation coefficient (R), normalized bias (nBias), and normalized root mean square deviation (nRMSD) were used to evaluate the agreements between the different configurations with the reference simulation. Specifically, nBias and nRMSD are calculated as

$$\text{nBias} = \frac{\sum_1^N (x_{\text{other}} - x_{\text{1km\_TOP}})}{N \cdot \text{IQR}} \tag{10}$$

$$\text{nRMSD} = \frac{\sqrt{\sum_1^N (x_{\text{other}} - x_{\text{1km\_TOP}})^2 / N}}{\text{IQR}} \tag{11}$$

where N is the available number of target variable x, $x_{\text{1km\_TOP}}$, and x_{other} are the target variable values for 1KM_TOP and other cases, respectively, and IQR is the difference between the 75th and the 25th percentiles of $x_{\text{1km\_TOP}}$.

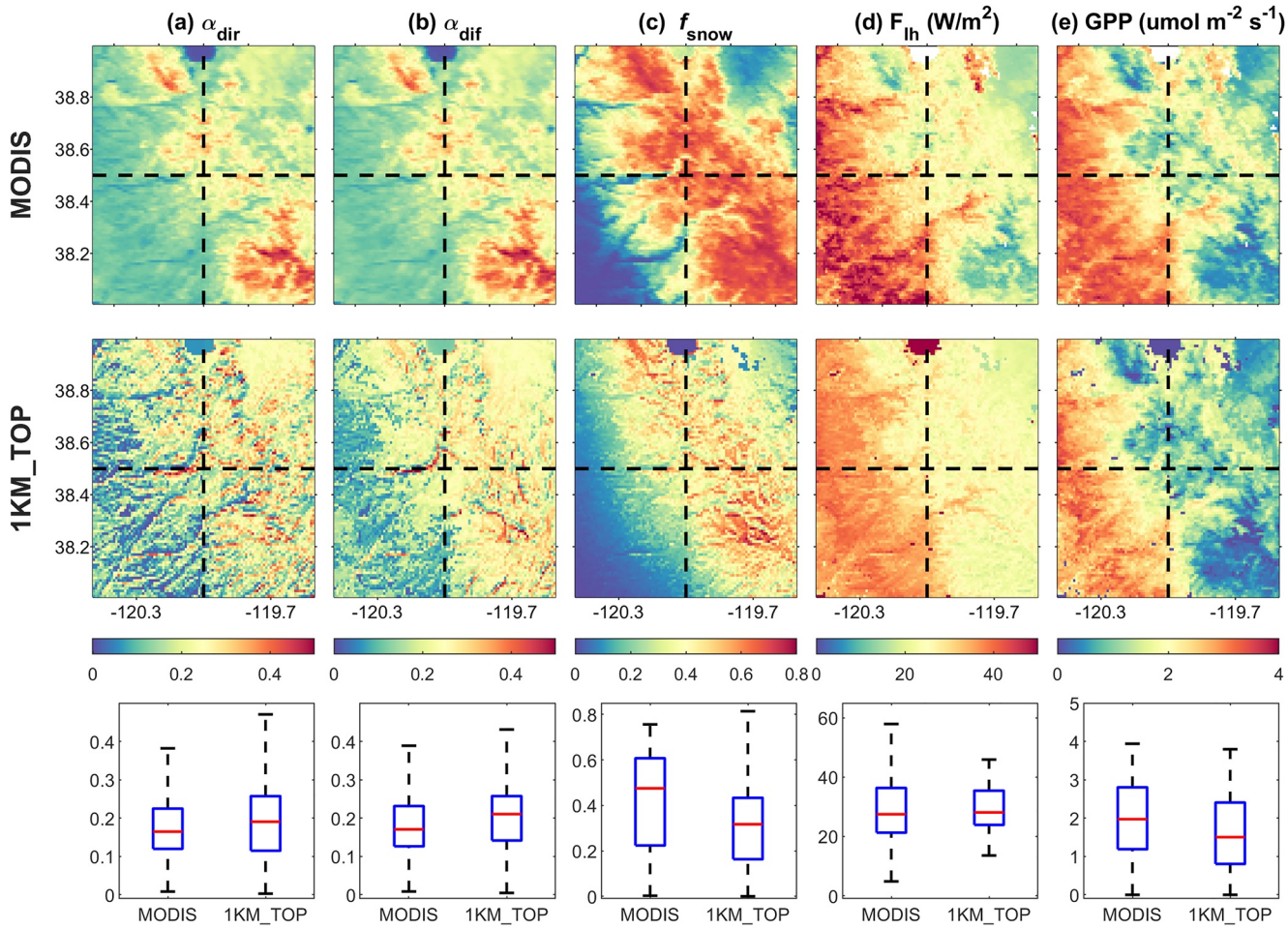

Figure 3. Spatial distributions of (a) direct albedo (α_{dir}), (b) diffuse albedo (α_{dif}), (c) snow cover fraction (f_{snow}), (d) latent heat (F_{lh}), and (e) gross primary productivity (GPP) for Moderate Resolution Imaging Spectroradiometer (MODIS; top row) and Earth System Model land model-simulations (middle row). Their statistical distributions are shown in the bottom row. The annual averaged values are shown except that f_{snow} is for winter.

3. Results

3.1. Comparison With Remote Sensing Data

Overall, 1KM_TOP well captures the spatial distributions of surface energy fluxes compared to MODIS data (Figure 3). For α_{dir} and α_{dif}, 1KM_TOP and MODIS show similar spatial distribution with large values in southeast regions and low values in southwestern and northeastern regions. MODIS has larger magnitudes than 1KM_TOP possibly due to the underestimated f_{snow} in 1KM_TOP. Compared to MODIS, 1KM_TOP is more heterogeneous, possibly because MODIS land surface albedo algorithms do not account for topography explicitly (Hao et al., 2018; Schaaf et al., 2002). For f_{snow}, F_{lh}, and GPP, 1KM_TOP has consistent spatial patterns with MODIS but some underestimations. The statistical boxplots in Figure 3 show that 1KM_TOP and MODIS have similar spread and median values. These results demonstrate that 1KM_TOP can serve as a good reference to evaluate the performance of other sub-grid structures.

3.2. High-Resolution Simulations

TOP and PP have large differences in R^s_{net} for the 1KM configuration. The spatial pattern of R^s_{net} in the winter for the four grids using PP (top row of Figure 4) is related to their PFT distributions (Figure 1c). TOP has more fragmented spatial distributions than PP for all the four grids (middle row of Figure 4). TOP may absorb larger

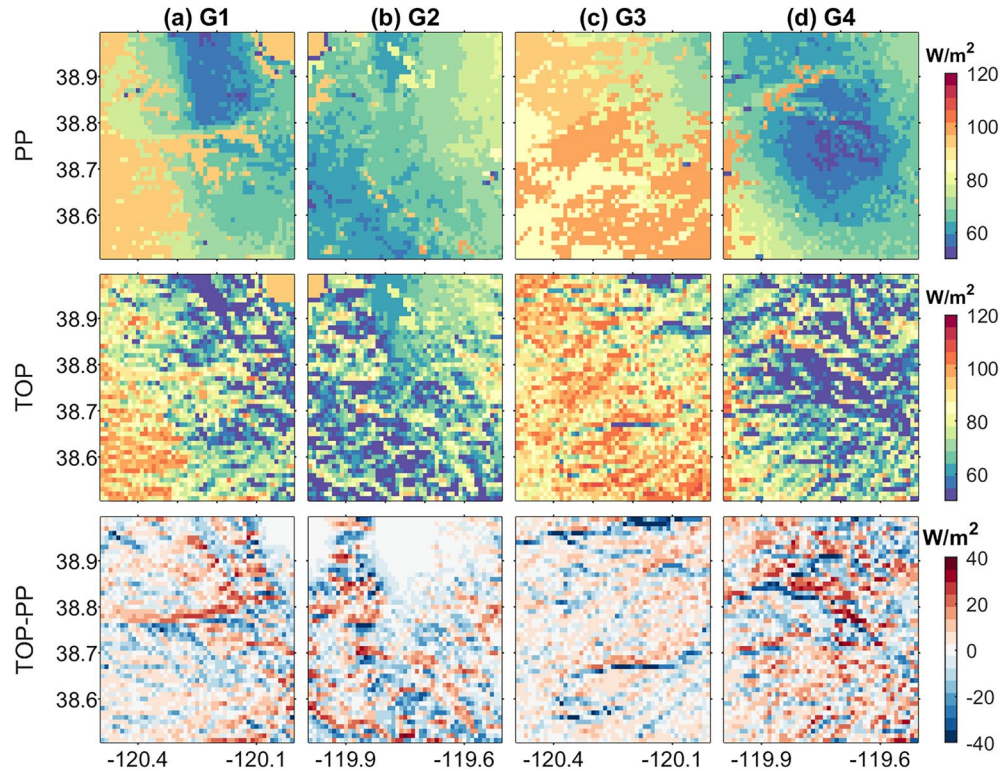

Figure 4. Simulated net solar radiation (R^s_{net}) for plane-parallel (PP; top row), TOP (middle row), and their differences (bottom row) during winter (DJF) for different grids: (a) G1, (b) G1, (c) G3, and (d) G4.

solar radiation due to the change of solar illumination geometry and reflected radiation from the surrounding terrain or may absorb less solar radiation due to the self-shadow or cast-shadow effects than PP depending on the local topography (bottom row of Figure 4). The difference between TOP and PP is less pronounced for G3 as compared to the other grids, which is related to their respective topographic characteristics (Figures 1a and 1b). The difference in R^s_{net} for G1 and G4 can reach up to 40 W/m^2. The seasonal variations and spatial patterns in TOP and PP are similar for G1 (top and middle rows of Figure 5). The southwestern regions have larger R^s_{net}, while the northeastern regions have smaller R^s_{net}, which follow the spatial distribution of the PFTs (Figure 1c). This spatial difference is because grass in the northeastern regions has higher land surface albedo than forest in the southwestern regions. The smaller R^s_{net} of the eastern regions in winter is caused by the larger f_{snow} (Figure 6a). The difference between TOP and PP shown in the bottom row of Figure 5 is larger in winter than summer because of higher f_{snow} and stronger shadowing effects due to large solar zenith angle in winter. The subsequent analysis mainly focuses on G1 for winter.

There are also large differences between TOP and PP in f_{snow}, T_s, turbulent heat flux, and GPP (Figure 6). The spatial patterns of all these variables for both TOP and PP in the top and middle rows of Figure 6 generally follow the spatial variabilities of PFTs (Figure 1c), but TOP is more fragmented than PP, as affected by local topography. The difference between TOP and PP in f_{snow} shows an opposite trend with R^s_{net} (bottom row of Figure 4), which can be larger than 0.2 (bottom row of Figure 6a), because larger net solar radiation can lead to larger snowmelt and thus lower snow f_{snow} and the positive snow albedo feedback can also contribute to it. In contrast, the difference in T_s presents a similar spatial pattern with R^s_{net} and can be as large as 2 K (bottom row of Figure 6b). These further lead to the differences in turbulent heat flux between TOP and PP with larger differences in F_{sh} than F_{lh} (bottom row of Figures 6c–6d). The differences in F_{sh} and F_{lh} can be large as 10 and 20 W/m^2, respectively. GPP is also affected by the sub-grid topographic effects although the difference between TOP and PP is not significant and is within 0.5 μmol m^{-2} s^{-1} due to the small LAI and dormancy in winter (bottom row of Figure 6e).

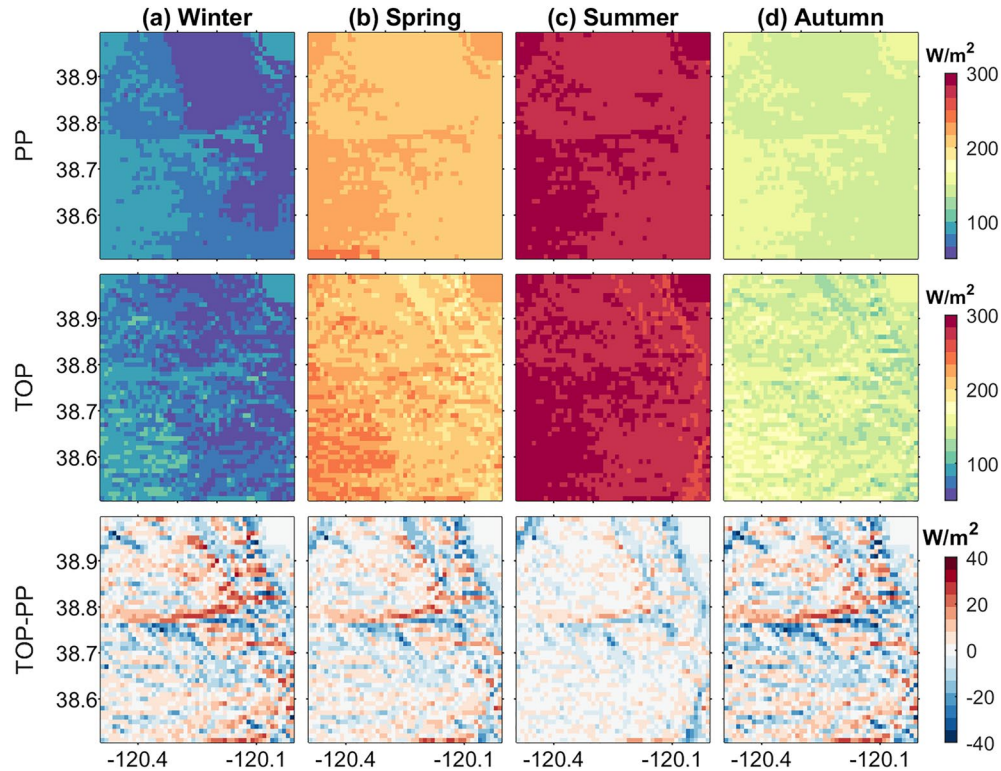

Figure 5. Simulated net solar radiation (R^s_{net}) for plane-parallel (PP; top row), TOP (middle row), and their differences (bottom row) for G1 in different seasons: (a) winter (DJF), (b) spring (MAM), (c) summer (JJA), and (d) autumn (SON).

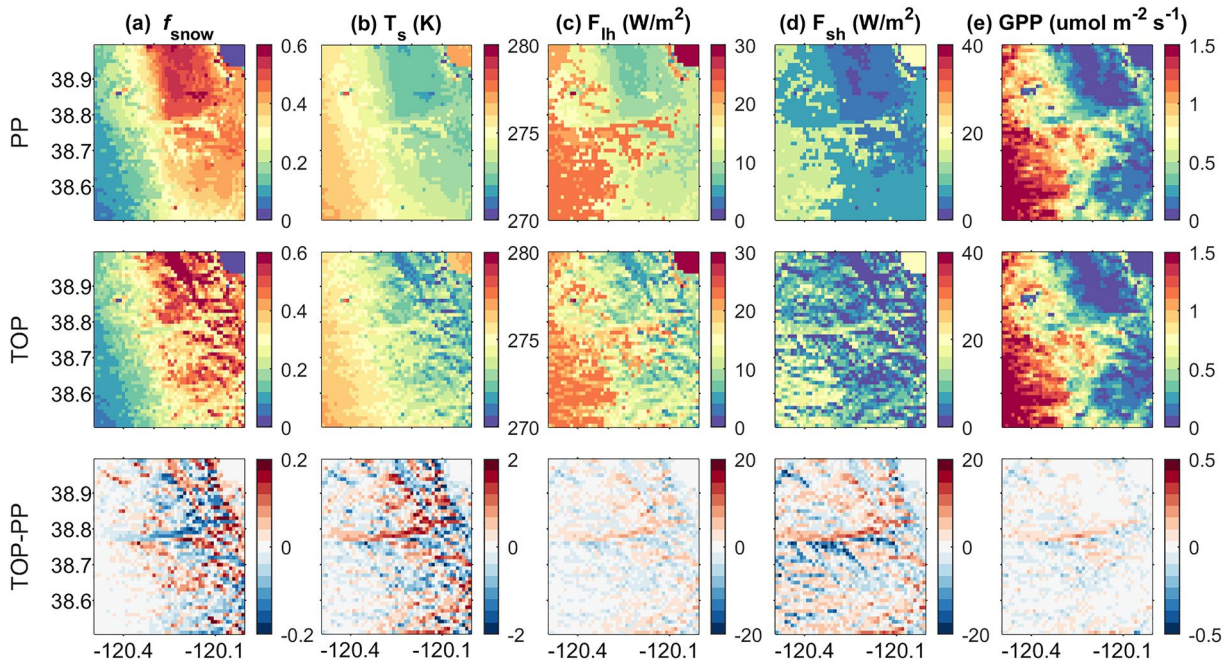

Figure 6. Plane-parallel (PP; top row) and TOP (middle row) simulation and their differences (bottom row) for (a) snow cover fraction (f_{snow}), (b) surface temperature (T_s), (c) latent heat (F_{lh}), (d) sensible heat (F_{sh}) flux, and (e) gross primary productivity (GPP) in winter (DJF) for G1.

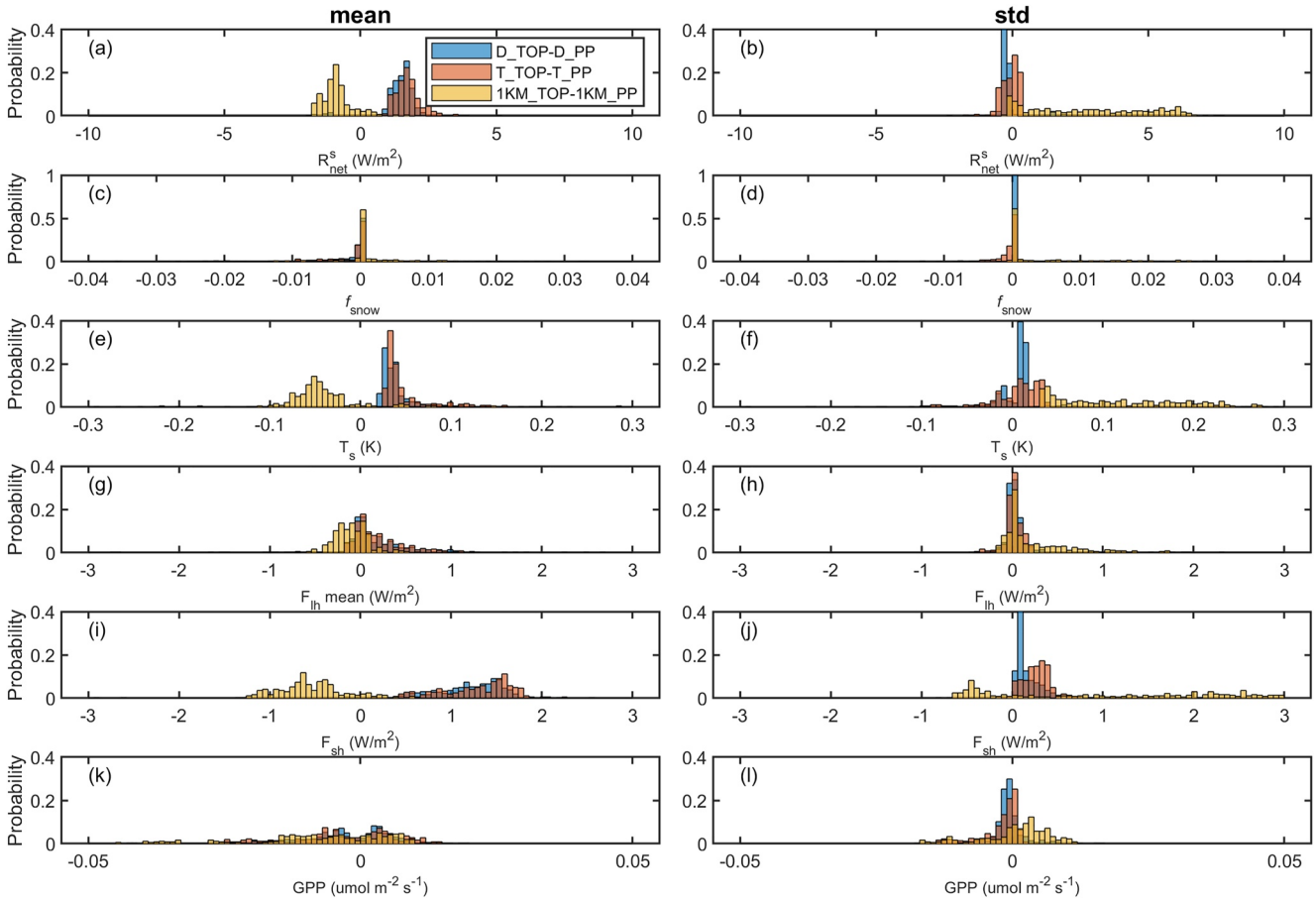

Figure 7. Statistical histograms of the mean and std differences between TOP and plane-parallel (PP) in R^s_{net}, f_{snow}, T_s, F_{lh}, F_{sh}, and GPP over G1 under different representations of sub-grid topography: D (blue), T (orange), and 1KM (yellow). The bar color is semitransparent to show the overlapping regions.

3.3. Effects on Surface Energy Balance

Solar radiation parameterizations (i.e., TOP and PP) have relatively small impacts on surface energy balance in the coarse grid scale (0.5°). The differences in mean and std between TOP and PP for the D and T configurations (represented by blue and orange color bars in Figure 7) are small and thus, the results for D_PP and T_PP are excluded from the subsequent analysis. For the 1KM simulation, the mean differences between TOP and PP are relatively small (Figure 7), but the std differences cannot be neglected. For instance, the std differences in R^s_{net} and T_s can reach up to 7.8 W/m^2 and 0.45 K, respectively (Figures 7b and 7f).

The representations of sub-grid topography have some impacts on the mean values of surface energy balance terms and turbulent heat flux for all four grids (i.e., G1–G4). Taking G1 as an example (Figure 8), for R^s_{net}, D_TOP can have differences larger than 14 W/m^2 compared to 1KM_TOP, and T_TOP is closer to 1KM_TOP in winter, while their differences are small in other seasons. Compared to 1KM_TOP, for f_{snow}, the biases of both D_TOP and T_TOP can be larger than 0.1 in winter. For T_s, their negative biases can reach up to 0.8 K for all seasons. For F_{lh} and F_{sh}, the biases can be as large as 6–8 W/m^2 for all seasons, and for GPP, D_TOP can have a large bias of above 0.7 μmol m^{-2} s^{-1} in summer, while the bias of T_TOP is smaller than 0.4 μmol m^{-2} s^{-1}. For all the variables, 1KM_PP shows small differences from 1KM_TOP for all seasons. Similar results are obtained from G2-G4 (Figures S1–S3 in Supporting Information S1). These grid-scale differences are generally smaller than the differences between TOP and PP at the 1 km scales (Figures 4–6).

The std and *cv* values of surface energy balance and turbulent heat flux within one grid show large differences under different representations of sub-grid topography (Figure 9 and Figures S4–S7 in Supporting Information S1). For std over G1, the magnitudes of nearly all variables under different cases have large differences. For

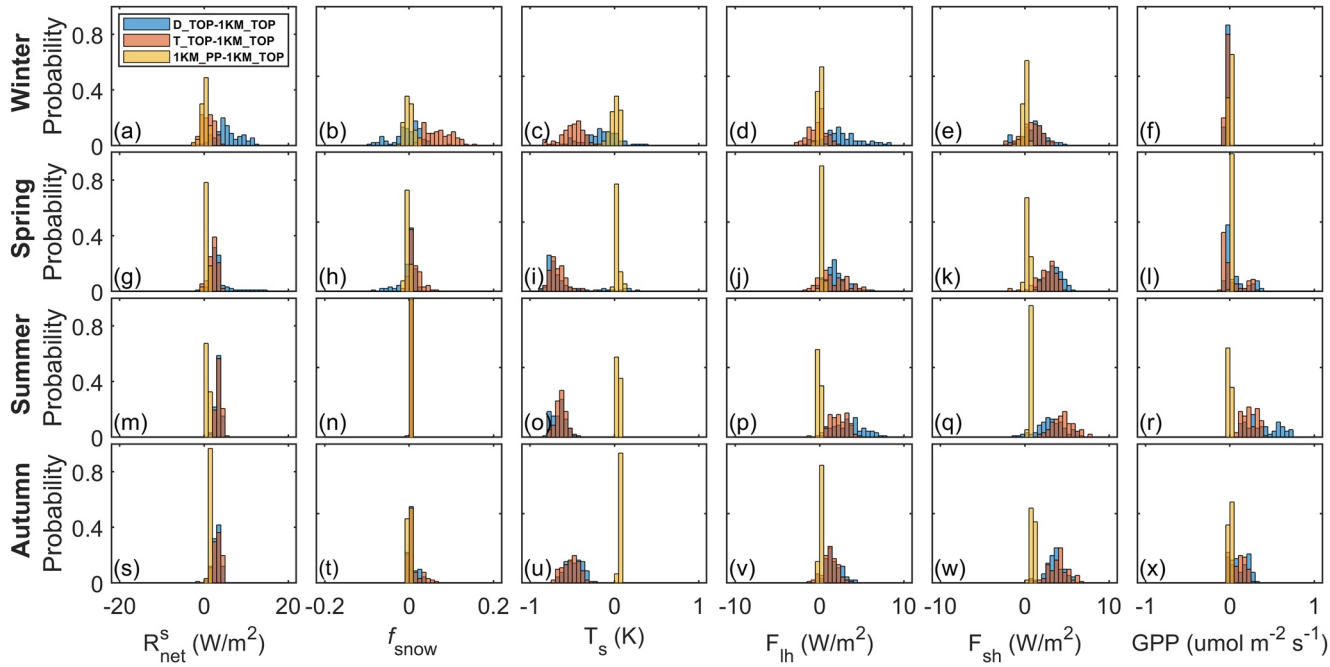

Figure 8. Statistical histograms of the mean differences in R^s_{net}, f_{snow}, T_s, F_{lh}, F_{sh}, and GPP between different cases of G1 for different seasons. Here, the differences were calculated as the differences between other cases and 1KM_TOP. The bar color is semitransparent to show the overlapping regions.

R^s_{net}, f_{snow}, T_s, and GPP, 1KM_TOP has larger std values than other cases, while for F_{lh} and F_{sh}, 1KM_TOP has smaller std values than D_TOP and T_TOP. Overall, T_TOP shows better agreements with the reference 1KM_TOP than D_TOP. For instance, the maximum f_{snow} biases of D_TOP and T_TOP are 0.22 and 0.11, respectively, and the GPP bias of D_TOP can be larger than 1 μmol m^{-2} s^{-1} in spring and summer, while the bias of T_TOP is within 0.3 μmol m^{-2} s^{-1}. Different from the mean, the std for 1KM_PP has some differences with 1KM_TOP

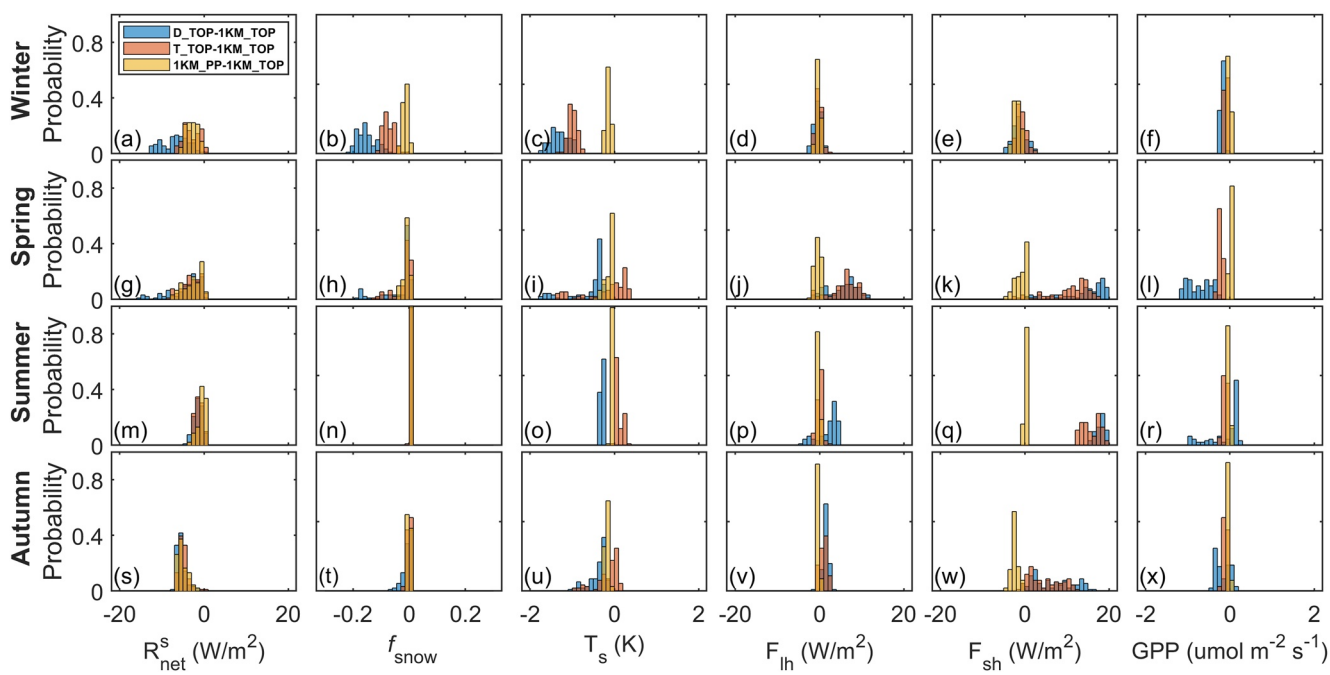

Figure 9. Same as Figure 6, except for standard deviation (std).

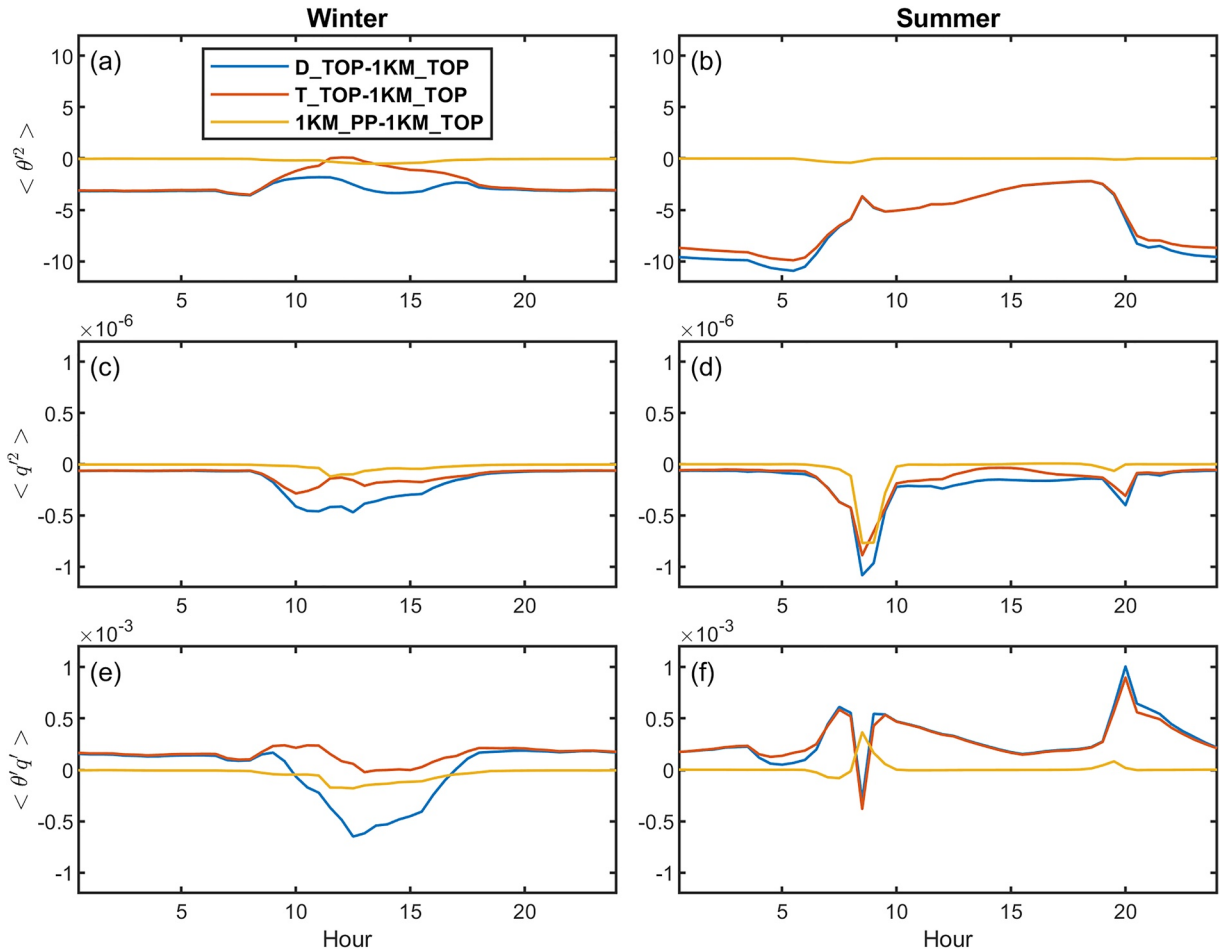

Figure 10. Seasonally averaged diurnal cycles of the difference in simulated (a and b) temperature variance ($\langle\theta'^2\rangle$), (c and d) humidity variance ($\langle q'^2\rangle$), and (e and f) temperature-humidity covariance ($\langle\theta' q'\rangle$) for G1 under different model configurations in winter (DJF) and summer (JJA). Here, the local solar time is used and the scalar (co-)variances were derived using the HET methods introduced in Section 2.4. The differences were calculated as the differences between other model configurations and 1KM_TOP.

especially for R^s_{net}. Specifically in winter and spring, the maximum mean and std differences in R^s_{net} between 1KM_PP and 1KM_TOP are 2.6 and 7.8 W/m^2, respectively. Similar results for G2–G4 are obtained (Figures S4–S6 in Supporting Information S1). For *cv*, the differences between different cases in nearly all variables are more significant in winter than in summer (Figure S7 in Supporting Information S1), and similarly, T_TOP is more consistent with 1KM_TOP than with D_TOP.

3.4. Effects on Surface Boundary Conditions for Scalar (Co-)variances

The sub-grid topographic representations in ELM have large impacts on the surface boundary conditions of scalar (co-)variances especially in winter (Figures 10 and 11). The comparisons of the seasonally averaged diurnal cycles of scalar (co-)variances under different cases show that T_TOP is closer to the reference case than D_TOP for both the HET and HOM methods in both summer and winter (Figure 10 and Figure S8 in Supporting Information S1). In winter, the biases of D_TOP and T_TOP in the three (co-)variances are smaller than zero and follow similar diurnal cycles. The biases of $\langle\theta'^2\rangle$ are the smallest at noon and the largest at night, while those of $\langle q'^2\rangle$ and $\langle\theta' q'\rangle$ are the largest at noon and the smallest at night. T_TOP shows smaller biases during daytime than D_TOP. In summer, the biases in $\langle q'^2\rangle$ and $\langle\theta' q'\rangle$ are larger in the morning and evening (Figures 10d and 10f), possibly caused by the large variability of humidity during these periods. For the accuracy assessment of the HET methods in winter, both D_TOP and T_TOP have lower agreement with 1KM_TOP than 1_km_PP, and D_TOP

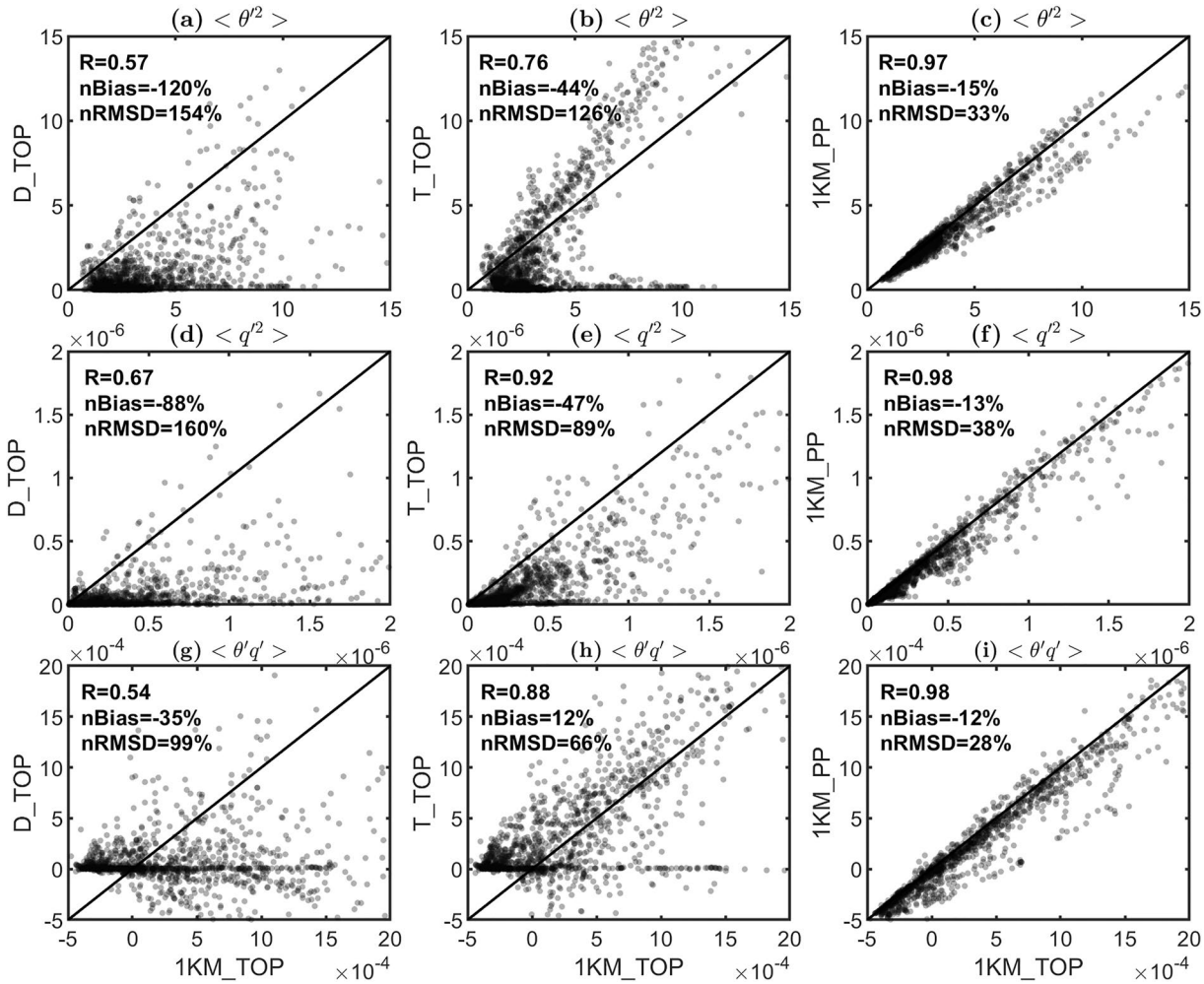

Figure 11. Comparisons of 1KM_TOP simulated (a–c) temperature variance ($\langle\theta'^2\rangle$), (d and e) humidity variance ($\langle q'^2\rangle$), and (g–i) temperature-humidity covariance ($\langle\theta' q'\rangle$) in winter (DJF) over G1 against simulated values from D_TOP, T_TOP, and 1KM_PP model configurations for the daytime (8 a.m.–16 p.m., local solar time). The (co-)variances were derived using the HET methods introduced in Section 2.4; and R, nBias, and nRMSD were shown in each subplot.

has the lowest R values and the highest nBias and nRMSD values for all scalar (co)-variances especially for the daytime (Figure 11 and Figure S9 in Supporting Information S1). For instance, for $\langle\theta'^2\rangle$ during the daytime and nighttime, the nBias value of D_TOP is −146%, while that of T_TOP is −115% (Figure S13 in Supporting Information S1). During the daytime (Figure 11), the nBias values of D_TOP and T_TOP are −120% and −44%, respectively. 1KM_PP is more consistent with 1KM_TOP and the R values are higher than 0.95 and the nBias values are ∼10% for all the three (co-)variances. The sub-grid topographic effects on local solar radiation in winter lead to some differences in scalar (co-)variances between 1KM_PP and 1KM_TOP (Figures 11c, 11f and 11i). Similar results are obtained for G2–G4 in winter (Figures S10–S12 in Supporting Information S1). For the HET methods in summer (Figure S13 in Supporting Information S1), T_TOP also shows slightly better agreements with 1KM_TOP than D_TOP in all three statistical metrics. The 1KM_PP has high correlations and low bias for all three scalar (co)-variances. While the nRMSD values for 1KM_PP are low for $\langle\theta'^2\rangle$ and $\langle\theta' q'\rangle$, they are large for $\langle q'^2\rangle$. For the HOM methods, T_TOP performs better than D_TOP in winter (Figure S14 in Supporting Information S1) and all cases show high correlations with 1KM_TOP in summer (Figure S15 in Supporting Information S1).

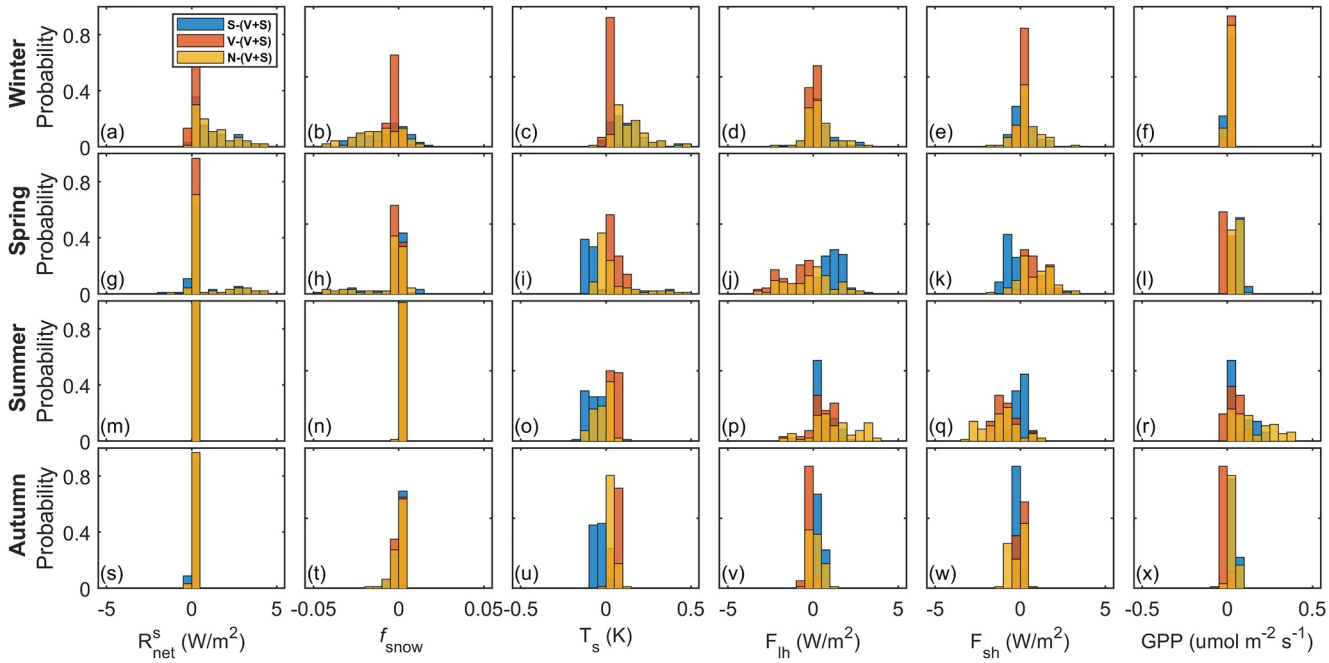

Figure 12. Statistical histograms of the differences between different cases (V + S, S, V, and N) in the mean values of R^s_{net}, f_{snow}, T_s, F_{lh}, F_{sh}, and GPP within one grid of G1 for different seasons. Here, the V + S case is used as a reference and thus the differences were calculated as the differences between other cases and V + S case. The bar color is semitransparent to show the overlapping regions.

3.5. Roles of Vegetation and Soil Heterogeneities

The topounit-scale sensitivity experiments show that vegetation heterogeneity accounts for most of the variations for nearly all the variables (Figures 12 and 13). Generally for R^s_{net}, f_{snow}, and T_s, the mean and std values of the V case are closer to the V + S case than to the S case, while those of the S case are closer to the N case.

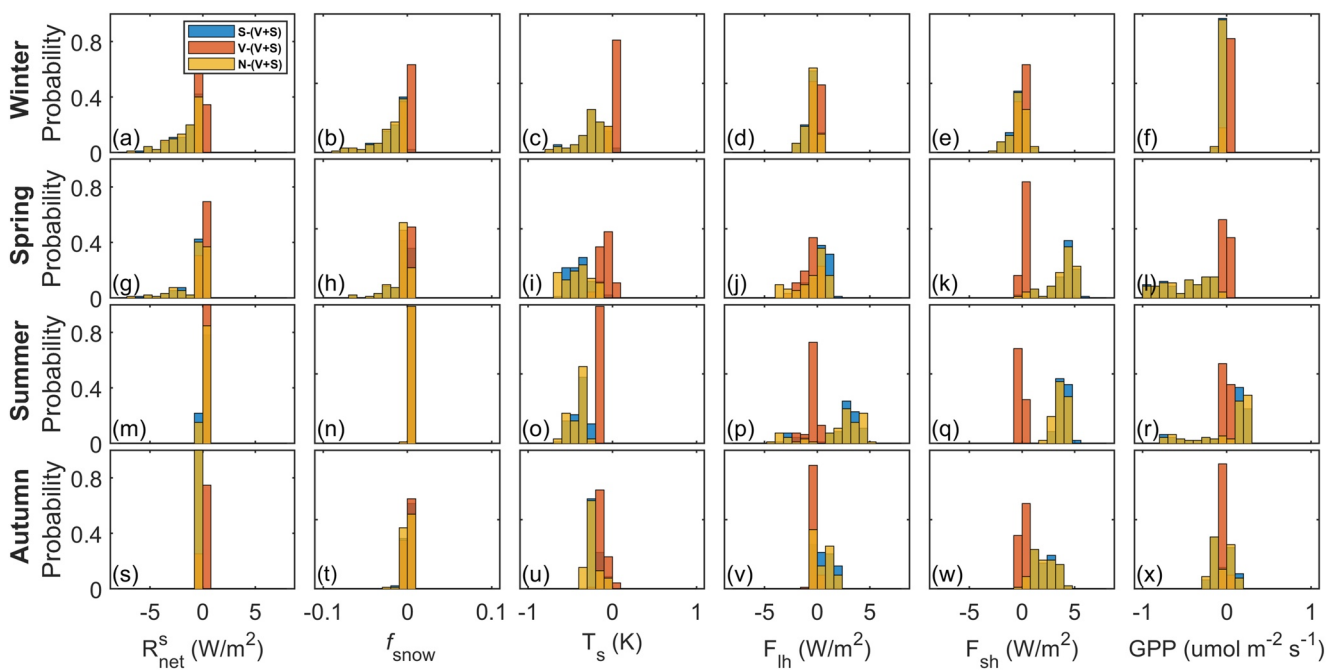

Figure 13. Same as Figure 12, except for std.

These demonstrate that vegetation heterogeneities contribute more to the variability of R^s_{net}, f_{snow}, and T_s than soil heterogeneity. For F_{lh}, F_{sh}, and GPP, both soil and vegetation heterogeneities contribute to the differences in mean values (Figure 12), but vegetation heterogeneity mainly accounts for the differences in std values (Figure 13). These demonstrate that topography-relevant PFT distributions, LAI and SAI characteristics contribute a lot to the mean and std values of the surface energy balance terms.

4. Discussion

Overall, 1KM_TOP well captures the spatial heterogeneity of surface energy fluxes compared to MODIS data. However, there are some differences between 1KM_TOP and MODIS (Figure 3). On one hand, uncertainties from the forcing data and model structure in ELM can contribute to the simulation bias. On the other hand, as discussed in Hao et al. (2021), most MODIS algorithms neglect the topographic effects and thus there are some uncertainties in the generated MODIS data especially over rugged terrain. Besides, the cloud and cloud shadow, illumination-viewing geometrical effects, and instrumental errors can also affect the accuracy of MODIS data (Bair et al., 2021). Further efforts are needed to improve the model performance of ELM and accuracy of MODIS data over rugged terrain.

Sub-grid topography has large impacts on the spatial distributions of surface energy balance. The spatial patterns of the differences between TOP and PP generally follow the topographic distributions (Figures 4–6). The sub-grid topographic effects on radiation are significant at a spatial resolution of 1 km. When aggregated to a coarse scale (e.g., 0.5° in this study), the differences between TOP and PP become small because the differences for sunny and shaded slopes offset each other (Zhao et al., 2016), but they still cannot be neglected especially in winter (Figure 8). The differences between TOP and PP are also dependent on the seasons (Figure 5) and are more pronounced in winter due to the snow cover and the strong shadowing effects caused by large solar zenith angles (Hao et al., 2021). These results are similar to previous studies over Sierra Nevada that used WRF (Gu et al., 2012; Liou et al., 2013) and CLM4 (Lee et al., 2015). These differences in surface energy balance can further affect atmospheric processes, such as cloud formation and precipitation (Lee et al., 2015; Liou et al., 2013), that need further investigations by performing E3SM simulations with active atmosphere and land components. GPP is also affected by the sub-grid topography (Figure 6), which underscores the necessity of accounting for sub-grid topographic heterogeneity when simulating GPP over complex terrain (Xie et al., 2021). However, the lateral surface and sub-surface flows from uplands to lowlands were neglected in the simulations reported here, which can also impact the simulation of energy and water fluxes in LSMs (Fan et al., 2019; Ji et al., 2017).

Different representations of sub-grid topography in LSMs affect the surface energy balance and surface boundary conditions. The mean values of surface energy balance terms and turbulent heat flux are affected by the representations of sub-grid topography (Figure 8). The std values also show large dependences on the representations of sub-grid topography (Figure 9) and these results are consistent with those reported in Liu et al. (2017). The high-order scalar (co-)variances for the HOM methods are identical for the different representations of sub-grid topography (Figures S14 and S15 in Supporting Information S1), while there are large differences between different representations of sub-grid topography for the HET methods (Figures 10 and 11). The large differences in the HET methods are expected to result in large differences in surface boundary conditions, which could further affect the PBL, atmospheric, and cloud dynamics (Chen et al., 2020). The T configuration shows better agreements with the 1KM simulations than the D configuration because topounit can better describe the topography-dependent vegetation and soil distributions. In our study area, the differences between the D and T configurations are mainly contributed by topography-dependent vegetation distribution (Figures 12 and 13). However, how the topography-dependent vegetation and soil heterogeneities affect the surface energy balance in other regions needs further investigation. These underline the feasibility and importance of implementing the sub-grid topographic structure in LSMs and are promising to guide model development and experimental designs.

Further optimizing the representations of sub-grid topography is possible as more high-resolution data sets of land surface parameters are available. In the current scheme of topounit in ELM, the topounit number for one grid is spatially variable, the maximum possible number of topounits is set as 12, and the actual number of topounit is determined only by the topographic complexity within a grid (Tesfa & Leung, 2017). For instance, the four grids over the study area include 11 topounits (Figure 2). A spatially variable number of topounits within a

grid will reduce the computational burden while attempting to accurately capture the sub-grid heterogeneity of topography. To what degree topounit can mimic the realistic conditions and the dependence on the number of topounits should be analyzed to further improve the performance of T_TOP in simulating surface energy balance and surface boundary conditions. There exist other approaches to generate topounits to parsimoniously capture sub-grid heterogeneity in topography that account for multiple sources of sub-grid heterogeneities. A topographic sub-grid tool for mountain (TopoSUB) was proposed based on a multidimensional informed sampling procedure, where the elevation, slope, aspect, and sky view factor were used to represent the topographic characteristics (Fiddes & Gruber, 2012). Similarly, a sub-grid classification method (SGC) was also developed to account for both topographic and vegetation variabilities, but it only considered the differences in elevation (Ke et al., 2013). In addition, a hydrological response unit (HRU)-based structure was developed based on the *K*-means clustering algorithm (Chaney et al., 2016). Different HRUs have different characteristics of LULC, soil, topography, and meteorology forcing. For improving our understanding of the Earth system over heterogeneous terrain, the sub-grid schemes in LSMs need to fully harness emerging big data from field measurements and remote sensing (Chaney et al., 2018).

There are still some limitations in the study. First, the 0.5° GSWP3 forcing data used in the offline ELM simulations are too coarse to capture the sub-grid heterogeneity due to topography in atmospheric forcings (Tesfa et al., 2020). Meteorological forcing data could account for a large proportion of turbulent heat flux in some cases (Jason Scot et al., 2021). Some atmospheric forcing downscaling methods for mountainous areas (Fiddes & Gruber, 2014; Fiddes et al., 2021; Tesfa et al., 2020) and high-resolution forcing data (e.g., Daymet) (Thornton et al., 2016) make it possible to further consider the topography-induced sub-grid heterogeneity of atmospheric forcings. Especially, Tesfa et al. (2020) developed topography-based methods for downscaling grid-scale precipitation to sub-grid scales, which are essentially coupled with the topounit-based sub-grid structure in ELM for improving land surface modeling over mountainous areas. Second, only offline ELM simulations were performed in this study and the effects of sub-grid topographic heterogeneity on the vertical structure of PBL and atmospheric dynamics need to be further investigated using the Cloud Layers Unified By Binormals model coupled with the E3SM atmospheric model (Xie et al., 2018). Third, generalization of the conclusion from this model-based analysis over four 0.5° × 0.5° grid cells to other regions with heterogeneous terrain needs to be further evaluated. Ongoing and future pioneering E3SM projects, for example, the 1 km gridded ELM implementation over the North American region under a hybrid CPU-GPU architecture of the Summit supercomputer and the global 3.25 km simulations in the DYnamics of the Atmospheric general circulation Modeled On Nonhydrostatic Domains Phase 2 model intercomparison (Caldwell et al., 2021), offer good opportunity to extend our findings.

5. Conclusions

The heterogeneity of sub-grid topography exerts large influence on many land surface processes and further affects atmospheric dynamics. This study used ELM to investigate the impacts of sub-grid topographic representations on surface energy balance and surface boundary conditions for turbulent heat flux and scalar (co-)variances. A series of offline simulations with three different sub-grid topographic representations (i.e., D, T, and 1KM) and two different treatments of the sub-grid topographic effects on solar radiation (i.e., TOP and PP) were carried out in four representative 0.5° grids. Compared to MODIS data, overall, the 1KM simulations with TOP well capture the spatial heterogeneity of surface fluxes. Topography has large impacts on the spatial distribution of surface energy balance terms and TOP and PP have significant differences at a resolution of 1 km in simulating surface energy balance, f_{snow} and T_s, which depend on seasons and local topography. The differences in the mean values are relatively small when aggregated to 0.5° grid scales because the positive-negative differences for sunny and shaded slopes offset each other. Different representations of sub-grid topography can also affect the surface energy balance and surface boundary conditions. As compared to the D configuration, the T configuration can more accurately capture the effects of sub-grid topographic heterogeneity on the mean values and standard deviations of surface energy balance terms as well as surface boundary conditions for turbulent heat flux and scalar (co-)variances. The findings in this study underline the importance of improving the representations of the sub-grid topographic heterogeneity in LSMs over complex terrain.

Data Availability Statement

The ELM codes are available publicly at https://github.com/E3SM-Project/E3SM (last accessed: 28 September 2021) and the code for TOP solar radiation parameterizations used in this paper is available at http://doi.org/10.5281/zenodo.4549401. Codes and data to reproduce results and plot figures are publicly available at http://doi.org/10.5281/zenodo.5908170. All MODIS, WorldClim V1, Soilgrid v2, and SRTM DEM data introduced in Table 1 are publicly available at Google Earth Engine (https://earthengine.google.com/).

Acknowledgments
This research has been supported by the U.S. Department of Energy (DOE), Office of Science, Office of Biological and Environmental Research, Earth System Model Development program area, as part of the Climate Process Team projects. This study was also supported by the Atmospheric Administration (NOAA) under award numbers NOAA-OAR-CPO-2019-2005530 and NA19OAR4310243. Y. Gu acknowledges the support by (while serving at) the National Science Foundation. The reported research was conducted at the Pacific Northwest National Laboratory, which is operated for the DOE by the Battelle Memorial Institute under contract DE-AC05-76RL01830. This research used the computational resources of the National Energy Research Scientific Computing Center, a DOE Office of Science User Facility supported by the Office of Science of the DOE.

References

André, J., De Moor, G., Lacarrere, P., & Du Vachat, R. (1978). Modeling the 24-hour evolution of the mean and turbulent structures of the planetary boundary layer. *Journal of the Atmospheric Sciences*, *35*(10), 1861–1883. https://doi.org/10.1175/1520-0469(1978)035<1861:mtheot>2.0.co;2

Avissar, R., & Chen, F. (1993). Development and analysis of prognostic equations for mesoscale kinetic energy and mesoscale (subgrid scale) fluxes for large-scale atmospheric models. *Journal of the Atmospheric Sciences*, *50*(22), 3751–3774. https://doi.org/10.1175/1520-0469(1993)050<3751:daaope>2.0.co;2

Bair, E. H., Stillinger, T., & Dozier, J. (2021). Snow property inversion from remote sensing (SPIReS): A generalized multispectral unmixing approach with examples from MODIS and Landsat 8 OLI. *IEEE Transactions on Geoscience and Remote Sensing*, *59*(9), 7270–7284. https://doi.org/10.1109/tgrs.2020.3040328

Bisht, G., Riley, W. J., Hammond, G. E., & Lorenzetti, D. M. (2018). Development and evaluation of a variably saturated flow model in the global E3SM land model (ELM) version 1.0. *Geoscientific Model Development Discussions*, *11*(10), 4085–4102. https://doi.org/10.5194/gmd-11-4085-2018

Bonan, G. B., Levis, S., Kergoat, L., & Oleson, K. W. (2002). Landscapes as patches of plant functional types: An integrating concept for climate and ecosystem models. *Global Biogeochemical Cycles*, *16*(2), 5-1–5-23. https://doi.org/10.1029/2000gb001360

Bou-Zeid, E., Anderson, W., Katul, G. G., & Mahrt, L. (2020). The persistent challenge of surface heterogeneity in boundary-layer meteorology: A review. *Boundary-Layer Meteorology*, *177*(2), 227–245. https://doi.org/10.1007/s10546-020-00551-8

Brunsell, N. A., Mechem, D. B., & Anderson, M. C. (2011). Surface heterogeneity impacts on boundary layer dynamics via energy balance partitioning. *Atmospheric Chemistry and Physics*, *11*(7), 3403–3416. https://doi.org/10.5194/acp-11-3403-2011

Caldwell, P. M., Terai, C. R., Hillman, B., Keen, N. D., Bogenschutz, P., Lin, W., et al. (2021). Convection-permitting simulations with the E3SM global atmosphere model. *Journal of Advances in Modeling Earth Systems*, *13*(11), e2021MS002544. https://doi.org/10.1029/2021ms002544

Chaney, N. W., Metcalfe, P., & Wood, E. F. (2016). HydroBlocks: A field-scale resolving land surface model for application over continental extents. *Hydrological Processes*, *30*(20), 3543–3559. https://doi.org/10.1002/hyp.10891

Chaney, N. W., Van Huijgevoort, M. H. J., Shevliakova, E., Malyshev, S., Milly, P. C. D., Gauthier, P. P. G., & Sulman, B. N. (2018). Harnessing big data to rethink land heterogeneity in Earth system models. *Hydrology and Earth System Sciences*, *22*(6), 3311–3330. https://doi.org/10.5194/hess-22-3311-2018

Chen, J., Hagos, S., Xiao, H., Fast, J. D., & Feng, Z. (2020). Characterization of surface heterogeneity-induced convection using cluster analysis. *Journal of Geophysical Research: Atmospheres*, *125*(20), e2020JD032550. https://doi.org/10.1029/2020jd032550

Dearborn, K. D., & Danby, R. K. (2017). Aspect and slope influence plant community composition more than elevation across forest–tundra ecotones in subarctic Canada. *Journal of Vegetation Science*, *28*(3), 595–604. https://doi.org/10.1111/jvs.12521

de Vrese, P., Schulz, J.-P., & Hagemann, S. (2016). On the representation of heterogeneity in land-surface–atmosphere coupling. *Boundary-Layer Meteorology*, *160*(1), 157–183. https://doi.org/10.1007/s10546-016-0133-1

Dirmeyer, P. A., Gao, X., Zhao, M., Guo, Z., Oki, T., & Hanasaki, N. (2006). GSWP-2: Multimodel analysis and implications for our perception of the land surface. *Bulletin of the American Meteorological Society*, *87*(10), 1381–1398. https://doi.org/10.1175/bams-87-10-1381

Dubayah, R. (1992). Estimating net solar radiation using Landsat Thematic Mapper and digital elevation data. *Water Resources Research*, *28*(9), 2469–2484. https://doi.org/10.1029/92wr00772

Fan, Y., Clark, M., Lawrence, D. M., Swenson, S., Band, L. E., Brantley, S. L., et al. (2019). Hillslope hydrology in global change research and Earth system modeling. *Water Resources Research*, *55*(2), 1737–1772. https://doi.org/10.1029/2018wr023903

Fiddes, J., Aalstad, K., & Lehning, M. (2021). TopoCLIM: Rapid topography-based downscaling of regional climate model output in complex terrain v.1.0. *Geoscientific Model Development*, *2021*, 1–27.

Fiddes, J., & Gruber, S. (2012). TopoSUB: A tool for efficient large area numerical modelling in complex topography at sub-grid scales. *Geoscientific Model Development*, *5*(5), 1245–1257. https://doi.org/10.5194/gmd-5-1245-2012

Fiddes, J., & Gruber, S. (2014). TopoSCALE v.1.0: Downscaling gridded climate data in complex terrain. *Geoscientific Model Development*, *7*(1), 387–405. https://doi.org/10.5194/gmd-7-387-2014

Fisher, R. A., & Koven, C. D. (2020). Perspectives on the future of land surface models and the challenges of representing complex terrestrial systems. *Journal of Advances in Modeling Earth Systems*, *12*(4), e2018MS001453. https://doi.org/10.1029/2018ms001453

Friedl, M. A., McIver, D., Hodges, J., Zhang, X., Muchoney, D., Strahler, A., et al. (2002). Global land cover mapping from MODIS: Algorithms and early results. *Remote Sensing of Environment*, *83*(1), 287–302. https://doi.org/10.1016/s0034-4257(02)00078-0

Giorgi, F., & Avissar, R. (1997). Representation of heterogeneity effects in Earth system modeling: Experience from land surface modeling. *Reviews of Geophysics*, *35*(4), 413–437. https://doi.org/10.1029/97rg01754

Golaz, J.-C., Caldwell, P. M., Van Roekel, L. P., Petersen, M. R., Tang, Q., Wolfe, J. D., et al. (2019). The DOE E3SM coupled model version 1: Overview and evaluation at standard resolution. *Journal of Advances in Modeling Earth Systems*, *11*(7), 2089–2129.

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., & Moore, R. (2017). Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sensing of Environment*, *202*, 18–27. https://doi.org/10.1016/j.rse.2017.06.031

Gu, Y., Liou, K. N., Lee, W. L., & Leung, L. R. (2012). Simulating 3-D radiative transfer effects over the Sierra Nevada Mountains using WRF. *Atmospheric Chemistry and Physics*, *12*(20), 9965–9976. https://doi.org/10.5194/acp-12-9965-2012

Guo, H., Golaz, J., Donner, L. J., Wyman, B., Zhao, M., & Ginoux, P. (2015). CLUBB as a unified cloud parameterization: Opportunities and challenges. *Geophysical Research Letters*, *42*(11), 4540–4547. https://doi.org/10.1002/2015gl063672

Hall, D. K., Riggs, G. A., Salomonson, V. V., DiGirolamo, N. E., & Bayr, K. J. (2002). MODIS snow-cover products. *Remote Sensing of Environment*, *83*(1), 181–194. https://doi.org/10.1016/s0034-4257(02)00095-0

Hao, D., Bisht, G., Gu, Y., Lee, W. L., Liou, K. N., & Leung, L. R. (2021). 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*, *14*(10), 1–23. https://doi.org/10.5194/gmd-14-6273-2021

Hao, D., Wen, J., Xiao, Q., Wu, S., Lin, X., You, D., & Tang, Y. (2018). Modeling anisotropic reflectance over composite sloping terrain. *IEEE Transactions on Geoscience and Remote Sensing*, *56*(7), 3903–3923. https://doi.org/10.1109/tgrs.2018.2816015

Hao, D., Wen, J., Xiao, Q., Wu, S., Lin, X., You, D., & Tang, Y. (2019). Impacts of DEM geolocation bias on downward surface shortwave radiation estimation over clear-sky rugged terrain: A case study in Dayekou basin, China. *IEEE Geoscience and Remote Sensing Letters*, *16*(1), 10–14. https://doi.org/10.1109/lgrs.2018.2868563

Hengl, T., Mendes de Jesus, J., Heuvelink, G. B. M., Ruiperez Gonzalez, M., Kilibarda, M., Blagotic, A., et al. (2017). SoilGrids250m: Global gridded soil information based on machine learning. *PLoS One*, *12*(2), e0169748. https://doi.org/10.1371/journal.pone.0169748

Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G., & Jarvis, A. (2005). Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, *25*(15), 1965–1978. https://doi.org/10.1002/joc.1276

Jason Scot, S., Andrew, D. B., Paul, A. D., & Nathaniel, W. C. (2021). Semi-coupling of a field-scale resolving land-surface model and WRF-LES to investigate the influence of land-surface heterogeneity on cloud development. *Earth and Space Science Open Archive*, *13*(10), e2021MS002602. https://doi.org/10.1029/2021MS002602

Ji, P., Yuan, X., & Liang, X.-Z. (2017). Do lateral flows matter for the hyperresolution land surface modeling? *Journal of Geophysical Research: Atmospheres*, *122*(2212), 12077–12092. https://doi.org/10.1002/2017jd027366

Ke, Y., Leung, L. R., Huang, M., Coleman, A. M., Li, H., & Wigmosta, M. S. (2012). Development of high resolution land surface parameters for the Community Land Model. *Geoscientific Model Development*, *5*(6), 1341–1362. https://doi.org/10.5194/gmd-5-1341-2012

Ke, Y., Leung, L. R., Huang, M., & Li, H. (2013). Enhancing the representation of subgrid land surface characteristics in land surface models. *Geoscientific Model Development*, *6*(5), 1609–1622. https://doi.org/10.5194/gmd-6-1609-2013

Lawrence, D. M., Fisher, R. A., Koven, C. D., Oleson, K. W., Swenson, S. C., Bonan, G., et al. (2019). The community land model version 5: Description of new features, benchmarking, and impact of forcing uncertainty. *Journal of Advances in Modeling Earth Systems*, *11*(12), 4245–4287. https://doi.org/10.1029/2018ms001583

Lee, J. M., Zhang, Y., & Klein, S. A. (2019). The effect of land surface heterogeneity and background wind on shallow cumulus clouds and the transition to deeper convection. *Journal of the Atmospheric Sciences*, *76*(2), 401–419. https://doi.org/10.1175/jas-d-18-0196.1

Lee, W. L., Gu, Y., Liou, K. N., Leung, L. R., & Hsu, H. H. (2015). A global model simulation for 3-D radiative transfer impact on surface hydrology over the Sierra Nevada and Rocky Mountains. *Atmospheric Chemistry and Physics*, *15*(10), 5405–5413. https://doi.org/10.5194/acp-15-5405-2015

Lee, W.-L., Liou, K., Wang, C., Gu, Y., Hsu, H., & Li, J. F. (2019). 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*(3), 1537–1549. https://doi.org/10.1029/2018jd029592

Lee, W.-L., Liou, K. N., & Hall, A. (2011). Parameterization of solar fluxes over mountain surfaces for application to climate models. *Journal of Geophysical Research*, *116*(D1), D01101. https://doi.org/10.1029/2010jd014722

Leung, L. R., Bader, D. C., Taylor, M. A., & McCoy, R. B. (2020). An introduction to the E3SM special collection: Goals, science drivers, development, and analysis. *Journal of Advances in Modeling Earth Systems*, *12*(11), e2019MS001821. https://doi.org/10.1029/2019ms001821

Levy, P., Drewer, J., Jammet, M., Leeson, S., Friborg, T., Skiba, U., & Oijen, M. v. (2020). Inference of spatial heterogeneity in surface fluxes from eddy covariance data: A case study from a subarctic mire ecosystem. *Agricultural and Forest Meteorology*, *280*, 107783. https://doi.org/10.1016/j.agrformet.2019.107783

Liou, K. N., Gu, Y., Leung, L. R., Lee, W. L., & Fovell, R. G. (2013). A WRF simulation of the impact of 3-D radiative transfer on surface hydrology over the Rocky Mountains and Sierra Nevada. *Atmospheric Chemistry and Physics*, *13*(23), 11709–11721. https://doi.org/10.5194/acp-13-11709-2013

Liu, S., Shao, Y., Kunoth, A., & Simmer, C. (2017). Impact of surface-heterogeneity on atmosphere and land-surface interactions. *Environmental Modelling & Software*, *88*, 35–47. https://doi.org/10.1016/j.envsoft.2016.11.006

Lyons, T. J., & Halldin, S. (2004). Surface heterogeneity and the spatial variation of fluxes. *Agricultural and Forest Meteorology*, *121*(3), 153–165. https://doi.org/10.1016/j.agrformet.2003.08.031

Machulskaya, E., & Mironov, D. (2018). Boundary conditions for scalar (co)variances over heterogeneous surfaces. *Boundary-Layer Meteorology*, *169*(1), 139–150. https://doi.org/10.1007/s10546-018-0354-6

Maronga, B., & Raasch, S. (2013). Large-Eddy simulations of surface heterogeneity effects on the convective boundary layer during the LITFASS-2003 experiment. *Boundary-Layer Meteorology*, *146*(1), 17–44. https://doi.org/10.1007/s10546-012-9748-z

Mu, Q., Heinsch, F. A., Zhao, M., & Running, S. W. (2007). Development of a global evapotranspiration algorithm based on MODIS and global meteorology data. *Remote Sensing of Environment*, *111*(4), 519–536. https://doi.org/10.1016/j.rse.2007.04.015

Myneni, R. B., Hoffman, S., Knyazikhin, Y., Privette, J., Glassy, J., Tian, Y., et al. (2002). Global products of vegetation leaf area and fraction absorbed PAR from year one of MODIS data. *Remote Sensing of Environment*, *83*(1), 214–231. https://doi.org/10.1016/s0034-4257(02)00074-3

Poggio, L., de Sousa, L. M., Batjes, N. H., Heuvelink, G. B. M., Kempen, B., Ribeiro, E., & Rossiter, D. (2021). SoilGrids 2.0: Producing soil information for the globe with quantified spatial uncertainty. *SOIL*, *7*(1), 217–240. https://doi.org/10.5194/soil-7-217-2021

Proy, C., Tanré, D., & Deschamps, P. Y. (1989). Evaluation of topographic effects in remotely sensed data. *Remote Sensing of Environment*, *30*(1), 21–32. https://doi.org/10.1016/0034-4257(89)90044-8

Rabus, B., Eineder, M., Roth, A., & Bamler, R. (2003). The shuttle radar topography mission—A new class of digital elevation models acquired by spaceborne radar. *ISPRS Journal of Photogrammetry and Remote Sensing*, *57*(4), 241–262. https://doi.org/10.1016/s0924-2716(02)00124-7

Román-Cascón, C., Marie, L., Fabienne, L., & Oscar, H. (2021). Surface representation impacts on turbulent heat fluxes in WRF (v.4.1.3). *Geoscientific Model Development Discussions*, *2021*, 1–51.

Ropars, P., & Boudreau, S. (2012). Shrub expansion at the forest–tundra ecotone: Spatial heterogeneity linked to local topography. *Environmental Research Letters*, *7*(1), 015501. https://doi.org/10.1088/1748-9326/7/1/015501

Running, S. W., Nemani, R. R., Heinsch, F. A., Zhao, M., Reeves, M., & Hashimoto, H. (2004). A continuous satellite-derived measure of global terrestrial primary production. *BioScience*, *54*(6), 547–560. https://doi.org/10.1641/0006-3568(2004)054[0547:acsmog]2.0.co;2

Schaaf, C. B., Gao, F., Strahler, A. H., Lucht, W., Li, X., Tsang, T., et al. (2002). First operational BRDF, albedo nadir reflectance products from MODIS. *Remote Sensing of Environment*, *83*(1), 135–148. https://doi.org/10.1016/s0034-4257(02)00091-3

Schneider, T., Lan, S., Stuart, A., & Teixeira, J. (2017). Earth system modeling 2.0: A blueprint for models that learn from observations and targeted high-resolution simulations. *Geophysical Research Letters*, *44*(24), 12396–12417. https://doi.org/10.1002/2017gl076101

Sellers, P. J. (1985). Canopy reflectance, photosynthesis and transpiration. *International Journal of Remote Sensing*, *6*(8), 1335–1372. https://doi.org/10.1080/01431168508948283

Still, C. J., Berry, J. A., Collatz, G. J., & DeFries, R. S. (2003). Global distribution of C3 and C4 vegetation: Carbon cycle implications. *Global Biogeochemical Cycles*, *17*(1), 6-1–6-14. https://doi.org/10.1029/2001gb001807

Tesfa, T. K., Leung, L. R., & Ghan, S. J. (2020). Exploring topography-based methods for downscaling subgrid precipitation for use in Earth system models. *Journal of Geophysical Research: Atmospheres*, *125*(5), e2019JD031456. https://doi.org/10.1029/2019jd031456

Tesfa, T. K., & Leung, L. Y. R. (2017). Exploring new topography-based subgrid spatial structures for improving land surface modeling. *Geoscientific Model Development*, *10*(2), 873–888. https://doi.org/10.5194/gmd-10-873-2017

Thornton, P., Thornton, M. M., Mayer, B. W., Wei, Y., Devarakonda, R., Vose, R. S., & Cook, R. B. (2016). Daymet: Daily surface weather data on a 1-km grid for North America, version 3. ORNL DAAC, Oak Ridge, Tennessee, USA. *USDA-NASS, 2019. 2017 Census of Agriculture, Summary and State Data, Geographic Area Series, Part 51, AC-17-A-51*.

Wu, Y., Nair, U. S., Pielke, R. A., McNider, R. T., Christopher, S. A., & Anantharaj, V. G. (2009). Impact of land surface heterogeneity on mesoscale Atmospheric dispersion. *Boundary-Layer Meteorology*, *133*(3), 367. https://doi.org/10.1007/s10546-009-9415-1

Xie, S., Lin, W., Rasch, P. J., Ma, P., Neale, R., Larson, V. E., et al. (2018). Understanding cloud and convective characteristics in version 1 of the E3SM atmosphere model. *Journal of Advances in Modeling Earth Systems*, *10*(10), 2618–2644. https://doi.org/10.1029/2018ms001350

Xie, X., Chen, J. M., Gong, P., & Li, A. (2021). Spatial scaling of gross primary productivity over sixteen mountainous watersheds using vegetation heterogeneity and surface topography. *Journal of Geophysical Research: Biogeosciences*, *126*(5), e2020JG005848. https://doi.org/10.1029/2020jg005848

Yang, X., Ricciuto, D. M., Thornton, P. E., Shi, X., Xu, M., Hoffman, F., & Norby, R. J. (2019). The effects of phosphorus cycle dynamics on carbon sources and sinks in the Amazon region: A modeling study using ELM v1. *Journal of Geophysical Research: Biogeosciences*, *124*(12), 3686–3698. https://doi.org/10.1029/2019jg005082

Zeng, X., Shaikh, M., Dai, Y., Dickinson, R. E., & Myneni, R. (2002). Coupling of the common land model to the NCAR community climate model. *Journal of Climate*, *15*(14), 1832–1854. https://doi.org/10.1175/1520-0442(2002)015<1832:cotclm>2.0.co;2

Zhang, N., Williams, Q. L., & Liu, H. (2010). Effects of land-surface heterogeneity on numerical simulations of mesoscale atmospheric boundary layer processes. *Theoretical and Applied Climatology*, *102*(3), 307–317. https://doi.org/10.1007/s00704-010-0268-9

Zhao, B., Liou, K. N., Gu, Y., He, C., Lee, W. L., Chang, X., et al. (2016). Impact of buildings on surface solar radiation over urban Beijing. *Atmospheric Chemistry and Physics*, *16*(9), 5841–5852. https://doi.org/10.5194/acp-16-5841-2016

Zheng, Y., Brunsell, N. A., Alfieri, J. G., & Niyogi, D. (2021). Impacts of land cover heterogeneity and land surface parameterizations on turbulent characteristics and mesoscale simulations. *Meteorology and Atmospheric Physics*, *133*(3), 589–610. https://doi.org/10.1007/s00703-020-00768-9