Scalability of grid- and subbasin-based land surface modeling approaches for hydrologic simulations

Teklu K. Tesfa, L. Ruby Leung, Maoyi Huang, Hong-Yi Li, Nathalie Voisin, and Mark S. Wigmosta

Pacific Northwest National Laboratory, Richland, WA 99352

Corresponding author: Teklu K. Tesfa

Email address: teklu.tesfa@pnnl.gov

Tel.: (509) 372 4479

Submitted to Journal of Geophysical Research - Atmosphere, July 2013

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1002/2013JD020493

Abstract

This paper investigates the relative merits of grid- and subbasin-based land surface modeling approaches for hydrologic simulations, with a focus on their scalability (i.e., ability to perform consistently across spatial resolutions) in simulating runoff generation. Simulations are produced by the grid- and subbasin-based Community Land Model (CLM) at 0.125^o, 0.25^o, 0.5^o and 1^o spatial resolutions over the U.S. Pacific Northwest. Using the 0.125^o simulation as the "reference" solution, statistical metrics are calculated by comparing simulations at 0.25^o, 0.5^o and 1^o resolutions with the 0.125^o simulation aggregated to the respective resolutions for each approach. Statistical significance test results suggest significant scalability advantage for the subbasin-based approach compared to the grid-based approach. Basin level annual average relative errors of surface runoff at 0.25^o, 0.5^o, and 1^o resolutions compared to the 0.125^o simulation are 3%, 4%, and 6% for the subbasin-based configuration and 4%, 7%, and 11% for the grid-based configuration, respectively. The scalability advantages are more pronounced during winter/spring and over mountainous regions. The source of runoff scalability is found to be related to the scalability of major meteorological and land surface parameters of runoff generation. More specifically, the subbasin-based approach is more consistent across spatial scales than the grid-based approach in snowfall/rainfall partitioning because of scalability related to air temperature and surface elevation. Scalability of a topographic parameter used in runoff parameterization also contributes to improved scalability of the rain driven saturated surface runoff component, particularly during winter. Hence this study demonstrates the importance of spatial structure for multi-scale modeling of hydrological processes.

Keywords: Grid-based, Hydrologic simulation, Land surface modeling, Model scalability, Spatial structure, Subbasin-based

1. Introduction

Realistic representation of land surface hydrologic processes is crucial to advancing modeling of land-atmosphere interactions in earth system models. It has been increasingly recognized that land-atmosphere interactions must be addressed from both hydrologic and atmospheric science perspectives [*Sridhar et al.*, 2003]. However, because of differences in modeling approaches used by the land surface and hydrologic science communities, advances made by each community have not been shared effectively to maximize the benefits. For example, most large scale hydrologic models are commonly configured to run on irregular computational units following watershed or subbasin boundaries, which are the natural units for representing hydrologic processes [e.g., *Bicknell et al.*, 1997; *Arnold et al.*, 1998]. In a watershed/subbasin, topography exerts important control on both surface and subsurface flows [*J Chen and Kumar*, 2001; *Beven*, 1997]. Land surface models, on the other hand, are typically configured to run on regular rectangular grids at uniform and relatively coarse spatial scales for coupling with atmospheric models to simulate climate and earth system processes [e.g., *Liang et al.*, 1994; *Chen et al.*, 1996].

The subbasin-based approach offers important advantages over the standard grid-based approach from a hydrologic perspective ([*Tesfa et al.*, 2014] and references therein) including: (1) conceptually, parameterizing runoff generation is more straightforward when the computational units follow topographic boundaries as opposed to the grid-based approach where grids often encompass areas from several natural subbasins, challenging the very concept of runoff generation formulation in parameterizations such as TOPMODEL in which topographic variation has major control in runoff generation; (2) a one to one correspondence between the subbasins and the river network structure, which makes it easy to parameterize runoff routing as opposed to the grid-based approach where the grids often cross over multiple river network reaches [*Tesfa et al.*, 2014]; and (3) a platform to share advances made by the land surface modeling and hydrologic science communities to improve hydrologic simulation in land surface models. Despite these distinct advantages, in the last decade only few land surface modeling efforts have attempted to apply land surface models using subbasin-based approach to improve parameterization of spatial variability of soil water [e.g., *Koster et al.*, 2000] and river transport [e.g., *Goteti et al.*, 2008], which, despite important advances, have several limitations in their routing approach and model inputs [*Tesfa et al.*, 2014]. As another attempt on the use of subbasin-based representation in a land surface model, Tesfa *et al.* [2014] introduced a new subbasin-based approach built upon version 4 of the Community Land Model (CLM4) [*Oleson et al.*, 2010; *Lawrence et al.*, 2011] with some important technical advances including meteorological and land surface inputs derived from high-resolution datasets coupled with a new physically based river routing model [*Li et al.*, 2013]. The Community Land Model version 4 (CLM4) [*Lawrence et al.*, 2011] has a large user community and its use of the TOPMODEL approach for parameterizing runoff may allow it to take more advantages of the subbasin-based land surface representation.

Motivated by the conceptual advantage of the subbasin-based approach for hydrologic modeling and the significance of scalable modeling approaches for providing reliable hydrologic predictions under changing climate and environmental conditions, this study aims to understand the relative merits of the grid- and subbasin-based modeling approaches on hydrologic simulations from scalability perspective. That is, instead of establishing a critical scale for optimal prediction [*Wood et al.*, 1988; *Famiglietti and Wood*, 1994; *Bruneau et al.*, 1995; *Woods et al.*, 1995; *Liang et al.*, 2004; *Shrestha et al.*, 2006], we evaluate the ability of the two modeling approaches to produce robust predictions across multiple spatial scales. For this purpose, we systematically compare the two modeling approaches: (1) to determine if they differ in their scalability in hydrologic flux simulations across multiple spatial resolutions; (2) to explore the sources of their scalability differences; and (3) to determine the significance of their scalability differences. In this paper, we focus more on runoff generation, which is closely related to soil moisture that plays an important role in land-atmosphere interactions in climate models through its controls on surface water and energy fluxes and stream flows that represent freshwater supply. To demonstrate scalability, model predictions made at increasingly higher spatial resolution should asymptotically approach the predictions made at very high spatial resolution. A modeling approach is more scalable if it exhibits the above behavior and model errors are less sensitive to spatial resolution than other approaches. To the best of our knowledge, this is the first attempt to document the relative merits of the grid- and subbasin-based modeling approaches on simulations of hydrologic fluxes from scalability perspective.

The remainder of the paper is organized as follows: Section 2 introduces the configurations of the two land surface modeling approaches. Section 3 describes input data processing. Methods of analysis are presented in Section 4. Results and discussions are presented in Section 5 and finally Section 6 summarizes the results and conclusions drawn from this study.

2. Land surface modeling approaches

In this study, two different land surface spatial representations of CLM4 are applied: the standard grid-based CLM (hereafter denoted as CLM) and the subbasin-based CLM (hereafter denoted as SCLM) described in Tesfa *et al.* [2014]. CLM4 is the latest version of the land component of the Community Earth System Model (CESM) [*Collins et al.*, 2006; *Gent et al.*, 2010; *Lawrence et al.*, 2011], which has been designed and used for studies of interannual and interdecadal variability, paleoclimate regimes, and projections of future changes of the global earth system. Compared to previous versions, CLM4 represents significant improvements in its model parameterizations and structure, including runoff generation, soil hydrology thermodynamics, snow and albedo parameters [*Lawrence et al.*, 2011]. In CLM4, runoff is generated based on the simplified TOPMODEL-based runoff formulation, where both surface and subsurface runoff generations are parameterized as exponential functions of the water table. For details on the runoff parameterizations in CLM4 the reader is referred to Niu *et al.*, [2005], Niu and Yang, [2006], Li *et al.*, [2011] and Huang *et al.*, [2013].

2.1 Grid-based CLM (CLM)

The grid-based CLM refers to the default modeling approach of the standard CLM4 [*Oleson et al.*, 2010] in which the study domain is divided into a number of regular latitude-longitude grid cells. In this study, CLM is set up at 0.125°, 0.25°, 0.5° and 1° spatial resolutions (Table 1 and Figure 1), where the study domain is, respectively, divided into regular latitude-longitude grids in dimensions of 104 x 136, 52 x 68, 26 x 34 and 13 x 17. Hereafter the 0.125°, 0.25°, 0.5° and 1° grid based CLM simulations are denoted as CLM0125, CLM025, CLM05 and CLM1, respectively.

2.2 Subbasin-based CLM (SCLM)

The subbasin-based CLM (SCLM) refers to the new CLM modeling framework introduced by Tesfa *et al.* [2014], where the study domain is divided into a number of irregular subbasins following the natural watershed divides. To be consistent with the grid-based modeling approach, the subbasins are organized into a two dimensional matrix of subbasins, each being treated as a single node for the resolution of interest. For detailed description of SCLM and its applications, interested readers are referred to Tesfa *et al.* [2014], Li *et al.* [2011] and Huang *et al.* [2013]. In this effort, we set up SCLM at four spatial resolutions (Table 1 and Figure 1) with average subbasin size equivalent to the grid-based modeling approach at

0.125°, 0.25°, 0.5° and 1° spatial resolutions, where the study domain is delineated into 5999, 1139, 299 and 75 subbasins that are organized as 77 x 78, 38 x 30, 20 x 15, and 10 x 8 matrices, respectively. At each resolution the extra grid cells of the matrix are masked out as non-land cells and therefore are excluded from the simulation. All the subbasins are delineated using ArcSWAT [*Neitsch et al.*, 2005] from the 90-m Digital Elevation Model (DEM) that was hydrologically conditioned using the 15 arcsec river networks extracted from the Hydrological data and maps based on Shuttle Elevation Derivatives (HydroSHEDS) [*Lehner et al.*, 2008]. To be able to compare objectively against the grid-based approach at equivalent spatial resolutions, the area threshold used for delineation was adjusted iteratively until the average basin size at each resolution is roughly equivalent to the corresponding grid-based average size (Table 1). Similar to the grid-based approach, hereafter the 0.125°, 0.25°, 0.5° and 1° SCLM simulations are denoted as SCLM0125, SCLM025, SCLM05 and SCLM1, respectively.

3. Study domain and input data

3.1 Study Domain

We conjecture that the more complex the topography, the more diverse the hydro-climatologic conditions and the more sensitive the hydrologic simulations are to spatial representations across spatial scales. To highlight potential scalability differences between the grid- and subbasin-based approaches, the Columbia River Basin (CRB) located in the U.S. Pacific Northwest is used as a case study. Figure 1 shows the delineation of the basin using regular latitude/longitude grids and subbasins at four spatial resolutions. The topography of CRB encompasses both mountainous and low-lying regions (Figures 1 and 2). The mountainous regions are characterized by low temperature and higher precipitation dominated by snowfall, while the low-lying regions have higher temperature and lower precipitation mainly in the form of rainfall. The majority of the precipitation in CRB falls during winter with its hydrology dominated by snow accumulation and melting. CRB encompasses the largest river in the Pacific Northwest region of North America, which is the fourth largest river in the United States by discharge volume, and plays a central role in the economy and culture of the region.

3.2 Input data

3.2.1 Land surface and topographic parameters

The land cover and vegetation parameters such as land and plant functional types, leaf area index (LAI), stem area index (SAI), and vegetation canopy top and bottom heights for both CLM and SCLM were extracted from the 0.05° resolution land parameter dataset developed by Ke *et al.*, [2012] based on the Moderate Resolution Imaging Spectroradiometer (MODIS) satellite data. Soil texture was extracted from the 30 arc-second State Soil Geographic Database (STATSGO, now referred to as the U.S. General Soil Map) [*Miller and White*, 1998] for grid cells within the conterminous US and 5-min Food and Agriculture Organization 16-category two-layer soil type data [*Chen et al.*, 2007] for grid cells outside of the conterminous US. Soil color and soil organic matter parameters were obtained from the 0.5° global input data provided with CLM4 due to lack of higher resolution information at the time of this study.

To avoid differences that can arise due to differences in data processing methods, surface parameters for both modeling approaches are generated using the same algorithms. An area dominant algorithm is used to derive soil parameters such as percentage of clay and sand, soil color and organic matter, where parameter value for each modeling unit (subbasin/grid) is assigned to the corresponding parameter value of the source grid cell covering its largest fraction. Figure 1 in the supplementary material shows the spatial distributions of soil types as an example of the land surface parameters at 0.125° resolution and the differences between the coarser resolutions and the finest (0.125°) resolution for both modeling approaches. Land cover and Plant Functional Type (PFT) parameters are generated using an area-weighted averaging algorithm. The algorithm computes parameter value for each modeling unit (PM_n) as the average of the corresponding parameter from all the source grid cells that intersect with the modeling unit (G_{ni}) weighted by the overlapping areas.

$$PM_n = \sum_i G_{ni} A_{ni} \tag{1}$$

where A_{ni} is ratio of the overlapping area between source grid cell *n* and the modeling unit divided by the modeling unit area. LAI, SAI, and top and bottom canopy heights are calculated using the area-weighted averaging algorithm weighted by PFT as:

$$LS_{n,m} = \sum_i A_{wt_i} LS_{nmi} P_{mi} \tag{2}$$

where $LS_{n,m}$ refers to the LAI/SAI of PFT *m* for modeling unit *n*; LS_{nmi} refers to the LAI/SAI of PFT *m* within source grid cell *i* which intersects with modeling unit *n*; P_{mi} is the fraction of PFT *m* within source grid cell *i*. A_{wt_i} is the area fraction weighted by PFT calculated as:

$$A_{wt_i} = \frac{A_{ni}}{\sum_i A_{ni} P_{mi}} \tag{3}$$

For CLM0125, CLM025, CLM05 and CLM1, the surface parameters are created using the CLM4 input preprocessing package [*Oleson et al.*, 2010], which uses the algorithms described above to derive surface parameters at each spatial resolution. For SCLM0125, SCLM025, SCLM05 and SCLM1, first the subbasin boundaries are overlaid on the source grids in ArcGIS to link the subbasins to the grids and calculate fraction of the subbasin area covered by each intersected source grid and area weights needed for the algorithms described above [*Tesfa et al.* 2014]. Then, a set of SCLM input preprocessing tools, consistent with the

CLM4 input preprocessing package, are used to generate land surface parameters at each spatial resolution

In addition to the land surface parameters described above, CLM requires a topographic parameter, F_{max}, which is used to calculate the saturated fraction (F_{sat}) of each modeling unit (grid cells/subbasins), which in turn is used in the calculation of the saturated component of surface runoff [*Oleson et al.*, 2010; *Li et al.*, 2011; *Huang et al.*, 2013] as:

$$SRF = \underbrace{F_{sat}P}_{\text{saturated}} + \underbrace{\left(1 - F_{sat}\right) max\left[0, \left(P - 1\right)\right]}_{\text{unsaturated}} \tag{4}$$

$$F_{sat} = F_{max} exp\left(-C_s f_{over} z\right) \tag{5}$$

where F_{sat} is the fraction of saturated area within a modeling unit, P is the effective rainfall intensity (in mm s^{-1}), which is the sum of liquid throughfall (L) and snowmelt (S), I is the soil infiltration capacity (in mm s^{-1}), F_{max} is the maximum possible saturated area fraction, C_s is a coefficient, f_{over} is a decay factor (m^{-1}) and z is the depth between the ground surface and the water table (in m).

To calculate F_{max}, compound topographic indices (CTIs) are first derived following the definition used in TOPMODEL [*Beven*, 1997; *Quinn et al.*, 1995] in ArcGIS from a 90-m DEM obtained from the HydroSHEDS [*Lehner et al.*, 2008]. We then derived F_{max} following the algorithms described in Niu *et al.* [2005] in the same manner for each spatial resolution of both CLM and SCLM, except that for SCLM the CTIs are clipped following the boundaries of the subbasins in ArcGIS.

3.2.2 Forcing data

The meteorological forcing data (1979-2008) used in this study are extracted from the North America Land Data Assimilation System Phase 2 (NLDAS-2) [*Xia et al.*, 2012]. These include hourly precipitation, air temperature, humidity, surface pressure, wind speed and shortwave and longwave radiations at 0.125° spatial resolution that were derived from the 32-km spatial resolution and 3-hourly temporal resolution North American Regional Reanalysis (NARR). The precipitation in NLDAS-2 was produced using a combination of observations from field stations, level 4 precipitation retrievals from NEXRAD system, and satellites, making it well suited for hydrologic studies. The NLDAS-2 forcing data are used directly for CLM0125, while for CLM025, CLM05 and CLM1 the NLDAS-2 forcing data are spatially aggregated to the corresponding resolution using the area-weighted averaging algorithm (equation 1). For SCLM0125, SCLM025, SCLM05 and SCLM1 the NLDAS-2 forcing data are remapped to the subbasins defined by their boundaries at each spatial resolution using the area-weighted averaging algorithm (equation 1) as described in Tesfa *et al.* [2014]. Figures 2 and 3 (supplementary material) respectively show the spatial distributions of long term mean air temperature and precipitation data at 0.125° resolution and the differences between the coarser resolutions and 0.125° resolution for both modeling approaches. Forcing data are generated using the same methods; thus, as can be seen from the figures, differences between the two modeling approaches at the finest (0.125°) resolution and across spatial resolutions are very small. In both modeling approaches, topography has dominant influence on both air temperature and precipitation in CRB. At the 0.5° and 1° resolutions, air temperature and precipitation can differ quite significantly from the 0.125° data along the Cascades range in the southeast and the northern Rocky range in the northwest and eastern sides of the basin.

4. Methods of analysis

The goal of this study is to investigate the relative merits of CLM and SCLM in hydrologic simulations in terms of their scalability in hydrologic fluxes simulation across multiple spatial resolutions. To compare simulation results from CLM and SCLM at multiple spatial resolutions, all the simulations are driven by the same meteorological forcing (NLDAS-2 1979 – 2008) and land surface parameters, generated using the same methods, and spun up until the state variables (soil moisture, soil temperature and ground water table depth) reached equilibriums. The results discussed in this study are generated by running each model from its spun up state.

By keeping all model parameterizations the same, predictions made at increasingly higher spatial resolutions should asymptotically approach the predictions made at the finest spatial resolution. On the other hand, the finest spatial resolution captures the topographic complexity and hydro-climatologic diversity of the basin better leading to better process representation. Thus, to demonstrate model scalability, in this study, the simulations at the 0.125° resolution (CLM0125 and SCLM0125) are used as the "reference" solution for assessing errors in simulations performed at increasingly coarser resolutions. Scalability of the two modeling approaches in simulated hydrologic fluxes is evaluated by comparing the ability of the coarse resolution simulations (CLM or SCLM) to reproduce the respective reference simulations at 0.125° spatial resolution.

Two types of comparisons are performed: aggregated and disaggregated. In the aggregated comparison, simulation results of CLM0125 (SCLM0125) are aggregated using the area-weighted averaging method (equation 1) to 0.25°, 0.5° and 1° resolution for comparison with CLM (SCLM) simulations performed at 0.25°, 0.5° and 1°, respectively. This comparison is used to assess the scalability of CLM and SCLM based on the ability of the coarser resolution simulations to reproduce the coarser or resolvable scale features upscaled from the corresponding fine (0.125°) resolution simulation and to investigate sources of scalability differences between CLM and SCLM. In the disaggregated comparison, the coarser resolution simulations are disaggregated (mapped) by simple linear spatial interpolation to 0.125° resolution for comparison with the corresponding fine resolution simulation (CLM0125 or SCLM0125). This type of comparison is performed for simulated results and major model input parameters with the main purpose to provide insights on factors that influence scalability of the simulations over various topographic regions rather than assessing model scalability. ArcGIS is used to map the grids/subbasins between different spatial resolutions to facilitate aggregation/disaggregation of input data and simulation results.

In both types of comparisons, by assuming the simulations from the fine scale as the "reference" solution, Mean Absolute Errors (MAEs) are calculated as the absolute difference between the monthly runoff from each coarse resolution simulation and CLM0125 (or SCLM0125) with aggregation or disaggregation. Relative Errors (REs) are calculated for each coarse resolution setup as follows:

$$RE = \left(\frac{RF_f - RF_c}{RF_f} \right) x100 \tag{6}$$

where, RF_f and RF_c are the long term average (e. g., of total runoff) from the fine (0.125°) and coarse resolutions, respectively. The two modeling approaches are compared at two levels: the whole study domain (basin) level using both aggregated and disaggregated comparisons; and the topographic region level using the disaggregated comparison only, as described below.

4.1 Basin level analysis

The basin level analyses are intended to investigate scalability differences between CLM and SCLM (with aggregated comparison only) and to explore sources of scalability differences (with both aggregated and disaggregated comparisons) using statistics from the whole study domain (basin). Scalability in simulated hydrologic fluxes is evaluated by comparing the average MAEs and REs calculated from the 0.25°, 0.5° and 1° resolutions CLM and SCLM for all grid cells or subbasins in the study domain. REs are also calculated using the long term seasonal runoff averaged over the study domain between the 0.125° resolution (CLM0125 and SCLM0125) and the corresponding coarse resolution setups (CLM025, CLM05 CLM1, SCLM025, SCLM05 and SCLM1). The same methods are then applied to runoff fluxes and major model inputs (such as precipitation, snow fall and air temperature) as an effort to identify inputs driving the scalability differences between the two modeling approaches. Finally, a non-parametric statistical significance T-test is performed on the MAE results from the aggregated comparison at each coarse resolution (0.25°, 0.5° and 1°) to evaluate the statistical significance of the differences in scalability between the two modeling approaches.

4.2 Topographic regions level analysis

CRB is characterized by high hydroclimatic diversity that is closely related to its high topographic heterogeneity. Hydrologic fluxes (e. g., runoff) are mainly controlled by the interactions between topographic and climatic properties. The topographic region level analyses are intended mainly to identify parameters causing scalability differences between the two modeling approaches over different topographic regions of the study domain. For this purpose, the study domain is classified into three topographic slope regions using average slope values calculated at 0.125° resolution for both modeling approaches as: gentle gradient (0~10 degrees); moderately steep gradient (10~20 degrees); and steep gradient (greater than 20.0 degrees) regions. Figure 2 shows the topographic regions of CLM and SCLM; while,

Table 2 compares the areas covered by each topographic region in the two modeling approaches. Note that the areas for each topographic region are similar in CLM and SCLM approaches. The average topographic slope values for the subbasins/grids at 0.125° resolution are calculated from a 90-m DEM in ArcGIS as elevation change rate along the steepest downhill descent from each grid cell. Disaggregated comparisons of MAEs and REs are performed over the topographic regions in a similar way as for the basin level analyses described above.

5. Results and Discussion

As discussed earlier, the focus of this study is on the scalability of CLM and SCLM in hydrologic fluxes simulation with greater emphasis on runoff generation. First, CLM and SCLM are evaluated for their scalability in simulated hydrologic fluxes at basin level. Then, sources of their scalability differences in hydrologic fluxes are investigated at basin level and over the topographic regions of the study domain.

5.1 Scalability of CLM and SCLM

In Figure 3, the two modeling approaches are compared for their scalability on hydrologic fluxes using the MAEs of runoff (surface, subsurface and total) and evapotranspiration (ET) fluxes, calculated at the coarse resolutions over the study domain. Results show that SCLM median (notch), average (plus sign) and the maximum (whiskers) MAEs are much smaller than that of CLM at all coarse resolutions for total runoff (Figure 3a), surface runoff (Figure 3b) and subsurface runoff (Figure 3c) showing a clear scalability advantage for SCLM over CLM in runoff generation. The contrast in MAEs increases from fine to coarse resolution, showing reduced sensitivity of SCLM to spatial resolution compared to CLM, although both modeling approaches show asymptotic convergence towards the fine resolution simulations as spatial resolution increases. Unlike runoff, scalability comparison in ET shows smaller

MAEs for CLM at 0.25° and 0.5° resolutions than SCLM but the opposite at 1° resolution (Figure 3d) suggesting that there is no clear scalability difference between CLM and SCLM in ET, which may be more specific to the study region because ET is limited more by energy than water availability. Reasons for the scalability differences between the two approaches at 0.25° and 0.5° resolutions in ET are explored in subsection 5.2.2. Thus, subsequent analyses and discussions focus on runoff generation, although the scalability results may have implications to surface fluxes in more water limited climate regimes.

Shown in Figure 4 are the long term averages of total runoff (left panel) and surface runoff (right panel) simulated by CLM/SCLM at 0.25° (Figure 4 a and d), 0.5° (Figure 4 b and e) and 1° (Figure 4 c and f) resolutions at each modeling grid/subbasin compared against the total and surface runoff simulated by the 0.125° resolution aggregated to the corresponding coarse resolution. Results show larger deviations in the coarse resolution simulations compared to results aggregated from 0.125° resolution in CLM. Hence SCLM is less sensitive to spatial resolution in runoff generation. The scalability advantage of SCLM becomes more pronounced as the spatial resolution coarsens.

Figure 5 compares the scalability of the two modeling approaches using the REs calculated from the long term seasonal averages of total runoff (Figure 5a) and surface runoff (Figure 5b). Long term seasonal runoff of each scale is averaged over the grids/subbasins and used to calculate REs following equation 6 between the coarse resolution simulations and the corresponding one at 0.125° resolution. The results generally show better scalability in SCLM than in CLM for both total runoff (Figure 5a) and surface runoff (Figure 5b). All coarse resolution simulations generate more surface runoff than the corresponding 0.125° resolution simulations (CLM0125/SCLM0125) during winter but the REs from CLM are more negative than SCLM at all coarse resolutions. During spring, surface runoff generated by all the coarse resolutions tend to be lower than the corresponding 0.125° resolution setup but the REs from

CLM are larger than that of SCLM. Hence, overall, SCLM has lower positive biases in runoff during the snow accumulation season (October – March) as well as lower negative biases during the snow melt season (April – July) compared to CLM, making SCLM more scalable at all spatial resolutions.

Comparison of CLM and SCLM for their scalability on surface and total runoff using MAEs averaged over the whole study domain at each coarse resolution (Table 3) generally suggests superior scalability advantage in runoff simulation for SCLM compared to CLM at all coarse resolutions. Also, comparison of the two modeling approaches using the average REs of surface and total runoff fluxes during winter, spring and the full-year separately (Table 4) generally show that SCLM is more scalable than CLM for both surface and total runoff confirming that coarser resolution SCLM simulations can reproduce the fine resolution SCLM0125 simulation much better than that of CLM. However, the improved scalability of SCLM over CLM is more pronounced during winter than averaged over the whole year suggesting that snow processes may play an important role in scalability differences between SCLM and CLM. For more insights on the sources of scalability differences between CLM and SCLM over various regions of the study domain, we discuss the results from basin and topographic region level analyses in the following section.

5.2 Exploring sources of scalability differences

Following the scalability differences between the two modeling approaches discussed in section 5.1, it is logical to ask why SCLM has superior scalability compared to CLM. In this section various potential sources of scalability differences are explored at basin and topographic region levels interchangeably.

5.2.1 Meteorological forcing

To understand why SCLM shows better scalability on runoff generation, we first explored differences in MAEs of major meteorological forcing (precipitation, air temperature and snow) and simulated surface and total runoff of the two modeling approaches by disaggregating the coarse resolution simulations to the corresponding 0.125° resolution at basin level. In general, at basin level, results showed little differences between CLM and SCLM; except a close resemblance of the differences in MAEs of snowfall to the differences in MAEs of total and surface runoff is noted between the two modeling approaches (i.e., slightly reduced MAE in SCLM at 0.5° and 1° resolution compared to CLM) (see supplementary Figures 4 & 5). This motivated us to extend the analysis to topographic regions to take a closer look at the role of meteorological forcing on the CLM and SCLM scalability differences.

Figure 6 shows the differences in MAEs of major meteorological forcing of CLM and SCLM coarse resolution simulations disaggregated to the corresponding 0.125° resolution over the topographic regions of the study domain. For comparison, the differences in MAEs of surface and total runoff of the coarse resolution simulations over the same topographic regions are shown in Figure 7. Note that the disaggregation approach is used for comparison because the topographic classification is more meaningfully defined based on elevation data at 0.125° resolution. In the gentle gradient region, similar to the total and surface runoff (Figures 7 a and d), differences in MAEs of precipitation (Figure 6a), air temperature (Figure 6b) and snowfall (Figure 6c) show slightly lower MAEs for CLM than SCLM. However, this region receives relatively little precipitation and thus contributes less to basin scale runoff compared to the moderately steep gradient and steep gradient regions (see Supplementary Figure 3). In the moderately steep and steep regions, differences in MAEs of precipitation between CLM and SCLM (Figure 6 d and g) is found to be similar to the gentle gradient region (Figure 6a), which are quite different from the differences shown in total and surface runoff MAEs (Figure 7 b, c, e, and f). On the other hand, a clear resemblance is noted for the differences in MAEs of snowfall (Figure 6 h and i) and air temperature (Figure 6 e and f) to the differences in MAEs of the total and surface runoff (Figure 7 b, c, e and f) in both moderately steep and steep regions of the study domain suggesting snow and air temperature as potential sources of the scalability differences between the two modeling approaches.

In Figures 6 and 7, it is important to note that MAE comparison at the 1° resolution over the steep gradient region is problematic because, although the average size of a subbasin over the study domain in SCLM1 is equivalent to the average size of a grid in CLM1, majority of the steep gradient region located in the northern area (Figure 2) is represented by a single subbasin in SCLM1 (Figure 1). This makes the MAE comparison at 1° resolution less consistent than that of 0.25° and 0.5° resolutions because many of the subbasins at 1° resolution are too coarse to represent areas with the complex topography in CRB. However, generally, the results over the moderately steep gradient and steep gradient regions suggest some advantages for SCLM for snowfall similar to the advantages in runoff simulation shown in Figure 7, which is consistent to the advantages shown in air temperature. This suggests that snowfall has contributed to the scalability differences in runoff generation (Figure 3). The slightly improved bias in air temperature for SCLM over the moderately steep gradient and steep gradient topographic regions is a consequence of improved consistency in elevation for SCLM. Plots of MAE in elevation show similar differences between CLM and SCLM across different spatial resolutions comparable to that of air temperature over the moderately steep gradient and steep gradient regions (Figure not shown).

Furthermore, consistent with Figure 7 (a and d), the differences in long term seasonal total runoff and surface runoff flux simulations at the 0.125° resolution averaged over the three topographic regions (Figure 8) generally shows CLM slightly lower bias for CLM over the gentle gradient region in both total runoff (Figure 8a) and surface runoff (Figure 8b). Over the moderately steep gradient region, SCLM results in lower bias on total runoff (Figure 8c), although CLM seem to have generally lower bias during winter season. For surface runoff (Figure 8d), again SCLM results in lower biases during spring, although its advantage during winter is less discernible. Over the steep gradient region, SCLM results in lower bias in both total (Figure 8e) and surface (Figure 8f) runoff simulations. In general, Figures 6, 7 and 8 show that the differences in bias between the two modeling approaches are more pronounced over the mountainous regions of the study domain suggesting potential impacts of snow on the scalability differences.

To discern more clearly the role of snowfall on CLM and SCLM scalability differences, we now return to analysis at basin level using aggregated comparison, where snowfall MAEs are calculated at each coarse resolution. The results in Figure 9 show a clear scalability advantage in snowfall and seasonal snowmelt for SCLM compared to CLM. Note that the runoff generation over the cold regions in CRB (moderately steep gradient and steep gradient regions in this study) is closely related to snowfall [Bowling *et al.* 2004]. Thus, at basin scale, the scalability differences in runoff generation shown in Figures 3 and 5 seem to be explained by the scalability differences in snowfall and seasonal snowmelt shown in Figure 9. Figure 10 shows the spatial distribution of long term averaged surface runoff simulated by CLM0125 (Figure 10a) and SCLM0125 (Figure 10b) and surface runoff MAEs of CLM (Figure 10 c, e and g) and SCLM (Figure 10 d, f, and h) at 0.25°, 0.5° and 1° resolutions. Hence, despite almost no discernible difference in surface runoff spatial distribution between CLM0125 and SCLM0125, the results show a clear difference in the MAEs of surface runoff between the two modeling approaches at all coarse resolutions. Differences are more pronounced over the moderately steep and steep gradient topographic regions (see Figure 2) where snowmelt runoff is important, resulting in the scalability advantage of SCLM over those regions.

In these analyses, it is important to note the difference between the two types of comparisons (aggregated vs. disaggregated). In the disaggregated comparison, MAEs are calculated at the fine scale that is not resolvable by the coarse resolution simulations resulting in less clear scalability differences between CLM and SCLM. On the other hand, the aggregated comparison, by calculating MAEs based on spatial features resolvable by the coarse resolution simulations rather than evaluating fine scale features not resolvable by the coarse simulations, allows us to see more clearly the scalability difference between the two modeling approaches (see Figures 3 & 9).

5.2.2 Land surface parameters

To further understand the sources of the scalability differences on ET as shown in Figure 3d, basin level land surface parameters are analyzed. We found that, due to lack of high resolution soil dataset, the area dominant algorithm used to generate soil color parameters resulted in exactly the same parameter values at 0.125°, 0.25°, and 0.5° resolutions (Figure not shown) and vary only at 1° for CLM; while SCLM soil color values vary across all resolutions. This resulted in less ET MAEs for CLM025 and CLM05 as compared to the corresponding SCLM simulations. At 1° resolution, SCLM results in smaller MAEs than CLM suggesting some scalability advantage on ET flux simulation (Figure 3d). However, overall, ET showed minimal sensitivity to differences in spatial resolution compared to runoff because of limited resolution soil color dataset combined with ET being more limited by energy rather than water availability in CRB. Thus, improved scalability in runoff generation is not reflected clearly in improved scalability in surface energy fluxes.

Another important potential source of scalability advantage for SCLM over CLM beside snowfall may be its advantage in the calculation of topographic indices, which are important parameters in the SIMTOP surface runoff generation scheme [*Niu et al.*, 2005*; Oleson et al.*, 2010]. In CLM4, surface runoff (*SRF*) is composed of runoff generated at the saturated fraction (saturation excess) (see equations 1 & 2) and unsaturated fraction (infiltration excess) of each modeling unit (subbasin or grid) [*Li et al.*, 2011].

For further insight into the role of topography on scalability differences between SCLM and CLM, we examined the two modeling approaches for their scalability differences on the (1) surface runoff generated by the saturated fraction of the modeling units (grids/subbasins) (i.e., $F_{sat}P$, Figure 11a), (2) saturated surface runoff driven by snowmelt (i.e., $F_{sat}S$, Figure 11b), (3) saturated surface runoff driven by rain (i.e., $F_{sat}L$, Figure 11c), (4) seasonal saturated fraction of the modeling units (grids/subbasins), F_{sat} (Figure 11d) and (5) MAEs of F_{sat} from aggregated comparison (Figure 11e). The results show a clear scalability advantage in the surface runoff component generated by the saturated fraction of the grid/subbasin as well as the snowmelt driven component of the saturated surface runoff (consistent with the results on snowmelt scalability shown in Figure 9b) for SCLM than CLM. SCLM also showed generally lower bias in the rain driven component of saturated surface runoff consistently in winter when rain-driven runoff is important. From the resemblance between Figures 11c and 11d, it is clear that the scalability difference between the two modeling approaches in the rain driven saturated runoff component is driven by the seasonal F_{sat}, which depends on the topographic parameter F_{max}. The scalability difference in F_{sat} is shown clearly from the MAEs of F_{sat} (Figure 11e). Here, it is important to note that the F_{sat} comparison (Figure 11e) at 1° resolution is less consistent than at 0.25° and 0.5° resolutions because, despite the average subbasin size being equivalent to 1°, many subbasins in SCLM1 are too large to capture the topographic complexity of CRB, which is reflected more in Figures 11c and 11d.

Investigation of the differences in surface runoff fraction across the coarse resolutions showed SCLM to be more consistent in partitioning of the total runoff between the surface and subsurface components (figure not shown), which is consistent with the lower bias of SCLM in F_{sat}. Further investigation of F_{max}, a parameter used in the calculation of F_{sat}, showed (figure not shown) that the basin level average F_{max} values increased as the spatial resolution coarsens in both modeling approaches resulting in larger saturated fraction in the coarser spatial resolutions than the corresponding 0.125° resolutions (Figure 11d). SCLM has slightly larger basin level average F_{max} values than CLM at all spatial resolutions, with slightly better consistency in its differences across spatial resolutions resulting in overall better consistency in the partitioning of runoff between surface and subsurface components and F_{sat}.

5.3 Significance of scalability differences

Following the results discussed so far, it is logical to ask whether the scalability differences between the two modeling approaches have any statistical significance. Shown in table 5 are p-values from a non-parametric statistical significance T-test on MAEs calculated at each coarse resolution by comparing the coarse resolution simulations with the corresponding fine resolution simulation (CLM0125 or SCLM0125) aggregated to each coarse resolution for runoff generation (total, surface and subsurface runoff), snow, evapotranspiration (ET) and F_{sat}. Using a confidence level of 95%, results show: (1) significant differences in MAEs between CLM and SCLM in all runoff components and snow at all coarse resolution (0.25°, 0.5°, and 1°); (2) significant difference in F_{sat} at 0.25° and 0.5° resolutions; (3) insignificant difference in F_{sat} at 1° resolution; and (4) insignificant difference in ET at all coarse resolutions. The results on runoff and snow suggest that SCLM has significant advantage in scalability on runoff generation, which is mainly driven by its superior scalability on snow simulation across all spatial resolutions. In addition, the results on F_{sat} suggest that the advantage of SCLM on F_{max} (topographic parameter) representation has contributed to the improved scalability on surface runoff simulation. The result on ET shows its low sensitivity to differences in spatial representation. Overall, these results show the importance of scalability advantage of SCLM compared to CLM modeling approach suggesting potential implications to future development of land surface models.

6. Summary and Conclusions

Scalability of land surface models is vital for multi-scale modeling of the terrestrial water cycle and land-atmosphere interactions. This study evaluates the relative merits of the grid-based and the subbasin-based spatial representations for land surface modeling using CLM4 as an example to investigate the scalability of CLM in hydrologic fluxes with greater focus on runoff generation. The two land surface modeling approaches are applied over the Columbia River Basin at four spatial resolutions (0.125°, 0.25°, 0.5°, and 1°), driven by the same meteorological and land surface input parameters mostly available at 0.125° resolution or higher. Using the simulations at 0.125° resolution as the "reference" solution, mean absolute and relative errors are calculated for the simulations at 0.25°, 0.5° and 1° spatial resolutions using aggregated and disaggregated comparisons at the whole basin and topographic regime levels. These metrics are compared between the grid-based (CLM) and subbasin-based (SCLM) simulations to evaluate their scalability in runoff generation. A non-parametric statistical significance T-test is employed to assess significance of the differences in scalability between the two modeling approaches at each coarse resolution. Major meteorological and land surface parameters of runoff generation are investigated to identify the sources of scalability differences between the two modeling approaches.

Results from the aggregated comparison suggest significant scalability advantage in runoff generation for SCLM compared to CLM. The aggregated comparison, by calculating MAEs based on spatial features resolvable by the coarse resolution simulations rather than evaluating fine scale features not resolvable by the coarse simulations, allows us to see more clearly the scalability difference between the two modeling approaches. Basin level annual average relative errors of surface runoff at 0.25°, 0.5°, and 1° resolutions are 3.02%, 4.32%, and 5.92% for SCLM and 3.70%, 7.36%, and 10.60% for CLM, respectively. Basin level annual average relative errors of total runoff at 0.25°, 0.5°, and 1° resolutions are 2.66%, 4.46%, and 6.13% for SCLM and 3.35%, 6.28%, and 9.36% for CLM, respectively. Topographic regime level analysis showed that scalability differences between SCLM and CLM are more pronounced over mountainous regions, especially during winter. As runoff generation is dominated by orographic precipitation and snow accumulation and melt in the mountains, SCLM showed overall improvements in scalability in runoff flux compared to CLM at the basin level. Overall, ET showed minimal sensitivity to differences in spatial resolution compared to runoff, which may be attributed to CRB being energy limited.

Investigation on the sources of the scalability differences between the two modeling approaches based on major meteorological and land surface parameters of runoff generation showed that SCLM has superior scalability in snowfall/rainfall partitioning, which is related to lower bias in air temperature over mountainous regions in SCLM that in turn is related to improved consistency of surface elevation in a subbasin-based representation over a grid-based representation. Further investigation of the scalability of SCLM and CLM on surface runoff generation indicates that SCLM has better scalability in both snowmelt and rain driven saturated components of surface runoff. In addition, scalability in the rain driven saturated surface runoff component is clearly related to F_{sat}, which depends on the topographic indices. However, the overall scalability advantages of SCLM are dominated by snowfall and snowmelt that disproportionately drive runoff processes in CRB; but rain driven runoff still plays a role in the winter.

Finally, statistical significance test results at 95% confidence level have shown that the scalability differences between SCLM and CLM on all runoff components and snow are statistically significant at all coarse resolutions (0.25°, 0.5°, and 1°). Differences in F_{sat} are significant only at 0.25° and 0.5° resolutions; while differences in ET are not significant at all coarse resolutions. The results on runoff and snow have shown superior scalability advantage for SCLM on runoff generation compared to CLM suggesting potential implications for future development of land surface models.

Despite the fact that the hydrologic formulations and parameters that govern runoff generation in CLM were originally calibrated for the grid-based modeling approach, it is interesting that SCLM demonstrates superior scalability to the grid-based CLM without changes to the existing parameterizations or model parameters. Hence this study demonstrates the importance of spatial structure for multi-scale modeling of hydrological processes. The subbasin-based modeling approach not only provides a logical way of modeling soil moisture heterogeneity and runoff but it can easily be extended to include modeling of river routing [e.g., *Tesfa et al.* 2014] and water management [*Voisin et al.*, 2013].

While it is possible that SCLM may exhibit improved scalability in river basins with hydrology dominated by complex terrain and snow processes similar to CRB, it is not clear if model structure and/or other factors (e.g., formulation of the runoff parameterizations) may contribute to scalability in rain-dominated river basins with lower topographic relief. The results from our systematic analyses motivate future research to further compare SCLM and CLM over different climatic and topographic regimes. Similar analyses should be extended to higher spatial resolution (e.g., including 1/16° when input data are available) and higher temporal resolution (e.g., investigate scalability of daily runoff). In addition to analyses using an idealized experimental framework where the high resolution simulations are used as the reference to evaluate the coarser resolution simulations, it would be useful to evaluate the simulations using observation data such as streamflow measurements from MOPEX basins to advance understanding of the relative merits of the grid-based and subbasin-based approach for hydrologic simulations.

In summary, this study demonstrated that by adopting a spatial structure that takes advantage of the spatial characteristics of meteorological and hydrologic processes, improved scalability could be achieved in simulating runoff generation, which has implications to simulating surface energy fluxes in coupled land-atmosphere models. The improved scalability reduces sensitivity of the simulations to spatial resolution and the needs for extensive parameter calibration when the models are applied at different spatial resolutions. We recognize, however, that parameter calibration is important in light of the myriad sources of uncertainties (e.g., associated with the forcing and surface parameters) so it would be useful to compare scalability differences between the grid- and subbasin-based approaches when parameter calibration is included in future research. Lastly, we emphasize that the subbasin-based approach improves scalability without adding computational cost compared to the grid-based approach. This is in contrast to other approaches such as introduction of subgrid elevation bands, as adopted in the Variable Infiltration Capacity (VIC) model [*Liang et al.*, 1994] and the subgrid parameterization of orographic precipitation of Leung and Ghan [1995; 1998], which also improves spatial structure that may lead to improved scalability, but at the expense of increased computational demand. Comparison of these approaches should also yield important insights on cost effective scalable approaches to multi-scale modeling of hydrologic processes.

Acknowledgements

This research is part of the Platform for Regional Integrated Modeling and Analysis (PRIMA) Initiative at Pacific Northwest National Laboratory (PNNL). It was conducted under the Laboratory Directed Research and Development Program at PNNL, a multi-program national laboratory operated by Battelle for the U.S. Department of Energy under Contract DE-AC05-76RL01830. Some analyses of the study were supported by the Office of Science of the U.S. Department of Energy as part of the Earth System Modeling and Regional and Global Climate Modeling programs.

Reference:

Arnold, J. G., R. Srinivasan, R. S. Muttiah, and J. R. Williams (1998), LARGE AREA HYDROLOGIC MODELING AND ASSESSMENT PART I: MODEL DEVELOPMENT1, *JAWRA Journal of the American Water Resources Association*, *34*(1), 73-89, doi:10.1111/j.1752-1688.1998.tb05961.x.

Beven, K. (1997), Topmodel: A Critique, *Hydrol. Processes*, *11*(9), 1069-1085.

Bicknell, B. R., J. C. Imhoff, J. L. Kittle Jr., A. S. Donigian, and R. C. Johanson (1997), *Hydrological Simulation Program—FORTRAN. User's Manual for Release 11*, Athens, GA.

Bruneau, P., C. Gascuel-Odoux, P. Robin, P. Merot, and K. Beven (1995), Sensitivity to Space and Time Resolution of a Hydrological Model Using Digital elevation Data, *Hydrol. Processes*, *9*, 69-81.

Cerdan, O., Y. Le Bissonnais, G. Govers, V. Lecomte, K. van Oost, A. Couturier, C. King, and N. Dubreuil (2004), Scale effect on runoff from experimental plots to catchments in agricultural areas in Normandy, *J. Hydrol.*, *299*(1–2), 4-14, doi:http://dx.doi.org/10.1016/j.jhydrol.2004.02.017.

Chen, F., K. Manning, M.A. LeMone, S.B. Trier, J.G. Alfieri, R. Roberts, M. Tewari, D. Niyogi, T. Horst, S. P. Oncley, Description and evaluation of the characteristics of the NCAR high-resolution land data assimilation system, *J. Appl. Meteorol. Clim.*, 46(6), 694-713.

Chen, F., K. Mitchell, J. Schaake, Y. Xue, H.-L. Pan, V. Koren, Q. Y. Duan, M. Ek, and A. Betts (1996), Modeling of Land Surface Evaporation by Four Schemes and Comparison with FIFE Observations, *Journal of Geophysical Research*, *101*(D3), 7251-7268.

Chen, J., and P. Kumar (2001), Topographic Influence on the Seasonal and Interannual Variation of Water and Energy Balance of Basins in North America, *Journal of Climate*, *14*(9), 1989-2014, doi:10.1175/1520-0442(2001)014<1989:TIOTSA>2.0.CO;2.

Collins, W. D., P. J. Rasch, B. A. Boville, J. J. Hack, J. R. McCaa, D. L. Williamson, B. P. Briegleb, C. M. Bitz, S.-J. Lin, and M. Zhang (2006), The Formulation and Atmospheric Simulation of the Community Atmosphere Model Version 3 (CAM3), *Journal of Climate*, *19*(11), 2144-2161, doi:10.1175/JCLI3760.1.

Famiglietti, J. S., and E. F. Wood (1994), Mutiscale modeling of spatially variable water and energy balances, *Water Resour. Res.*, *30*(11), 3061-3078.

Gent, P., S. Yeager, R. Neale, S. Levis, and D. Bailey (2010), Improvements in a half degree atmosphere/land version of the CCSM, *Clim Dyn*, *34*(6), 819-833, doi:10.1007/s00382-009-0614-8.

Goteti, G., J. S. Famiglietti, and K. Asante (2008), A Catchment-Based Hydrologic and Routing Modeling System with explicit river channels, *Journal of Geophysical Research: Atmospheres*, *113*(D14), D14116, doi:10.1029/2007JD009691.

Huang, M., Z. Hou, L. R. Leung, Y. Ke, Y. Liu, Z. Fang, and Y. Sun (2013), Uncertainty Analysis of Runoff Simulations and Parameter Identifiability in the Community Land Model – Evidence from MOPEX Basins, *J. Hydrometeor.*, doi:10.1175/JHM-D-12-0138.1.

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

Koster, R. D., M. J. Suarez, A. Ducharne, M. Stieglitz, and P. Kumar (2000), A catchment-based approach to modeling land surface processes in a general circulation model: 1. Model structure, *J. Geophys. Res.*, *105*(D20), 24809-24822, doi:10.1029/2000JD900327.

Lawrence, D. M., et al. (2011), Parameterization improvements and functional and structural advances in Version 4 of the Community Land Model, *Journal of Advances in Modeling Earth Systems*, *3*(1), M03001, doi:10.1029/2011MS00045.

Lehner, B., K. Verdin, and A. Jarvis (2008), New Global Hydrography Derived From Spaceborne Elevation Data, *Eos, Transactions American Geophysical Union*, *89*(10), 93-94, doi:10.1029/2008EO100001.

Leung, L. R., and S. J. Ghan (1995), A subgrid parameterization of orographic precipitation, *Theoretical and Applied Climatology*, *52*(1-2), 95-118, doi:10.1007/BF00865510.

Leung, L. R., and S. J. Ghan (1998), Parameterizing Subgrid Orographic Precipitation and Surface Cover in Climate Models, *Mon. Weather Rev.*, *126*, 3271-3291.

Tesfa, T. K., H. Li, L.-Y. R. Leung, M. Huang, Y. Ke, Y. Sun, Y. Liu, and (2014), A Subbasin-based framework to represent land surface processes in an Earth System Model., *Geosci. Model Dev. Discuss.*, 6, 2699-2730, doi:10.5194/gmdd-6-2699-2013.

Li, H., M. Huang, M. S. Wigmosta, Y. Ke, A. M. Coleman, L. R. Leung, A. Wang, and D. M. Ricciuto (2011), Evaluating runoff simulations from the Community Land Model 4.0 using observations from flux towers and a mountainous watershed, *J. Geophys. Res.*, *116*(D24), D24120, doi:10.1029/2011JD016276.

Li, H., M. S. Wigmosta, H. Wu, M. Huang, Y. Ke, A. M. Coleman, and L. R. Leung (2013), A physically based runoff routing model for land surface and earth system models, *Journal of Hydrometeorology*, doi:10.1175/JHM-D-12-015.1.

Liang, X., J. Guo, and L. R. Leung (2004), Assessment of the effects of spatial resolutions on daily water flux simulations, *J. Hydrol.*, *298*(1–4), 287-310, doi:http://dx.doi.org/10.1016/j.jhydrol.2003.07.007.

Liang, X., D. P. Lettenmaier, E. F. Wood, and S. J. Burges (1994), A Simple Hydrologically Based Model of Land Surface Water and Energy Fluxes for General Circulation Models, *J. Geophys. Res.*, *99*(D7), 14415-14428.

Miller, D. A., and R. A. White (1998), A Conterminous United States Multilayer Soil Characteristics Dataset for Regional Climate and Hydrology Modeling, *Earth Interactions*, 2(2), 1-26.

Neitsch, S. L., J. G. Arnold, J. R. Kiniry, R. Srinivasan, and J. R. Williams (2005), Soil and

Water Assessment Tool Theoretical Documentation, version 2005, edited, Temple, TX: Grassland, Soil and Water Research Laboratory, Agricultural Research Service.

Niu, G.-Y., and Z.-L. Yang (2006), Effects of Frozen Soil on Snowmelt Runoff and Soil Water Storage at a Continental Scale, *J. Hydrometeor.*, *7*(5), 937-952, doi:10.1175/JHM538.1.

Niu, G.-Y., Z.-L. Yang, R. E. Dickinson, and L. E. Gulden (2005), A simple TOPMODEL-based runoff parameterization (SIMTOP) for use in global climate models, *J. Geophys. Res.*, *110*(D21), D21106, doi:10.1029/2005JD006111.

Oleson, K. W., D. M. Lawrence, G. B. Bonan, M. G. Flanner, E. Kluzek, P. J. Lawrence, S. Levis, S. C. Swenson, and P. E. Thornton (2010), Technical Description of version 4.0 of the Community Land Model (CLM), NCAR/TN-478+STR, National Center for Atmospheric Research, Boulder.

Quinn, P., K. Beven, and A. Culf (1995), The Introduction of Macroscale Hydrological Complexity into Land Surface-Atmosphere Transfer ;Models and the Effect on Planetary Boundary Layer Development, *J. Hydrol.*, *166*(1995), 421-444.

Shrestha, R., Y. Tachikawa, and K. Takara (2006), Input data resolution analysis for distributed hydrological modeling, *J. Hydrol.*, *319*(1-4), 36-50.

Sridhar, V., R. L. Elliott, and F. Chen (2003), Scaling effects on modeled surface energy-balance components using the NOAH-OSU land surface model, *J. Hydrol.*, *280*(1–4), 105-123, doi:http://dx.doi.org/10.1016/S0022-1694(03)00220-8.

Voisin, N., L. Liu, M. Hejazi, T. K. Tesfa, L. H., M. Huang, Y. Liu, and L. R. Leung (2013), One-way coupling of an integrated assessment model and a water resources model:

evaluation and implications of future changes over the US Midwest, *Hydrol. Earth Syst. Sci.*, *10*, 6359-6406, doi:doi:10.5194/hessd-10-6359-2013.

Wood, E. F., M. Sivapalan, K. Beven, and L. Band (1988), Effects of spatial variability and scale with implications to hydrologic modeling, *J. Hydrol.*, *102*, 29-47.

Woods, R., M. Sivapalan, and M. Duncan (1995), Investigating the Representative Elementary Area Concept - an Approach Based on Field Data, *Hydrol. Processes*, *9*(3-4), 291-312.

Xia, Y., M. Ek, H. Wei, and J. Meng (2012), Comparative analysis of relationships between NLDAS-2 forcings and model outputs, *Hydrol. Processes*, *26*(3), 467-474, doi:10.1002/hyp.8240.

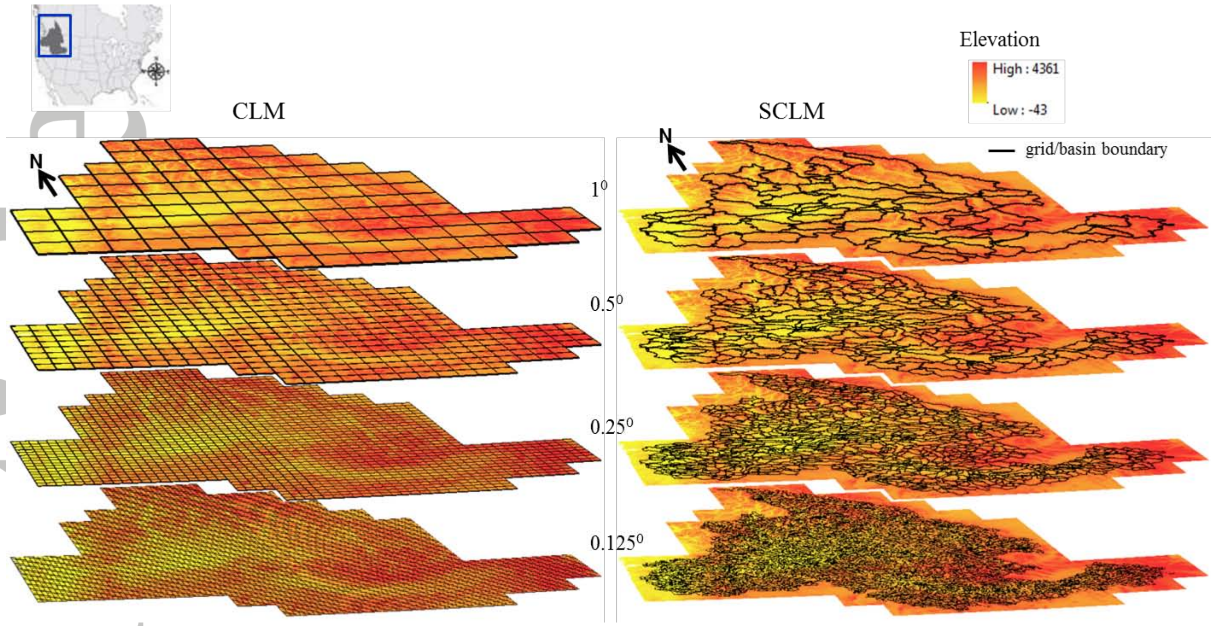

Figure 1: The Columbia River Basin (CRB) location and perspective view of grid- and subbasin-based boundaries.

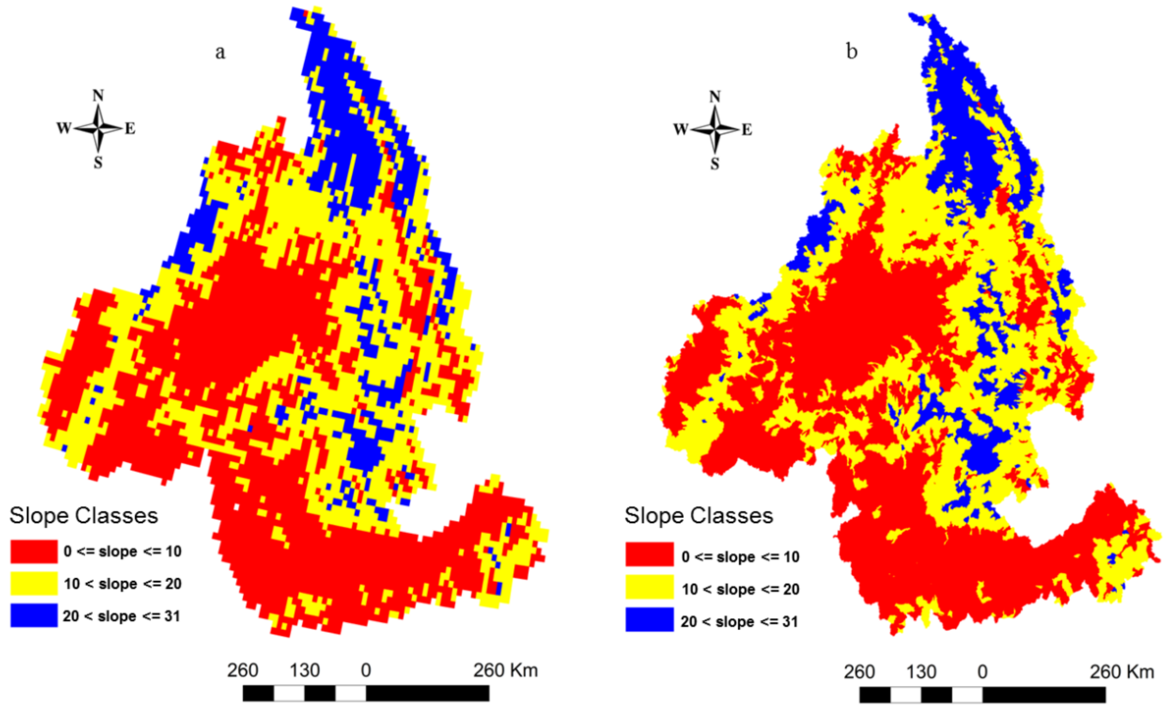

Figure 2: Gentle gradient (blue), moderately steep gradient (yellow) and steep gradient (red) topographic regions for the 0.125° resolution CLM (a) and SCLM (b) classified based on average slope in degrees.

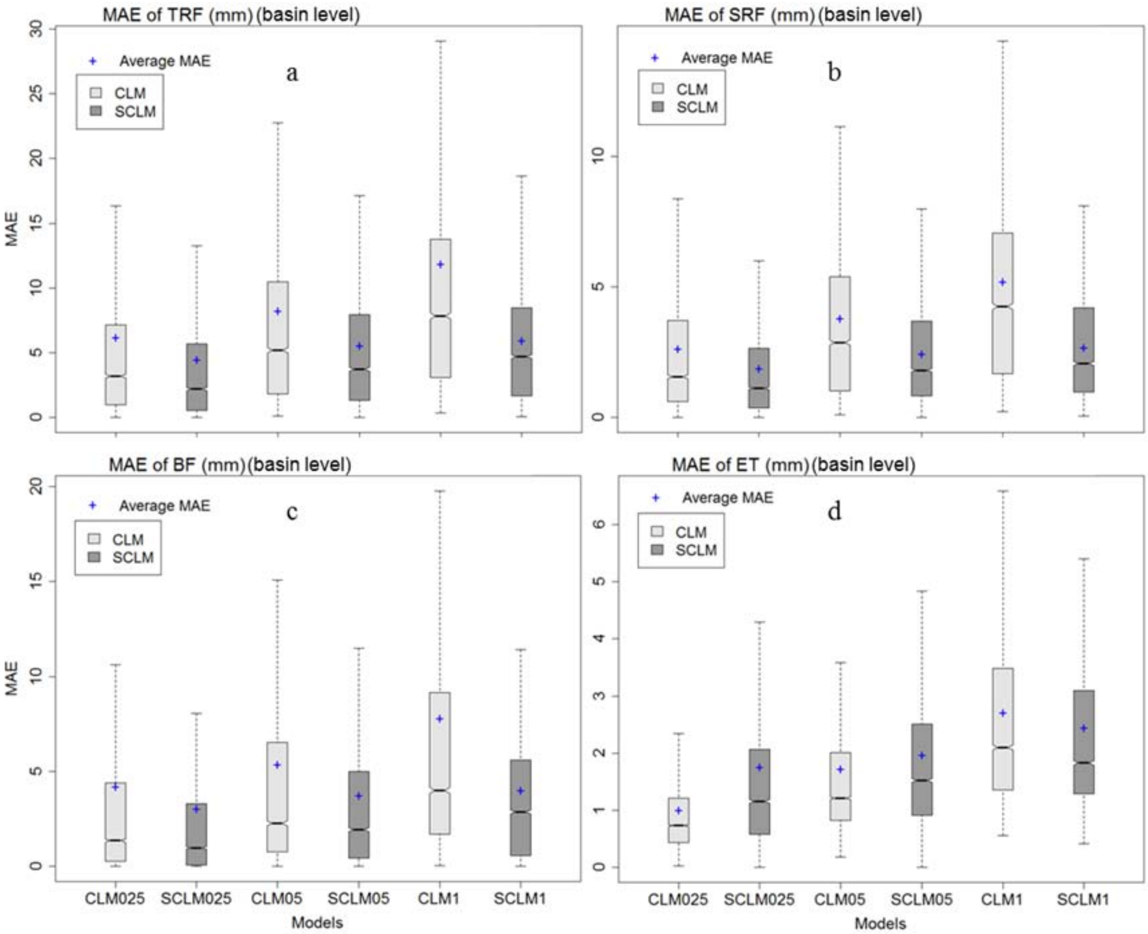

Figure 3: Scalability of CLM and SCLM total runoff (a), surface runoff (b), subsurface runoff (c) and evapotranspiration (d) based on mean absolute error (MAE) of the coarser scales in comparison to CLM0125 and SCLM0125 aggregated to each coarse scale. On each box, the central mark (notch) is the median (q_2), the edges of the boxplot are the 25th (q_1) and 75th (q_3) percentiles, and the whiskers extend to the most extreme data points (q_3 + 1.5 x interquartile range ($q_3 - q_1$) and q_1 – 1.5 x interquartile range ($q_3 - q_1$)); outliers are not considered.

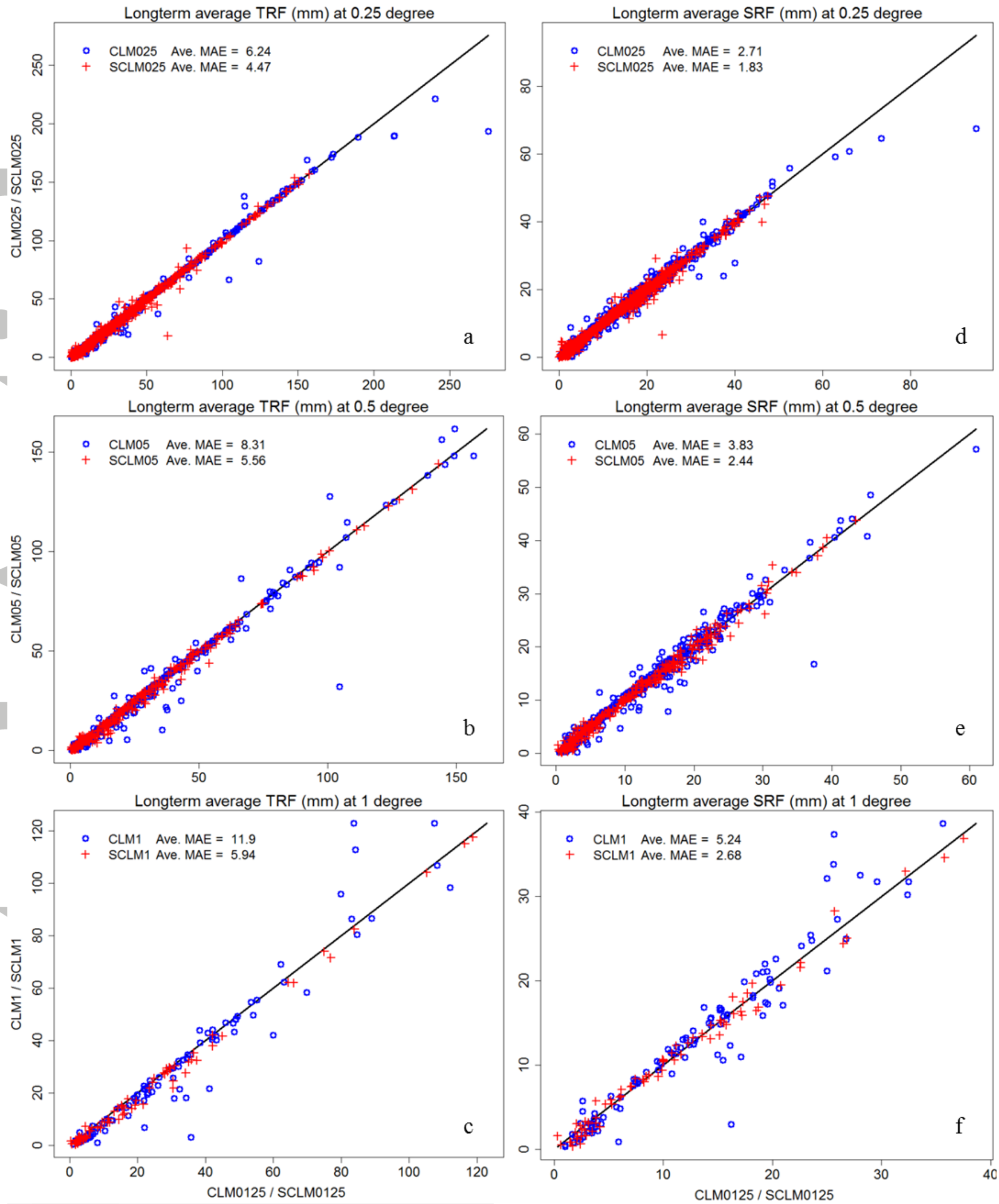

Figure 4: Comparison of long term average total runoff (left panel) and surface runoff (right panel), calculated at each modeling unit (grid/subbasin), generated by the 0.25° (a and d), 0.5° (b and e) and 1° (c and f) resolutions against long term average total and surface runoff generated by the corresponding 0.125° resolution aggregated to each coarse resolution.

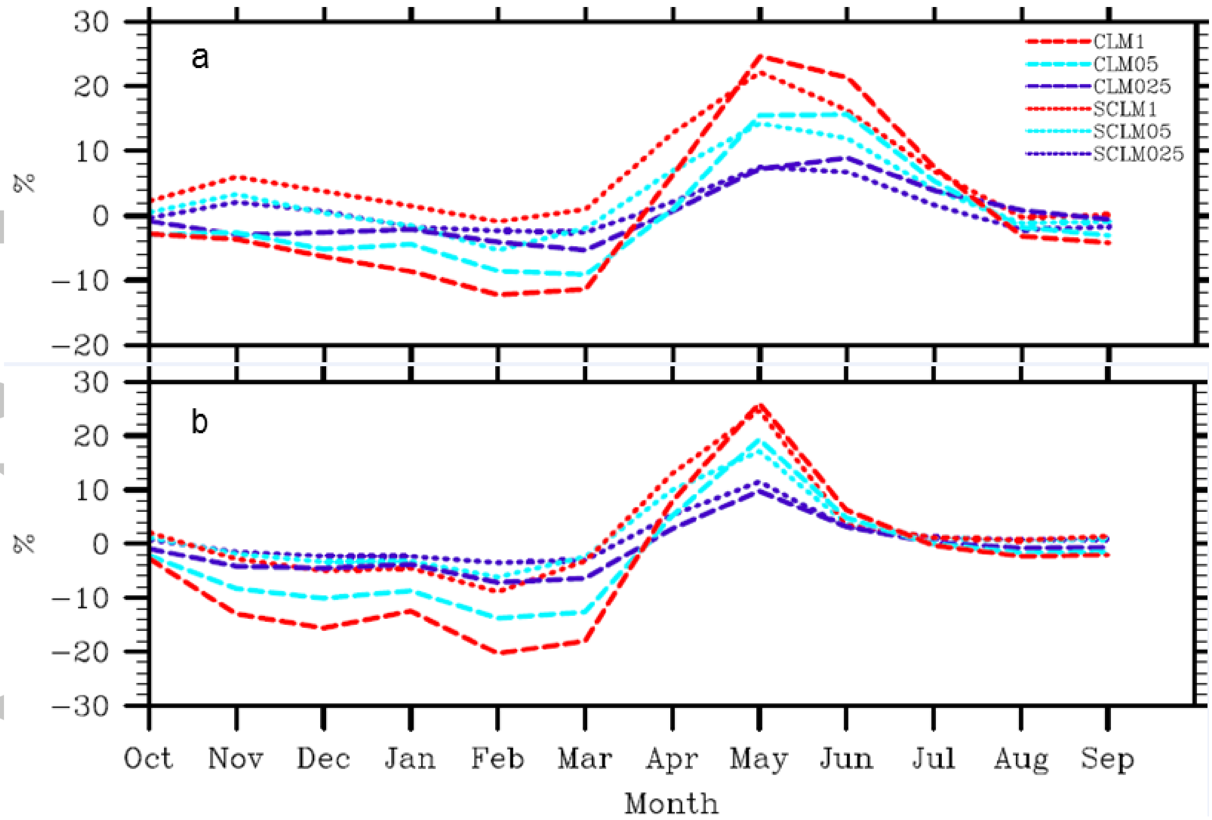

Figure 5: Relative Error (RE) in long term monthly mean total runoff (a) and surface runoff (b) simulated by CLM025, CLM05 and CLM1 as compared to results from CLM0125; and SCLM025, SCLM05 and SCLM1 as compared to results from SCLM0125 over the whole CRB.

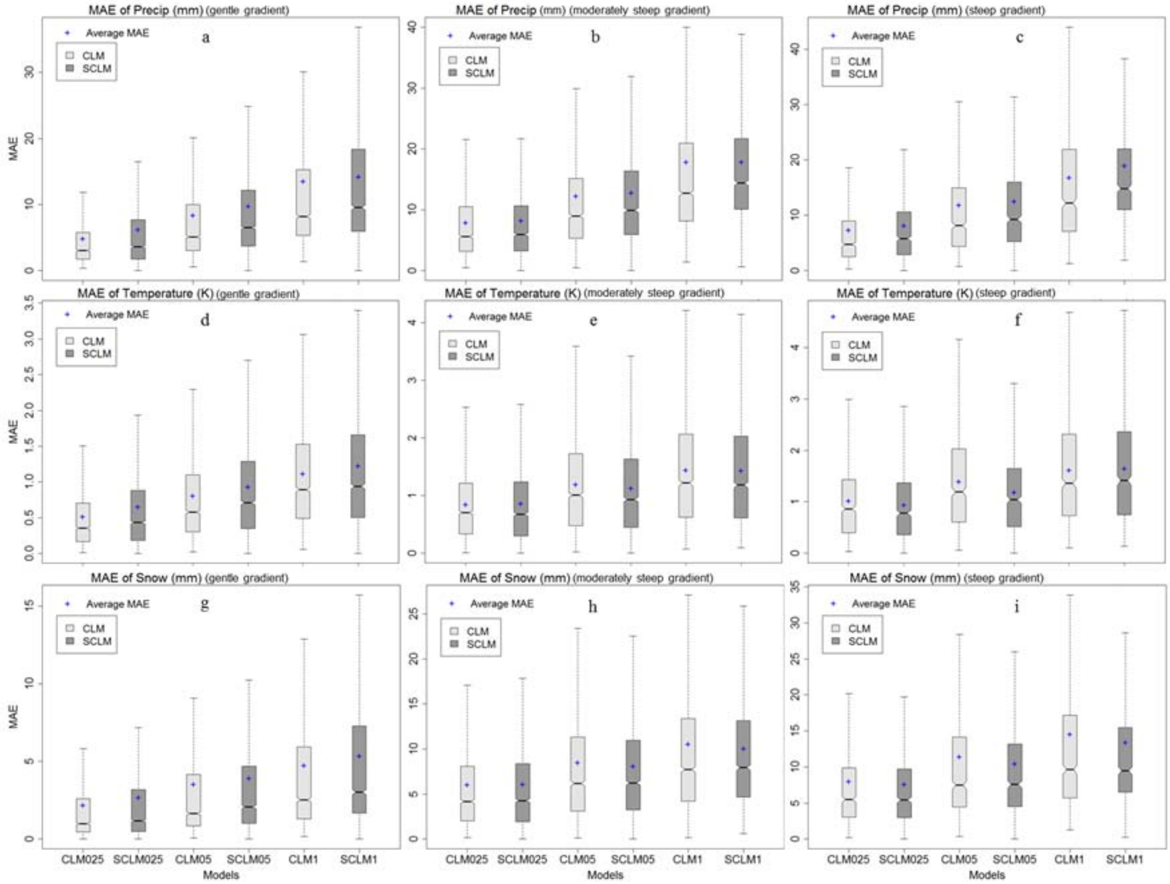

Figure 6: Differences in MAEs between CLM and SCLM on precipitation (top panel), air temperature (middle panel) and atmospheric snow (bottom panel) based on MAEs calaculated by comparing the coarser scales (disaggregated to 0.125^{o} resolution) to CLM0125 and SCLM0125 over gentle gradient (a, d, and g), moderately steep gradient (b, e and h) and steep gradient (c, f and i) topographic regions. On each boxplot, the central mark (notch) is the median (q_2), the edges of the boxplot are the 25th (q_1) and 75th (q_3) percentiles, and the whiskers extend to the most extreme data points (q_3 + 1.5 x interquartile range ($q_3 - q_1$) and q_1 – 1.5 x interquartile range ($q_3 - q_1$)); outliers are not considered.

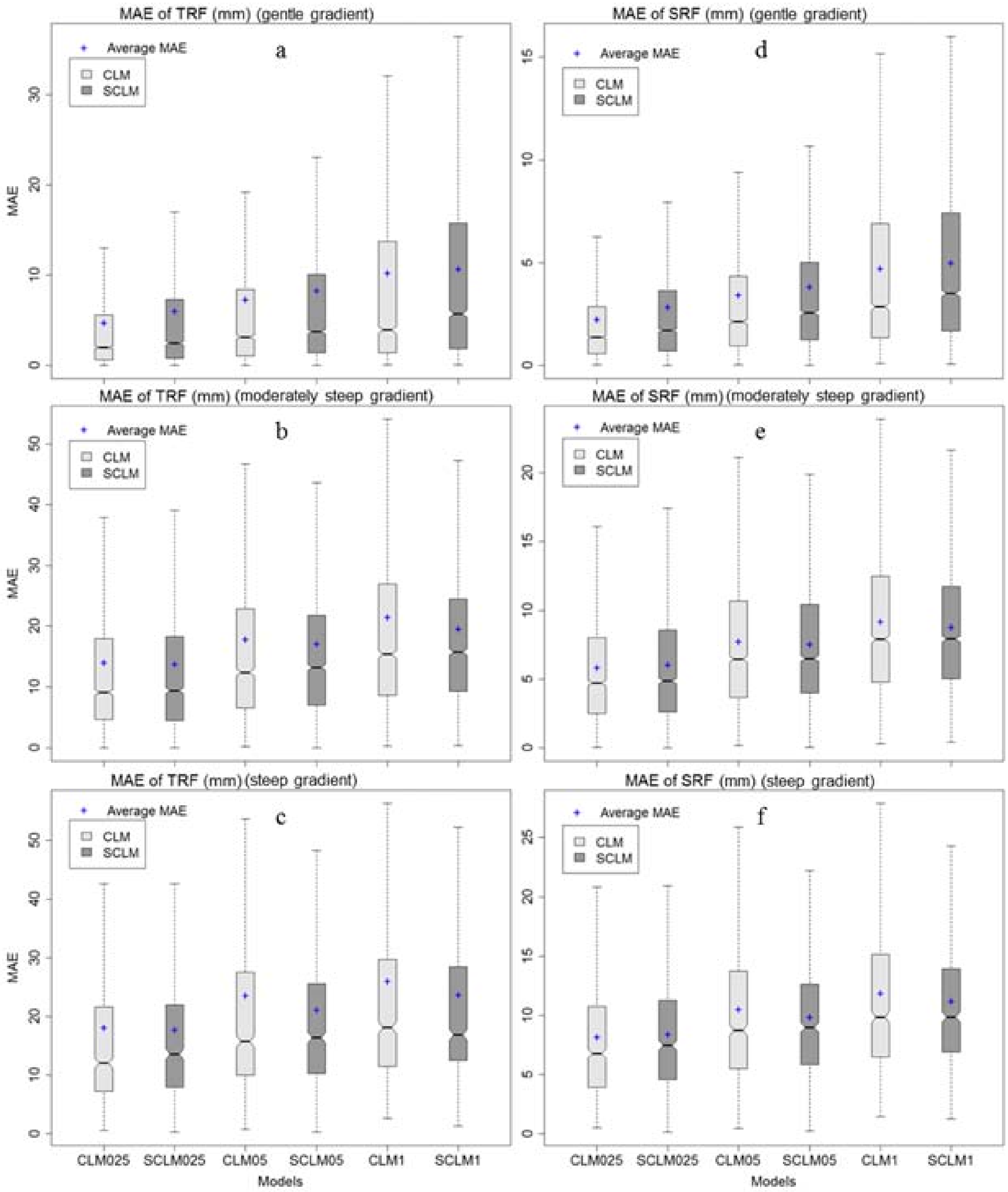

Figure 7: Differences in MAEs between CLM and SCLM on total runoff (left panel) and surface runoff (right panel) calculated by comparing the coarser scales (disaggregated to 0.125^{o} resolution) to CLM0125 and SCLM0125 over gentle gradient (a and d), moderately steep gradient (b and e) and steep gradient (c and f) topographic regions. On each boxplot, the central mark (notch) is the median (q_2), the edges of the boxplot are the 25th (q_1) and 75th (q_3) percentiles, and the whiskers extend to the most extreme data points (q_3 + 1.5 x interquartile range ($q_3 - q_1$) and q_1 – 1.5 x interquartile range ($q_3 - q_1$)); outliers are not considered.

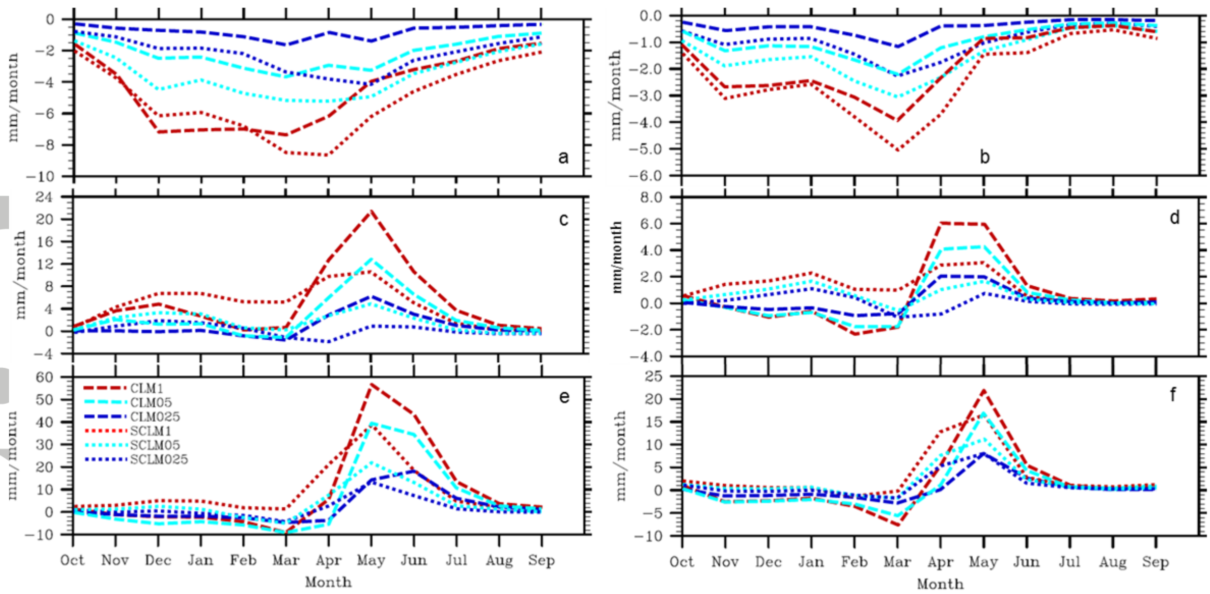

Figure 8: Differences in total runoff (left panel) and surface runoff (right panel) simulations from coarser scale CLMs and SCLMs as compared to the corresponding simulations from the fine scale CLM and SCLM averaged over the gentle gradient (a and b), moderately steep gradient (c and d) and steep gradient (e and f) topographic regions, respectively. In this comparison, results from the coarser scales CLMs and SCLMs were disaggregated to the corresponding 0.125^{o} degree scales.

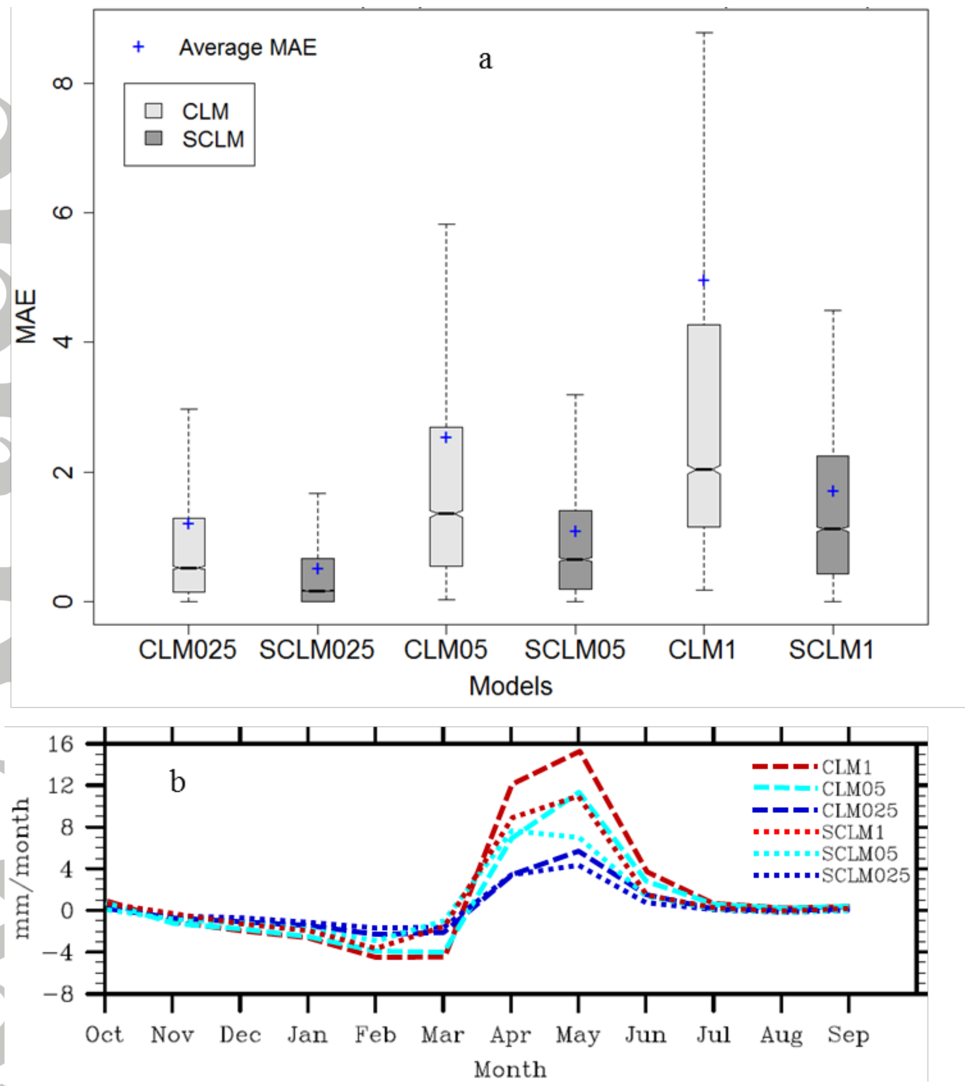

Figure 9: Differences between SCLM and CLM on (a) MAEs (mm) of snowfall and (b) bias in long term monthly average snowmelt (mm/month) at basin level. On each boxplot, the central mark (notch) is the median (q_2), the edges of the boxplot are the 25th (q_1) and 75th (q_3) percentiles, and the whiskers extend to the most extreme data points (q_3 + 1.5 x interquartile range ($q_3 - q_1$) and q_1 – 1.5 x interquartile range ($q_3 - q_1$)); outliers are not considered. Figure 10: Spatial distribution of long term average surface runoff (SRF) in mm/year for CLM0125 (a) and SCLM0125 (b); and MAEs (mm) of the corresponding coarser resolutions: c) CLM025, d) SCLM025, e) CLM05, f) SCLM05, g) CLM1 and h) SCLM1.

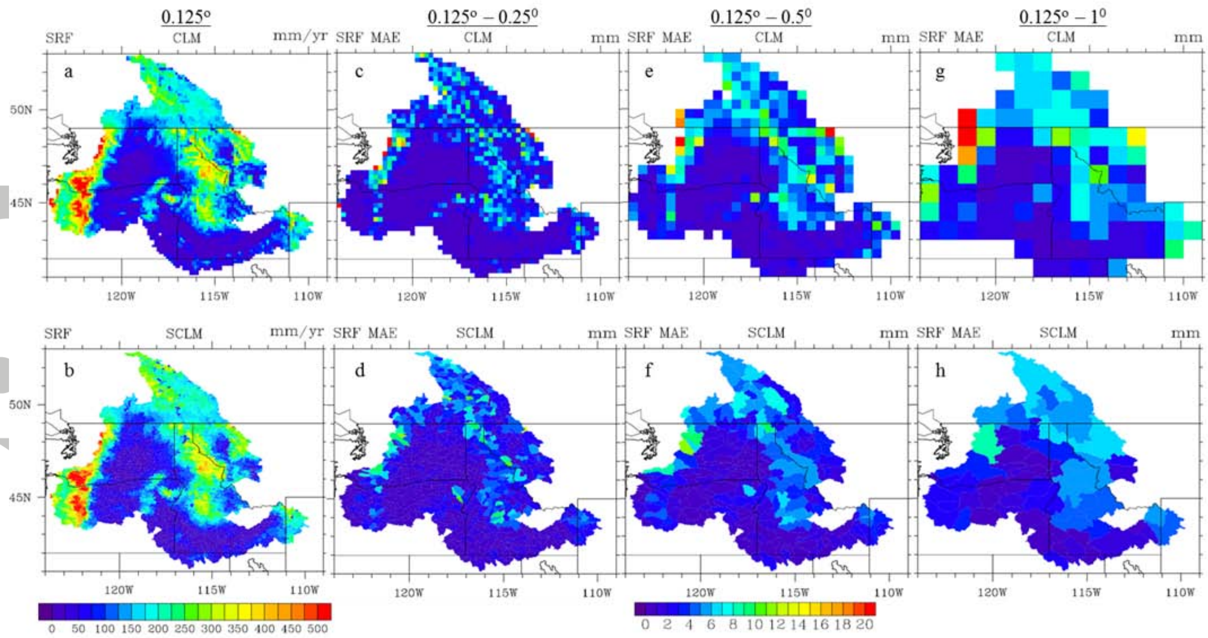

Figure 10 Spatial distribution of long term average surface runoff (SRF) in mm/yearCLM0125 (a) and SCLM0125 (b); and MAEs mm of the corresponding coarser resolutions: c) CLM025, d) SCLM025, e) CLM05, f) SCLM05, g) CLM1 and h) SCLM1.

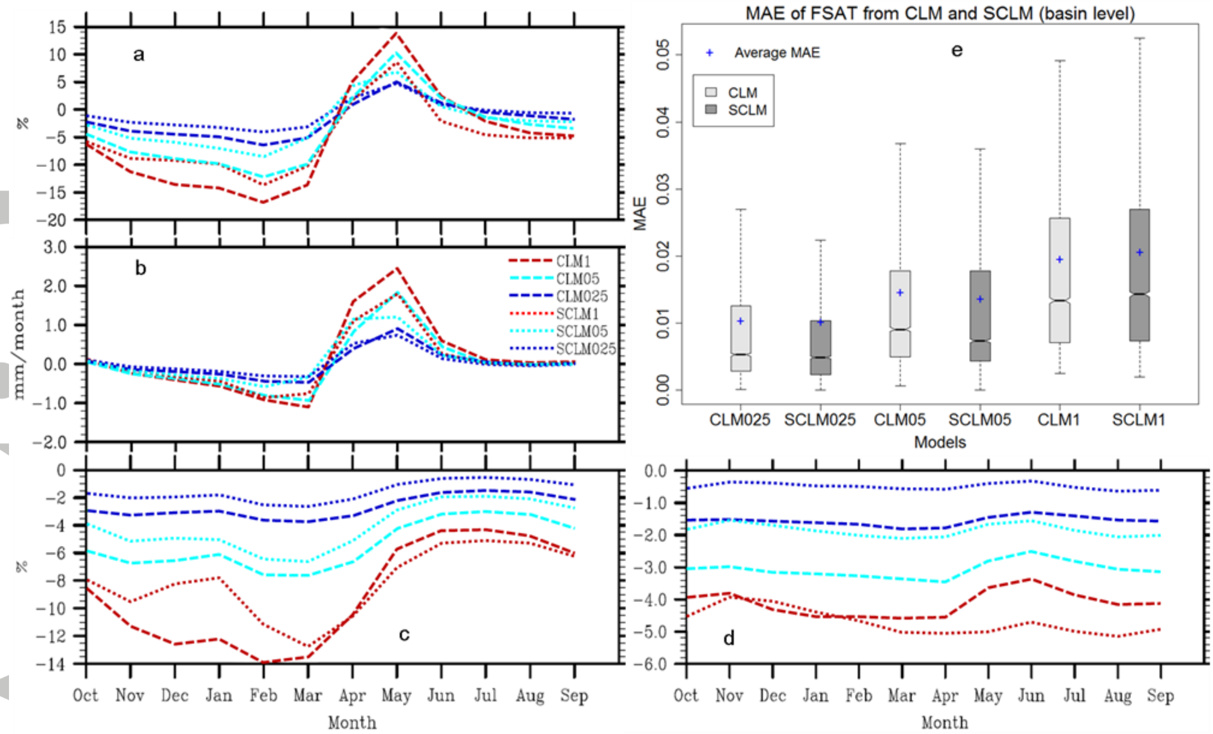

Figure 11: The difference between the coarse scale CLM and SCLM compared to the CLM0125 or SCLM0125 aggregated to the coarse resolutions for the (a) saturated fraction of surface runoff (%), (b) snowmelt component of saturated surface runoff (mm/month), (c) rain component of saturated surface runoff of the seasonal saturated fraction (%) and (d) FSAT (%). The MAEs (fraction) of saturated fraction, FSAT, are shown in (e). On each boxplot, the central mark (notch) is the median (q_2), the edges of the boxplot are the 25th (q_1) and 75th (q_3) percentiles, and the whiskers extend to the most extreme data points (q_3 + 1.5 x interquartile range ($q_3 - q_1$) and q_1 – 1.5 x interquartile range ($q_3 - q_1$)); outliers are not considered.

Table 1: CRB grid- and basin-based area comparison

Model resolution	Average grid size (km^2)	Average subbasin size (km^2)	Average size ratio (grid/subbasin)
0.125°	132.571	108.85	1.22
0.25°	530.317	573.33	0.93
0.5°	2121.61	2184.01	0.97
1°	8492.59	8706.9	0.98

Table 2: Area covered within each topographic class (km^2)

	CLM	SCLM
Steep gradient topography	88139.34	85510.59
Moderately steep gradient topography	245414.2	248216.1
Gentle gradient topography	317885.0	319291.4
Total	651438.54	653018.09

Table 3: Comparison of scalability on surface (SRF) and total (TRF) runoff using average MAEs (mm/month) calculated at the 0.125^{o} resolution grids/subbasins

Model resolution	SRF		TRF	
	CLM	SCLM	CLM	SCLM
0.25°	2.63	1.89	6.22	4.51
0.5°	3.79	2.45	8.28	5.57
1°	5.22	2.68	11.86	5.94

Table 4: Comparison of scalability on surface runoff (SRF) and total runoff (TRF) using Average Relative Error (RE)

Model	Average Winter RE (%)		Average Spring RE (%)		Average Annual RE (%)	
	SRF	TRF	SRF	TRF	SRF	TRF
SCLM025	2.42	1.73	6.61	4.09	3.02	2.66
SCLM05	3.70	2.64	9.87	7.73	4.32	4.36
SCLM1	5.33	3.02	13.65	11.96	5.92	6.13
CLM025	4.94	2.98	6.31	4.43	3.70	3.35
CLM05	10.22	5.22	12.38	8.55	7.36	6.28
CLM1	15.38	7.73	17.30	14.08	10.60	9.36

Table 5: Significance test results (p-values) on MAE differences between the two approaches on major fluxes and related parameters at the coarse (0.25^o, 0.5^o, and 1^o) resolutions compared to the 0.125^o resolution aggregated to the respective coarse resolution

Modeled Variables	**Spatial Scale**		
	0.25°	**0.5°**	**1°**
Total Runoff	1.06E-12	1.42E-04	4.34E-04
Surface Runoff	3.19E-18	3.43E-07	1.91E-05
Subsurface Runoff	1.15E-08	5.47E-03	4.39E-03
Evapotranspiration	1	0.9994457	0.213864
Snow	5.20E-56	4.70E-14	6.26E-06
F_{sat}	0.01736	0.0356	0.8199