Geosci. Model Dev. Discuss., 6, 2699–2730, 2013 www.geosci-model-dev-discuss.net/6/2699/2013/ doi:10.5194/gmdd-6-2699-2013 © Author(s) 2013. CC Attribution 3.0 License.



This discussion paper is/has been under review for the journal Geoscientific Model Development (GMD). Please refer to the corresponding final paper in GMD if available.

A subbasin-based framework to represent land surface processes in an Earth System Model

H.-Y. Li¹, M. Huang¹, T. Tesfa¹, Y. Ke², Y. Sun³, Y. Liu¹, and L. R. Leung¹

¹Pacific Northwest National Laboratory, 902 Battelle Boulevard, Richland, WA 99352, USA ²Base of the State Key Laboratory of Urban Environment Process and Digital Modelling, Department of Resource Environment and Tourism, Capital Normal University, 105 Xi San Huan Bei Lu, Beijing 100048, China

³Department of Hydraulic Engineering, State Key Laboratory of Hydroscience and Engineering, Tsinghua University, Beijing 100084, China

Received: 10 April 2013 - Accepted: 23 April 2013 - Published: 3 May 2013

Correspondence to: H.-Y. Li (hongyi.li@pnl.gov)

Published by Copernicus Publications on behalf of the European Geosciences Union.





Abstract

Realistically representing spatial heterogeneity and lateral land surface processes within and between modeling units in earth system models is important because of their implications to surface energy and water exchanges. The traditional approach of using

- ⁵ regular grids as computational units in land surface models and earth system models may lead to inadequate representation of subgrid heterogeneity and lateral movements of water, energy and carbon fluxes, especially when the grid resolution increases. Here a new subbasin-based framework is introduced in the Community Land Model (CLM), which is the land component of the Community Earth System Model (CESM). Local
- processes are represented assuming each subbasin as a grid cell on a pseudo grid matrix with no significant modifications to the existing CLM modeling structure. Lateral routing of water within and between subbasins is simulated with the subbasin version of a recently-developed physically based routing model, Model for Scale Adaptive River Routing (MOSART). As an illustration, this new framework is implemented in the topo-
- ¹⁵ graphically diverse region of the US Pacific Northwest. The modeling units (subbasins) are delineated from high-resolution Digital Elevation Models (DEMs) while atmospheric forcing and surface parameters are remapped from the corresponding high resolution datasets. The impacts of this representation on simulating hydrologic processes are explored by comparing it with the default (grid-based) CLM representation. In addition,
- the effects of DEM resolution on parameterizing topography and the subsequent effects on runoff processes are investigated. Limited model evaluation and comparison showed that small difference between the averaged forcing can lead to more significant difference in the simulated runoff and streamflow because of nonlinear lateral processes. Topographic indices derived from high resolution DEMs may not improve
- the overall water balance, but affect the partitioning between surface and subsurface runoff. More systematic analyses are needed to determine the relative merits of the subbasin representation compared to the commonly used grid-based representation, especially when land surface models are approaching higher resolutions.





1 Introduction

The representation of land surface processes is a key component of earth system models. It has been widely recognized that development of planetary boundary layer, initiation of shallow and deep convections, and cloud formation and precipitation are

- ⁵ sensitive to spatial heterogeneity of hydrologic state variables such as soil water distribution and snow cover (Quinn et al., 1995; Leung and Ghan, 1995, 1998). Hence parameterizations of spatial heterogeneity in land surface models must account for the lateral redistribution of water that subsequently affects the simulated water and energy exchange with the atmosphere (Li et al., 2011; Liang et al., 1996; Niu et al., 2005).
- ¹⁰ The most common practice in land surface modeling to resolve spatial heterogeneity is to divide a study domain into a number of regular latitude-longitude or other quasiuniform rectangular grids for computational convenience. However, subbasin-based representation, i.e., dividing the study domain into irregular subbasins, offers distinct advantages over the traditional grid-based representation. First, the subbasin-based
- ¹⁵ representation follows the natural topographic divides and river network structure that strongly govern hydrological processes such as surface runoff and streamflow. Figure 1 shows an example of the Columbia River basin in the grid- and subbasin-based representations overlaid by a river network. As highlighted in Fig. 1 by the dashed red line, a single grid cell in the grid-based representation very often crosses over sev-
- eral channel reaches which leads to great difficulties in parameterizing runoff routing (Guo et al., 2004; Wu et al., 2011, 2012; Wen et al., 2012). Second, a single grid cell in the grid-based representation also often encompasses areas from several subbasins, which challenges the conceptual basis of runoff schemes such as the TOP-MODEL formulation in which topographic variations are of primary importance to runoff
- generation (Beven, 1997). For example, the key parameter of such runoff schemes, the topographic index and its statistical distribution (within a spatial unit), essentially measures the accumulated contributing areas from natural divides to the valley and then to the channels. Third, at very high resolution, the grid-based approach must be





modified to account for lateral redistribution of water from neighboring grid cells, which becomes important in determining the soil moisture states, but in a subbasin-based approach, such requirements are to some extent relaxed because surface water is not redistributed across topographic boundaries of the subbasins. Last but not least, the
⁵ subbasin-based representation provides a bridge that may enhance co-development of the hydrologic component in land surface models with contributions from the land surface modeling and hydrologic science communities because the latter has mostly

focused their theoretical and modeling advances on catchment or subbasin scales.

There have been a few attempts to implement subbasin-based representation in land surface models (Koster et al., 2000; Goteti et al., 2008). Koster et al. (2000) was among the first to adopt this approach to improve parameterizations of spatial variability of soil water in land surface and earth system models. Their study focused more on representing soil moisture while surface water movements and storages, which are closely related to soil moisture, were not discussed explicitly. Goteti et al. (2008) developed

- ¹⁵ a catchment-based hydrologic and routing modeling system (CHARMS) with explicit treatment of surface water bodies and storages. Despite of the important advances, their approach has several limitations in that: (1) routing was essentially based on the unit-hydrograph approach so channel velocity and depth were not directly linked to the discharge; (2) model inputs including forcing and land surface parameters were
- ²⁰ remapped, or disaggregated, from the default CLM input dataset provided by the National Center for Atmospheric Research (NCAR) at a resolution of 0.5° (Lawrence and Chase, 2007) which is too coarse comparing to the average size of subbasins. Lastly, although catchments were used as the fundamental modeling units in Koster et al. (2000) and Goteti et al. (2008), the subbasin representation has not been systematically compared with the grid-based representation to evaluate its potential advantages in land surface modeling.

In this study we present another attempt on the use of subbasin-based representation in a land surface modeling framework after Koster et al. (2000) and Goteti et al. (2008). More specifically we present technical advances in (1) deriving input





forcings and land parameters from high resolution datasets; and (2) coupling a new physically based river routing model in the subbasin-based framework. We choose CLM4 as the basis for our development because it has a large user community and its use of the TOPMODEL approach for parameterizing runoff may allow it to take more

⁵ advantages of the subbasin representation, as discussed earlier. CLM4 includes extensive modifications compared to earlier versions in parameterizing runoff generation, soil hydrology thermodynamics, and snow (Lawrence et al., 2011). In this work, all modifications are applied to CLM4. For brevity, hereafter we denote the subbasin-based representation of CLM as Subbasin-based CLM, SCLM, while CLM strictly refers to the grid-based representation of CLM4.

To illustrate the subbasin-based framework, we apply this new modeling framework to the Columbia River basin located in the Pacific Northwest Region of the United States, and compare the model simulations using CLM and SCLM. The average size of the subbasins delineated in this study is chosen to be equivalent to a 0.125° lat/lon grid because this resolution is sufficiently high for CLM applications at the regional scale (the default CLM setting provided by the National Center for Atmospheric Research uses a global grid of 0.5° resolution). This technical paper focuses on the technical aspects of SCLM with illustrative comparison with CLM. More in-depth comparison of

SCLM and CLM will be reported later in a separate study.
 The rest of this paper is organized as follow: Sect. 2 introduces the implementation of the subbasin-based framework. Section 3 describes the sources and preprocessing of forcings and land surface parameters to drive the land model. Section 4 provides preliminary analysis of the application over the Columbia River basin as a case study. Section 5 closes with summary and discussion on future directions.





Subbasin-based representation of CLM 2

Subbasins as spatial units 2.1

15

20

2.2

Discussion Pape CLM applications at regional or global scales involve a large number of computational units. A customized parallel algorithm is embedded in CLM to facilitate such large simulations on clustering computers. To accommodate this parallel algorithm, all CLM units are organized into a two dimensional matrix with each node containing a single grid cell. To take advantage of the parallel algorithm without significantly modifying the original **Discussion** Paper computational structure of CLM, the subbasins of a study domain are also organized into a two dimensional matrix on which each subbasin is treated as a single node, to be consistent with the grid-based representation. Using this subbasin-based representation, all the local land surface processes such as water and energy transfer between the land surface and the atmosphere, as well as subgrid (or within-subbasin) processes such as runoff generation, are represented assuming each subbasin as a pseudo grid cell without significantly modifying the existing CLM modeling structure. Note that the Discussion Paper SCLM structure has been tested in a preliminary manner at small watersheds (Li et al., 2011; Huang et al., 2013), but without invoking the routing component because the river transport model (RTM) embedded in CLM4 and its supporting parameters are not intended for the subbasin-based representation. The next section introduces the coupling of a new river routing model to SCLM. Coupling a new river routing model to SCLM The above matrix representation of CLM grids does not distort the real spatial arrange-Discussion ments among the grid cells, i.e., grid cells that are neighbors in a model domain of a region are still neighbors in the matrix because the grids have regular structure (e.g., each rectangular has exactly 8 neighbors). For SCLM, each subbasin can have differ-

ent numbers of neighboring subbasins, so the (2-D) matrix structure cannot reflect the 25 real spatial arrangements or linkages among the subbasins. We therefore impose an





extra indexing system by assigning a unique index to each individual subbasin. The linkages between the subbasins, i.e., upstream/downstream relationships and other parameters needed for the runoff routing are all preprocessed and identified by their indices defined using the same indexing system and stored in the input datasets.

- The Model for Scale Adaptive River Transport, MOSART, is a large scale rout-5 ing model recently developed with explicit treatment of routing processes at hillslope, across tributaries (within a spatial unit) and through the main channel (for more details please refer to Li et al., 2013). MOSART can be applied at both grid- and subbasin-based representations. Li et al. (2013) describes the concept and framework of MOSART and evaluation of its grid-based representation in the US Pacific Northwest
- 10 region at multiple spatial resolutions. In this work, MOSART was modified to follow the subbasin structure for direct coupling with SCLM. The surface and subsurface runoff produced from the SCLM units is fed into the MOSART units by a one-to-one mapping based on the indexing system described above. MOSART then routes the runoff within
- and between the subbasins all the way to the ocean or basin outlet.

2.3 Subbasin delineation

To set up SCLM, we utilized the 90 m Digital Elevation Model (DEM), the 15 arcsec river networks, and the basin boundary of CRB from the Hydrological data and maps based on SHuttle Elevation Derivatives (HydroSHEDS) (Lehner et al., 2008). Although

- DEMs at resolutions of 30 m or higher over the study area can be obtained from the 20 United States Geological Survey (USGS) database, the goal of our study is to develop a framework suitable for SCLM applications worldwide. Therefore, a global database (i.e., HydroSHEDS) is used in this study. The 15 arcsec river network is reconciled with the 90 m DEM over CRB for hydrologic conditioning to ensure a consistent delineation
- of the river network with HydroSHEDS. Using ArcSWAT (Neitsch et al., 2005), we de-25 lineate sub-basins within CRB, as well as a river network consistent with the sub-basin boundaries, using the hydrologically-conditioned DEMs as inputs. For comparison with the grid-based application of CLM4 at 0.125° resolution, the threshold area for the





subbasin delineation is adjusted iteratively until the average subbasin size is roughly equivalent to the 0.125° grid as shown in Fig. 1. In addition, the resulting subbasin dataset was edited in ArcGIS to merge extremely small subbasins with their neighboring subbasins. We eventually obtain 5999 sub-basins with an average drainage area

- as about ~ 100 km². These subbasins are then organized into a 77 × 78 matrix, with the additional seven grid cells masked out as non-land cells and therefore excluded from the simulations. Given that subbasins could only be defined over land, any pseudo grid cell in the matrix is either 100 % or 0 % land. This is different from the grid-based applications, in which a single grid can be occupied fractionally by land or ocean.
- ¹⁰ ArcSWAT also provides the subbasin parameters needed for the routing model (MOSART) including subbasin upstream/downstream dependence information, accumulated contributing area and other channel parameters such as slope and length. The bankfull channel width and depth values are derived based on the empirical hydraulic geometry relationships estimated in Li et al. (2013).

3 Preparation of model inputs: forcing and land surface parameters

3.1 Forcing

The meteorological forcing in this study was extracted from the phase two North America Land Data Assimilation System (NLDAS-2) at an hourly time step from 1979–2008 (Xia et al., 2012), including precipitation, shortwave and long-wave radiation, air tem-²⁰ perature, humidity, surface pressure and wind speed at 0.125° resolution derived from the 32 km resolution 3 hourly North American Regional Reanalysis (NARR). The precipitation fields in NLDAS-2 were produced by combining observations from field stations, level 4 precipitation retrievals from NEXRAD system over the county, and satellites, and are well suited for hydrologic studies. For grid based applications, the NLDAS-2 faraing data are applied directly. For SCLM, on area average algorithm is applied to

²⁵ 2 forcing data are applied directly. For SCLM, an area-average algorithm is applied to remap the NLDAS-2 forcing to the subbasins defined by their boundaries. That is, the





algorithm computes the value of each meteorological variable in a subbasin (S_n) as the average of the corresponding variable from all the 0.125° grid cells that intersect with the subbasin (G_{ni}) weighted by the overlapping areas.

$$S_n = \sum_i G_{ni} A_{ni}$$

⁵ where A_{ni} is ratio of the overlapping area between grid cell *n* and the subbasin divided by the subbasin area. ArcGIS is used to link the subbasins to the intersecting grids and to compute the fractional areas of the intersecting grids. Figure 2 shows the spatial distributions of annual mean precipitation, surface temperature and mean downward shortwave radiation. It can be seen from the figure that the differences between the 10 two representations are very small.

3.2 Land surface parameters

Ke et al. (2012) has developed a new land parameter dataset at 0.05° resolution based on the most recent MODIS land cover and improved MODIS leaf area index (LAI) products. Soil texture was generated based on a hybrid of 30 arc-second State Soil
Geographic Database (STATSGO, now referred to as the US General Soil Map) (for CONUS) and 5 min Food and Agriculture Organization (outside CONUS) 16-category two-layer soil type data (Chen et al., 2007; Miller and White, 1998). The two-layer soil type data was then converted to composition of clay and sand (Cosby et al., 1984; Dai et al., 2003) within each 30 arc-second grid cells and interpolated to 10 vertical layers down to 3.8 m depth. Other land surface parameters such as soil color and soil organic matter were derived from the default 0.5° CLM4 global input dataset provided by National Center for Atmospheric Research (NCAR). The CLM4 preprocessing package

(Oleson et al., 2010) was used for the grid based CLM simulations in this study.

The SCLM soil, land cover, and vegetation parameters were derived by overlaying the subbasin boundaries over the above grids. Similar to the forcing parameters, ArcGIS is used to link the subbasins to the grids and calculate the area weights of the



(1)



intersecting grids. Consistent with the CLM4 preprocessing package (Oleson et al., 2010), soil properties such as percent clay, percent sand etc. are calculated using area dominant algorithm, where each parameter value for the subbasins is assigned to the value covering largest fraction of the subbasin. Land cover characteristics and Plant
 ⁵ Functional Type (PFTs) for each subbasin are determined using Eq. (1). Leaf Area Index (LAI) and Stem Area Index (SAI) parameters for each subbasin are calculated

using PFT weighted area-average algorithm as:

$$LS_{n,m} = \sum_{i} A_{wti} LS_{nmi} P_{mi}$$

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

$$A_{\text{wt}i} = \frac{A_{ni}}{\sum A_{ni} P_{mi}}$$

where A_{ni} is the same as in Eq. (1). Figure 3 shows some of the surface parameters such as the total leaf area index, percentage of clay in the soil and soil color. Between the two representations, the differences of most land surface parameters are small due to the high resolution of source data. But this is not the case for soil color since the resolution of source data is 0.5° which is much coarser.

3.3 Topography

 $f_{\rm max}$ is an important hydrologic parameters which could affect the partitioning of precipitation into surface runoff and infiltration (Hou et al., 2012; Huang et al., 2013; Li et al., 2011; Niu et al., 2005). In this work, the values of $f_{\rm max}$ are derived following the algorithm described in Niu et al. (2005) and Niu and Yang (2006) for both SCLM and CLM simulations. Based on the DEMs, compound topographic indices (CTIs) are first

(2)

(3)



derived following the definition used in TOPMODEL (Beven; 1997; Quinn et al., 1995) by ArcGIS. The CTIs are then fitted to a distribution, within each subbasin, to estimate the topographic relevant hydrologic parameters (i.e., f_{max} and C_s). In CLM4 and its previous versions (e.g., CLM3, CLM3.5), these parameters are derived from coarse resolution (e.g., 1 km) DEM (Niu et al., 2005) due to the lack of higher resolution DEMs in global domain. However, as discussed in our previous study (Li et al., 2011), the estimation of these parameters using 1 km DEMs is problematic due to its inconsistency with hydrology theory. Interested readers are referred to Li et al. (2011) for details. With the newly available HydroSHEDS (Lehner et al., 2008) database, 90 m DEM data are now available globally. We therefore estimated f_{max} values using the 90 m DEM from 10 HydroSHEDS, shown in Fig. 4 as SCLM-exp1. Similar approach is used to derive topographic relevant parameters for CLM, shown in Fig. 4 as CLM-exp1. Overall, f_{max} values derived at the grid-based representation are slightly larger than those at the subbasin-based representation with less dispersed distribution. Most of the simulation results hereinafter (i.e., shown in Figs. 5–9) are based on the f_{max} fields derived from 15 the 90 m DEM. The f_{max} fields noted as SCLM-exp2 and CLM-exp2 are used in Fig. 10 only, and will be discussed later.

4 Coupled simulation and results

To spin-up, the meteorological forcing (1979–2008) is fed to the model repeatedly un-

til the state variables (soil moisture, soil temperature and groundwater table depth) reached equilibrium. The resulted state variables are then used as the initial conditions for the final model run from which the simulation results are analyzed and discussed in the following.

An important aspect to compare the coupled SCLM-MOSART and CLM-MOSART ²⁵ models is how they simulate streamflow. Figure 5 shows the comparison of the simulated streamflow from SCLM-MOSART and CLM-MOSART at a number of major USGS stream gauges (as shown in Fig. 1a). There is a systematic phase shift between





the streamflow simulated by both models, with SCLM-MOSART producing higher streamflow in January–March and lower streamflow in April–June. Also shown in Fig. 5 is the naturalized streamflow obtained from the Surface Hydrology Group, University of Washington (http://www.hydro.washington.edu/2860/). Here the naturalized streamflow

- ⁵ can be viewed as the observation because human influences such as reservoir operation have been removed, although the removal procedure introduces uncertainties besides measurement errors. Due to the availability of the naturalized streamflow data, the analysis period is chosen as October 1979–September 1989 for Fig. 5 and other related figures herein. Compared to the naturalized streamflow, both SCLM-MOSART
- and CLM-MOSART underestimate streamflow in the northern part of the Columbia River basin (as indicated by the comparison at CORRA, CHIEF and PRIRA) and overestimate streamflow in the southern part (as indicated by the comparison at BROWN and ICEHA). Since the northern part is wetter and provides more runoff than the southern part (e.g., comparing the streamflow at PRIRA and ICEHA), both SCLM-MOSART
- and CLM-MOSART underestimate streamflow at DALLE, which is downstream of the conjunction where the two parts join each other. Also, both SCLM-MOSART and CLM-MOSART produce peak flow one month earlier than the naturalized/observed streamflow. This can be attributed to the parameterization of snow processes in CLM that very often leads to earlier snowmelt (Wang et al., 2008; Li et al., 2011).
- Due to the earlier snowmelt, one may conclude that SCLM-MOSART does not necessarily perform better than CLM-MOSART in simulating streamflow, although SCLM-MOSART reproduces the shape of the monthly mean hydrographs better (i.e., SCLM-MOSART more often simulates a flatter peaks compared to CLM-MOSART that mostly simulates single dominant peaks). One possible reason SCLM-MOSART is not per-
- forming better than CLM-MOSART is that the runoff simulations in SCLM and CLM are governed by the same set of hydrological formulations and parameters originally calibrated for CLM based on the grid-based configuration at the global scale. If reproducing the observed streamflow is the target, a fair comparison between SCLM-MOSART and CLM-MOSART should be conducted with separate parameter calibration for each,





which has been a topic of research in separate studies (Hou et al., 2013; Huang et al., 2013; Sun et al., 2013), and is beyond the scope of this study. Here, our main objective is to demonstrate the new configuration of SCLM-MOSART and identify differences in the hydrological responses caused purely by different approaches to delineating

the fundamental spatial units without changes in model parameters or adjustments of model parameterizations to take advantage of one representation over the other. The phase shift between the monthly mean hydrographs simulated by SCLM-MOSART and CLM-MOSART is thus a subject for further investigation.

Streamflow is a direct product of runoff routing processes that are fed by, and therefore directly controlled by, runoff generation in terms of both magnitude and timing. Runoff generation itself is controlled by the interactions between climate and landscape properties and the latter two are very often closely interrelated to each other. In order to understand the simulated runoff generation in different climate regimes, the subbasins/grids in the Columbia River basin are grouped into different regimes by

- rainfall fraction (fraction of rainfall to the total precipitation). It is found that the spatial distributions of rainfall fraction for CLM and SCLM are largely consistent with the spatial distribution of elevation as shown in Fig. 1. This is not surprising since rainfall/snowfall partitioning of precipitation is dominated by near surface air temperature, which in turn is closely related to elevation variation. Figure 6 shows the three regimes
- defined based on rainfall fraction: snow dominated areas with rainfall fraction ranging between 0.1–0.5; snow-rain transition areas with rainfall fraction ranging between 0.5–0.75; and rain-dominated areas with rainfall fraction ranging between 0.75–1.0. The grids in CLM and subbasins in SCLM are classified based on the same criteria. The total area of each regime is listed in Table 1. Using different thresholds of 0.1, 0.4,
- 0.7 and 1.0 does not change the conclusions except the snow-dominate area is rather small as the two middle thresholds decrease. Hence subsequent analysis is based on the classification with thresholds of 0.1, 0.5, 0.75, and 1.0.

Figure 7 shows the seasonal variation of runoff averaged over the whole Columbia River basin and the three different regimes. The phase shift between SCLM and CLM is





consistently shown among different runoff components, i.e., surface runoff, subsurface runoff, and the total. This phase shift is more apparent in the snow-dominated areas than in the rain-dominated areas and more apparent in the surface runoff than the subsurface runoff. In the snow-dominated areas, SCLM produces less subsurface runoff

⁵ due to drier soil. The latter is because that the evaporation from bare soil and canopy simulated by SCLM is overall slightly higher than that simulated by CLM. Nevertheless, the transpiration simulations by SCLM and CLM show almost no difference. The total evapotranspiration simulations thus show slight difference (figures not shown).

Comparing with the phase shift in the simulated streamflow as shown in Fig. 5, how-¹⁰ ever, the phase shift in the simulated runoff is less significant. The transformation from runoff to streamflow is captured by the routing process which is nonlinear in nature. Therefore one could infer that it is this nonlinear transformation that has amplified the phase difference between the runoff simulated by SCLM and CLM. It is then logical to ask whether and how this phase difference exists in the climatic forcings that are major ¹⁵ drivers of runoff generation processes.

Figure 8 shows the seasonal variation of precipitation, temperature, and the partitioning of precipitation into rainfall and snowfall. From the plots for the whole Columbia River basin, one can see that there is no difference between the mean precipitation and temperature averaged over all SCLM subbasins and CLM grids, which is expected

- ²⁰ because the remapping from grids to subbasins conserves the area averaged forcings which are inputs to the model. The partitioning of precipitation into rainfall/snowfall is not the inputs, but rather the results of model simulations. One could observe that the mean snowfall averaged over all SCLM subbasins is slightly lower than that averaged over all CLM grids, and vice versa for rainfall. From the plots for the different regimes,
- this difference is even clearer for snowfall in snow-dominated areas and rainfall in raindominated areas. The near surface air temperature is the dominant control of partitioning precipitation into rainfall/snowfall. There are some differences of the mean temperatures over snow-dominated areas between SCLM and CLM, but they are barely discernible. One could now infer from Fig. 8 that the different representations lead to very





small differences in the mean temperature in snow-dominated areas between SCLM and CLM, which lead to some discernible differences of mean snowfall/rainfall over the whole Columbia River basin between SCLM and CLM. The latter then leads to the obvious difference in runoff as shown in Fig. 7 due to the nonlinear runoff generation ⁵ processes.

Based on the above discussion, we conclude that the slight difference in the forcings such as temperature between SCLM and CLM, although from the same sources, has been augmented first via runoff generation and then routing processes, leading to significant phase difference in the simulated streamflow. Both runoff generation and routing processes are nonlinear, and more importantly, involve horizontal movements of water that can indeed be influenced by the different representations of subbasin versus grid, although they are differentiated only by the horizontal boundaries between the fundamental units.

10

Figure 9 shows the energy balance simulated by SCLM and CLM. The difference between the two different representations is obviously less significant than that indicated by the simulated streamflow. This is because the energy balance is more directly controlled by vertical processes such as transfer of moisture through the soil column and only indirectly affected by lateral hydrological processes (through the interaction between runoff generation and soil moisture variation).

The above numerical simulations using SCLM and CLM use a number of land surface parameters as partially listed in Sect. 3. SCLM and CLM may exhibit different sensitivities to those surface parameters. To examine the role of surface parameters on the simulations, we choose $f_{\rm max}$ as an example since $f_{\rm max}$ is an important parameter of runoff generation. For this purpose, we have also derived the $f_{\rm max}$ fields based

²⁵ on the standard 1-km HYDRO1K dataset as used the default CLM4 input dataset provided by National Center for Atmospheric Research (NCAR) (Oelson et al., 2010; Niu et al., 2005), shown in Fig. 4 as SCLM-exp2 and CLM-exp2. One can observe that the f_{max} values derived from the 90 m DEM are apparently larger than those derived from the 1 km DEM and the distribution of the former is less dispersed than the latter.





Figure 10 shows the comparison of runoff generation simulations resulted from different f_{max} fields. Overall, the difference in runoff simulations caused by differences in f_{max} fields lies largely in the partitioning between surface and subsurface runoff. That is, f_{max} derived from 90 m DEM leads to higher surface runoff, which is compensated by lower subsurface runoff. Hence the difference between the total runoff simulated using two sets of f_{max} values is negligible. It is therefore reasonable to infer that the difference between simulated streamflow is also not significant (at least true at the monthly scale or larger). Interestingly, this conclusion holds for both SCLM and CLM.

5 Summary

In this study we have described the development of a subbasin-based representation of CLM including preprocessing of inputs from high resolution datasets and coupling with a physically based river routing model. With this we have compared the results from this new framework with those from the default grid-based CLM over the mountainous Columbia River basin. With the limited comparison, we preliminarily found that: (1)
 Small difference between the averaged forcing leads to more significant difference in simulated runoff and streamflow because of nonlinear horizontal processes; and (2) Topographic indices derived from high resolution DEM may not improve the overall water balance, but does affect the partitioning between surface and subsurface runoff.

This paper is a first step towards a systematic investigation of the subbasin-based representation of land surface models. We intend to continue with several studies to further improve our understanding of the relative merits of this approach. For example, in a parallel study, we are able to show that SCLM is less sensitive to the change of spatial resolutions than CLM. Such analyses covering a wider range of model resolutions will be particularly useful as the land surface modeling community is exploring the feasibility and advantages of ultra-high resolution (e.g., 1 km resolution globally) (Wood et al., 2011), and conceptually the differences between the two approaches may be larger as lateral water redistribution becomes increasingly





important in determining soil moisture states at smaller spatial scales. The results in this study also suggest that without proper calibration, SCLM may not necessarily perform better than CLM, so some future research to include parameter calibration of SCLM and CLM on smaller basins with good forcing and evaluation data would be useful to test this point. Given that CLM is the land component of an earth system model and can interact with the atmosphere component and ecoan component

- tem model and can interact with the atmosphere component and ocean component, it would be valuable to examine how the subbasin-based representation of terrestrial processes may affect the global cycling of water and energy among land, atmosphere and ocean. Such a scientific pursuit is now supported by recent progresses made in
- ¹⁰ software engineering. With the latest public versions of the Community Climate System model and the Community Earth System model (i.e., CCSM4 and CESM1) there are a suite of new coupling capabilities in the CPL7 coupler that allow more flexibility and extensibility to address the challenges faced by the earth system modeling community (Gent et al., 2010; Lawrence et al., 2011). As documented in Craig et al. (2012),
- the CPL7 coupling architecture adapts a totally new framework by including a toplevel driver that calls model component initialization, run, and finalization methods through the specified interfaces. This framework embraces new flexibility in two aspects: (1) placing the model components on relatively arbitrary hardware processor layouts; (2) running them sequentially, concurrently, or in a mixed way. Significant im-
- ²⁰ provements have been made to the memory and performance scaling of the coupler to support configurations with resolution much higher than before so that CCSM4/CESM1 could handle global simulations with high resolutions (e.g., 0.1°). Moreover, under the CPL7 structure, any components within CESM can be run on unstructured grids (http://www.cesm.ucar.edu/models/cesm1.1/notable_improvements.html) so it is possible to couple a land model that uses subhasin representation.
- $_{\rm 25}$ $\,$ ble to couple a land model that uses subbasin representation.

5

Acknowledgements. This study is supported by the Platform for Regional Integrated Modeling and Analysis (PRIMA) initiative. This research was also partly supported by the Office of Science of the US Department of Energy as part of the Earth System Modeling and Integrated





Assessment Research programs. The Pacific Northwest National Laboratory is operated by Battelle for the US Department of Energy under Contract DE-AC06-76RLO1830.

References

10

20

25

Beven, K.: TOPMODEL: a ritique, Hydrol. Process., 11, 1069–1085, 1997.

- ⁵ Chen, F., Manning, K. W., LeMone, M. A., Trier, S. B., Alfieri, J. G., Roberts, R., Tewari, M., Niyogi, D., Horst, T. W., Oncley, S. P., Basara, J. B., and Blanken, P. D.: Description and evaluation of the characteristics of the NCAR high-resolution land data assimilation system, J. Appl. Meteorol. Clim., 46, 694–713, 2007.
 - Cosby, B. J., Hornberger, G. M., Clapp, R. B., and Ginn, T. R.: A statistical exploration of the relationships of soil moisture characteristics to the physical properties of soils, Water Resour. Res., 20, 682–690, 1984.
 - Craig, A. P., Vertenstein, M., and Jacob, R.: A new flexible coupler for earth system modeling developed for CCSM4 and CESM1, Int. J. High Perform. C., 26, 19, 309–328, 2012.
 - Dai, Y. J., Zeng, X., Dickinson, R. E., Baker, I., Bonan, G. B., Bosilovich, M. G., Denning, A.
- S., Dirmeyer, P. A., Houser, P. R., Niu, G.-Y., Oleson, K. W., Schlosser, C. A., Yang, Z.-L., Dickinson, R. E., Baker, I., Bonan, G. B., Bosilovich, M. G., Denning, A. S., Dirmeyer, P. A., Houser, P. R., Oleson, K. W., Schlosser, C. A., and Yang Z-.L.: The common land model, B. Am. Meteorol. Soc., 84, 1013–1023, doi:10.1175/BAMS-84-8-1013, 2003.

Gent, P., Yeager, S., Neale, R., Levis, S., and Bailey, D.: Improvements in a half degree atmosphere/land version of the CCSM, Clim. Dynam., 34, 819–833, 2010.

Gesch, D. B., Verdin, K. L., and Greenlee, S. K.: New land surface digital elevation model covers the Earth, Eos Trans. AGU, 80, 69–70, 1999.

Goteti, G., Famiglietti, J. S., and Asante, K.: A catchment-based hydrologic and routing modeling system with explicit river channels, J. Geophys. Res., 113, D14116, doi:10.1029/2007JD009691, 2008.

- Guo, J., Liang, X., and Leung, L. R.: A new multi-scale flow network generation scheme for land surface models, Geophys. Res. Lett., 31, L23502, doi:10.1029/2004GL021381, 2004.
- Hou, Z., Huang, M., Leung, L. R., Lin, G., and Ricciuto, D. M.: Sensitivity of surface flux simulations to hydrologic parameters based on an uncertainty quantification framework applied to





the Community Land Model, J. Geophys. Res., 117, D15108, doi:10.1029/2012JD017521, 2012.

- Huang, M. and Liang, X.: On the assessment of the impact of reducing parameters and identification of parameter uncertainties for a hydrologic model with applications to ungauged
- 5 basins, J. Hydrol., 320, 37–61, 2006.
 - Huang, M., Hou, Z., Leung, L. Y. R., Ke, Y., Liu, Y., Fang, Z., and Sun, Y.: Uncertainty analysis of runoff simulations and parameter identifiability in the Community Land Model evidence from MOPEX basins, J. Hydrometeorol., submitted, 2013.
 - Ke, Y., Leung, R. L., Huang, M., Coleman, A. M., Li, H., and Wigmosta, M. S.: Development of
- high resolution land surface parameters for the Community Land Model, Geosci. Model Dev.,
 5, 1341–1362, doi:10.5194/gmd-5-1341-2012, 2012.
 - Lawrence, D., Oleson, K. W., Flanner, M. G., Thorton, P. E., Swenson, S. C., Lawrence, P. J., Zeng X., Yang, Z.-L., Levis, S., Skaguchi, K., Bonan, G. B., and Slater, A. G.: Parameterization improvements and functional and structural advances in version 4 of the Community
- Land Model, J. Adv. Model. Earth Sys., 3, M03001, doi:10.1029/2011MS000045, 2011. Lehner, B., Verdin, K., and Jarvis, A.: New global hydrography derived from spaceborne elevation data, Eos Trans. AGU, 89, 93–94, 2008.
 - Leung, L. R. and Ghan, S. J.: A subgrid parameterization of orographic precipitation, Theor. Appl. Climatol., 52, 95–118, 1995.
- Leung, L. R. and Ghan, S. J.: Parameterizing subgrid orographic precipitation and surface cover in climate models, Mon. Weather Rev., 126, 3271–3291, 1998.
 - Li, H., Wigmosta, M. S., Wu, H., Huang, M., Ke, Y., Coleman, A. M., and Leung, L. R.: A physically based runoff routing model for land surface and earth system models, J. Hydrometeorol., in press, 2013.
- Li, H., Huang, M., Wigmosta M. S., Ke, Y., Coleman, A. M., Leung, L. R., Wang, A., and Ricciuto, D. M.: Evaluating runoff simulations from the Community Land Model 4.0 using observations from flux towers and a mountainous watershed, J. Geophys. Res., 116, D24120, doi:10.1029/2011JD016276, 2011.

Liang, X., Wood, E. F., and Lettenmaier, D. P.: Surface soil moisture parameterization of the VIC-2L model: evaluation and modifications, Global Planet. Change, 13, 195–206, 1996.

VIC-2L model: evaluation and modifications, Global Planet. Change, 13, 195–206, 1996.
 Miller, D. A. and White, R. A.: A conterminous united states multilayer soil characteristics dataset for regional climate and hydrology modeling, Earth Interact., 2, 1–26, 1998.





Koster, R. D., Suarez, M. J., Ducharne, A., Stieglitz, M., and Kumar, P.: A catchment-based approach to modeling land surface processes in a GCM: 1. Model structure, J. Geophys. Res., 105, 24809–24822, 2000.

Neitsch, S. L., Arnold, J. G., Kiniry, J. R., Srinivasan, R., and Williams, J. R.: Soil and Wa-

- ter Assessment Tool Theoretical Documentation, version 2005, Grassland, Soil and Water Research Laboratory, Agricultural Research Service, Temple, Texas, available at: http: //swatmodel.tamu.edu/documentation, last access: April 2013, 2005.
 - Niu, G.-Y., Yang, Z.-L., Dickinson, R. E., and Gulden, L. E.: A simple TOPMODEL-based runoff parameterization (SIMTOP) for use in global climate models, J. Geophys. Res., 110, D21106, doi:10.1029/2005JD006111, 2005.

10

25

- Oleson, K. W., Lawrence, D. M., Bonan, G. B., Flanner, M. G., Kluzek, E., Lawrence, P. J., Levis, S., Swenson, S. C., Thornton, P. E., Dai, A., Decker, M., Dickinson, R., Feddema, J., Heald, C. L., Hoffman, F., Lamarque, J.-F., Mahowald, N., Niu, G.-Y., Qian, T., Randerson, J., Running, S., Sakaguchi, K., Slater, A., Stöckli, R., Wang, A., Yang, Z.-L., Zeng, X., and
- ¹⁵ Zeng, X.: Technical description of version 4.0 of the Community Land Model (CLM), National Center for Atmospheric Research, Boulder, 257 pp., 2010.
 - Quinn, P., Beven, K., and Culf, A.: The introduction of macroscale hydrological complexity into land surface–atmosphere transfer models and the effect on the planetary boundary layer development, J. Hydrol., 166, 421–444, 1995.
- Sun, Y., Hou, Z., Huang, M., Tian, F., and Leung, L.-Y. R.: Inverse modeling of hydrologic parameters using surface flux and streamflow observations in the Community Land Model, Hydrol. Earth Syst. Sci., in review, 2013.
 - US Geological Survey: HYDRO1K Elevation Derivative Database, Center for Earth Resources Observation and Sciences, Sioux Falls, SD, available at: http://eros.usgs.gov/#/Find_Data/ Products_and_Data_Available/gtopo30/hydro, 2000.
 - Wang, A., Li, K. Y., and Lettenmaier, D. P.: Integration of the variable infiltration capacity model soil hydrology scheme into the community land model, J. Geophys. Res., 113, D09111, doi:10.1029/2007JD009246, 2008.

Wen, Z., Liang, X., and Yang, S.: A new multiscale routing framework and its evaluation for land

- surface modeling applications, Water Resour. Res., W08528, doi:10.1029/2011WR011337, 2012.
 - Wood, E. F., Roundy, J. K., Troy, T. J., van Beek, L. P. H., Bierkens, M. F. P., Blyth, E., de Roo, A., Döll, P., Ek, M., Famiglietti, J., Gochis, D., van de Giesen, N., Houser, P., Jaffé,





P. R., Kollet, S., Lehner, B., Lettenmaier, D. P., Peters-Lidard, C., Sivapalan, M., Sheffield, J., Wade, A., and Whitehead, P.: Hyperresolution global land surface modeling: meeting a grand challenge for monitoring Earth's terrestrial water, Water Resour. Res., 47, W05301, doi:10.1029/2010WR010090, 2011.

- ⁵ Wu, H., Kimball, J. S., Mantua, N., and Stanford, J.: Automated upscaling of river networks for macroscale hydrological modeling, Water Resour. Res., 47, W03517, doi:10.1029/2009WR008871, 2011.
 - Wu, H., Kimball, J. S., Li, H., Huang, M., Leung, L. R., and Adler, R. F.: A new global river network database for macroscale hydrologic modeling, Water Resour. Res., 48, W09701, doi:10.1029/2012WR012313, 2012.
 - Xia, Y., Ek, M., Wei, H., and Meng, J.: Comparative analysis of relationships between NLDAS-2 forcings and model outputs, Hydrol. Process., 26, 467–474, doi:10.1002/hyp.8240, 2012.

10





Table 1.	The	portion	of	areas	within	three	climate	regimes	(km ²)	١.
		portion	01	uicuo	wwith in i	11100	omnato	regimes	(iviii)	

	SCLM	CLM
Snow-dominated	71 064	87 756
Rain-snow mixing	254 267	223 440
Rain-dominated	327 687	340 240
Total	653018	651 437

Discussion Pa	GN 6, 2699–2	1DD 2730, 2013						
per Discussion F	A subba framewor surface pr Earth Syst HY.	A subbasin-based framework for land surface processes in Earth System Models HY. Li et al.						
Paper	Title	Title Page						
_	Abstract	Introduction						
Disc	Conclusions	References						
ussion	Tables	Figures						
Pap	I	►I.						
er	•	•						
Discussio	Back	Close						
	Full Scr	Full Screen / Esc						
n Pap	Printer-frie	Printer-friendly Version						
ber	Interactive Discussion							





Fig. 1. (a) Subbasin delineation of the Columbia River Basin and locations of USGS stream gauges; **(b)** subbasin delineation overlaid with regular grids at 0.125° resolution (highlighted with red dashed line are example grids containing multiple river channels).







Fig. 2. Forcing at subbasin- and grid-based representations. From top to bottom, annual mean precipitation (PREC), annual mean surface temperature (TBOT) and annual mean downward shortwave radiation (FSDS).





Fig. 3. Surface parameters at subbasin- and grid-based representations. From top to bottom: total leaf area index (TLAI), percent of clay (% CLAY) and soil color (S. COLOR).





Fig. 4. f_{max} distribution at subbasin- and grid-based representations. Here CLM-exp1 and SCLM-exp1 are for the f_{max} fields derived based on the 90 m DEM. CLM-exp2 and SCLM-exp2 are for the f_{max} fields derived based on the 1 km DEM.







Fig. 5. Streamflow at different spatiotemporal scales.







Fig. 6. Model simulated rainfall/snowfall partitioning. The blue color is for the snow-dominated areas, yellow color is for the snow-rain mixing areas, and red color is for the rain-dominated areas.







Fig. 7. Seasonality of runoff at three sub-regions.





Fig. 8. Seasonality of forcing at three sub-regions.







Fig. 9. Seasonality of energy fluxes at three sub-regions.







Fig. 10. Impacts of f_{max} on runoff generation (Exp1 refers to the simulations based on f_{max} values derived from 90 m DEM; Exp2 refers to the simulations based on the f_{max} values derived from 1 km DEM).



