https://doi.org/10.5194/hess-2019-598
Preprint. Discussion started: 4 December 2019

Hydrology and Earth System Sciences
Discussions

Multi-variable, multi-configuration testing of ORCHIDEE land surface model water flux and storage estimates across semi-arid sites in the southwestern US

Natasha MacBean[1*], Russell L. Scott[2], Joel A. Biederman[2], Catherine Ottlé[3], Nicolas Vuichard[3], Agnès Ducharne[4], Thomas Kolb[5], Sabina Dore[6], Marcy Litvak[7], David J.P. Moore[8].

[1]Department of Geography, Indiana University, Bloomington, IN 47405, USA.
[2]Southwest Watershed Research Center, United States Agricultural Department, Agricultural Research Service, Tucson, AZ 85719, USA.
[3]Laboratoire des Sciences du Climat et de l'Environnement, LSCE/IPSL, CEA-CNRS-UVSQ, Université Paris-Saclay, Gif-sur-Yvette, F-91191, France.
[4]UMR METIS, Sorbonne Université, CNRS, EPHE, Paris, F-75005, France
[5]School of Forestry, Northern Arizona University, Flagstaff, AZ, 86011, USA.
[6]Hydrofocus, Inc., Davis, CA, 95618, USA.
[7]Department of Biology, University of New Mexico, Albuquerque, NM, 87131, USA.
[8]School of Natural Resources and the Environment, University of Arizona, Tucson, AZ, 85721, USA.

[*]*Correspondence to*: Natasha MacBean (nlmacbean@gmail.com)

Abstract

Plant activity in semi-arid ecosystems is largely controlled by pulses of precipitation, making them particularly vulnerable to increased aridity expected with climate change. Simple bucket-model hydrology schemes in land surface models (LSMs) have had limited ability in accurately capturing semi-arid water stores and fluxes. Recent, more complex, LSM hydrology models have not been widely evaluated against semi-arid ecosystem *in situ* data. We hypothesize that the failure of older LSM versions to represent evapotranspiration, ET, in arid lands is because simple bucket models do not capture realistic fluctuations in upper layer soil moisture. We therefore predict that including a discretized soil hydrology scheme based on a mechanistic description of moisture diffusion will result in an improvement in model ET when compared to data because the temporal variability of upper layer soil moisture content better corresponds to that of precipitation inputs. To test this prediction, we compared ORCHIDEE LSM simulations from 1) a simple conceptual 2-layer bucket scheme with fixed hydrological parameters; and 2) a 11-layer discretized mechanistic scheme of moisture diffusion in unsaturated soil based on Richards equations against daily and monthly soil moisture and ET observations, together with data-derived transpiration/ evaporation, T/ET, ratios, from six semi-arid grass, shrub and forest sites in the southwestern USA. The 11-layer scheme also has modified calculations of surface runoff, bare soil evaporation, and water limitation to be compatible with the more complex hydrology configuration. To

https://doi.org/10.5194/hess-2019-598
Preprint. Discussion started: 4 December 2019

Hydrology and Earth System Sciences
Discussions

diagnose remaining discrepancies in the 11-layer model, we tested two further configurations: i) the addition of a term that captures bare soil evaporation resistance to dry soil; and ii) reduced bare soil fraction. We found that the more mechanistic 11-layer model results better representation of the daily and monthly ET observations. We show that is likely because of improved simulation of soil moisture in the upper layers of soil (top 5cm). Some discrepancies between observed and modelled soil moisture and ET may allow us to prioritize future model development. Adding a soil resistance term generally decreased simulated E and increased soil moisture content, thus increasing T and T/ET ratios and reducing the negative T/ET model-data bias. By reducing the bare soil fraction in the model, we illustrated that modelled leaf T is too low at sparsely vegetated sites. We conclude that a discretized soil hydrology scheme and associated developments improves estimates of ET by allowing the model to more closely match the pulse precipitation dynamics of these semi-arid ecosystems; however, the partitioning of T from bare soil evaporation is not solved by this modification alone.

1 Introduction

Semi-arid ecosystems – which cover ~40% of the Earth's terrestrial surface and include rangelands, shrublands, grasslands, savannas, and seasonally dry forests – are in zones of transition between humid and arid climates and are characterized by sparse, patchy vegetation cover and limited water availability. Moisture availability in these ecosystems is therefore a major control on the complex interactions between vegetation dynamics and surface energy, water, and carbon exchange (Biederman et al., 2017; Haverd et al., 2016). Given the sensitivity to water availability, semi-arid ecosystem functioning may be particularly vulnerable to projected changes in climate (Tietjen et al., 2009; Maestre et al., 2012; Gremer et al., 2015). IPCC Earth System Model (ESM) projections and observation-based datasets indicate these regions will likely experience more intense warming and droughts, increases in extreme rainfall events, and a greater contrast between wet and dry seasons in the future (IPCC, 2013; Donat et al., 2016; Sippel et al., 2017; Huang et al., 2017).

To simulate the impact of climate change on semi-arid ecosystem functioning, it is essential that the land surface model (LSM) component of ESMs accurately represent semi-arid water flux and storage budgets (and all associated processes). In the last two to three decades, LSM groups have progressively updated their hydrology schemes from the more simplistic "bucket" type models included in earlier versions (Manabe, 1969). The resulting schemes typically include more physically-based representations of vertical diffusion of water in unsaturated soils (Clark et al., 2015). In addition to increasing the complexity of soil hydrology, several studies have attempted to address the issue that models tend to miscalculate partitioning of evapotranspiration (ET) into transpiration (T) and bare soil evaporation (E), with models systematically underestimating T/ET ratios (Wei et al., 2017; Chang et al., 2018). One such mechanism that models have introduced is a evaporation resistance term that reduces the rate of water evaporation from bare soil surfaces (Swenson and Lawrence, 2014; Decker et al., 2017). The development of these more mechanistic soil hydrology schemes should mean that LSMs better capture high temporal frequency to seasonal and long-term temporal variability of water stores and fluxes. However, it is not always apparent that increasing

https://doi.org/10.5194/hess-2019-598
Preprint. Discussion started: 4 December 2019


Hydrology and Earth System Sciences
Discussions

model complexity provides more accurate representations of reality (as encapsulated by observations of different variables at multiple spatio-temporal scales). Further, increasing model complexity comes at a cost of increased computational resources and unknown parameters. Therefore, it is imperative that we test models of increasing complexity against multiple types of observations at a variety of sites.

New generation LSM water flux and storage estimates have been extensively tested at multiple scales from the site level to the globe (Abramowitz et al., 2008; Dirmeyer et al., 2011; Guimberteau et al., 2014; Mueller et al., 2014; Best et al., 2015; Ukkola et al., 2016b; Raoult et al., 2018; Scanlon et al., 2018, 2019). Model-data biases are observed across all biomes; however, a key finding common to these studies is that models do not capture seasonal to inter-annual water stores and fluxes well during dry periods and/or at drier sites (Mueller et al., 2014; Swenson et al., 2014; De Kauwe et al., 2015; Best et al., 2015; Ukkola et al., 2016a; Humphrey et al., 2018; Scanlon et al., 2019). Mueller et al. (2014) showed that CMIP5 models overestimated multiyear mean daily ET in many regions, with the strongest bias in dryland regions (particularly western North America). Likewise, Grippa et al. (2011) and Scanlon et al. (2019) demonstrated that LSMs underestimate seasonal amplitude of total water storage in semi-arid (and tropical) regions. However, compared to more mesic ecosystems, semi-arid ecosystem LSM water flux and storage simulations have been rarely been tested extensively against *in situ* observations, apart from a few exceptions (Hogue et al., 2005; Abramowitz et al., 2008; Whitley et al., 2016). Whitley et al. (2016) compared carbon and water flux simulations from six LSMs at five OzFlux savanna sites. Their study highlighted two key deficiencies in modeling water fluxes: i) modeled C4 grass T is too low; and ii) models with shallow rooting depths typically underestimate woody plant dry season ET. As part of a model inter-comparison for West Africa (the AMMA LSM Intercomparison Project – ALMIP), LSM water storage, fluxes, runoff, and land surface temperature were evaluated against *in situ* and remote sensing data in the Malian Gourma region of the central Sahel (Boone et al., 2009; De Kauwe et al., 2013; Lohou et al., 2014; Grippa et al., 2011; Grippa et al., 2017). These studies highlight that temporal characteristics of water storage and fluxes in this monsoon-driven semi-arid region are captured fairly well by models; however, the studies also point to various model issues, including: difficulties in simulating bare soil evaporation response to rainfall events (Lohou et al., 2014); underestimation of dry season ET (Grippa et al., 2011); the need for greater water and energy exchange sensitivity to different vegetation types and soil characteristics (De Kauwe et al., 2013; Lohou et al., 2014; Grippa et at., 2017); and overestimation of surface runoff (Grippa et al., 2017). How models prescribe or predict leaf area index (LAI) has also been highlighted as a driver of hydrological model-data differences (Ha et al., 2015; Grippa et al., 2017).

The aim of this study was to contribute a new LSM hydrology model evaluation in a semi-arid region not previously investigated: the monsoon-driven semi-arid southwestern United States (hereafter, the SW US). The density and diversity of ecosystem research sites in the SW US provides a rare opportunity to test an LSM across a range of semi-arid ecosystems. The semi-arid SW US has also been identified as one of the key regions of global land-atmosphere coupling (Koster et al., 2004) and the most persistent climate change hotspot in the US (Diffenbaugh et al., 2008; Allen et al., 2016). Expected future soil moisture deficits in this region will result in strong atmospheric feedbacks with consequent high temperature increases (Senerivatne et al., 2013) and a potential weakening of the terrestrial biosphere C sink (Berg et al., 2016; Green et al., 2019).

https://doi.org/10.5194/hess-2019-598
Preprint. Discussion started: 4 December 2019


Hydrology and Earth System Sciences
Discussions

Several studies based on model predictions, instrumental records, and paleoclimatic data analyses have suggested that over the coming century the risk of more severe, multi-decadal drought in the SW US will increase considerably (Ault et al., 2014, 2016; Cook et al., 2015). In fact, models suggest that a transition to drier conditions is already underway (Seager et al., 2007; Archer and Predick, 2008; Seager and Vecchi, 2010). Investigating how well LSMs capture hydrological stores and fluxes in this region therefore provides a crucial test for how well models can produce accurate global climate change projections.

Here, we tested the ability of the ORCHIDEE (ORganizing Carbon and Hydrology in Dynamic EcosystEms) LSM to simulate multiple water flux and storage related variables at six SW US semi-arid Ameriflux eddy covariance sites spanning forest, shrub- and grass-dominated ecosystems (Biederman et al., 2017). We tested two versions of the ORCHIDEE LSM with hydrological schemes of differing complexity: 1) a simple 2-layer conceptual bucket scheme (hereafter, 2LAY) with constant water holding capacity (de Rosnay and Polcher 1998); and 2) a 11-layer mechanistic scheme (hereafter, 11LAY) based on the combination of Richards and Darcy's equations with hydraulic parameters based on soil texture (de Rosnay et al., 2002). Besides the change in the soil hydrology between the 2LAY and 11LAY versions, several other hydrology-related processes have also been modified due to increases in the complexity of the model. In the 2LAY scheme, runoff occurred when the soil reached saturation; whereas in the 11LAY scheme, surface runoff and drainage are treated more mechanistically (see Section 2.2.2). In the 2LAY scheme, there was an implicit resistance to bare soil evaporation based on the depth of the dry soil for the bare soil PFT. In the 11LAY scheme, there is an optional bare soil evaporation resistance term based on the relative soil water content of the first four soil layers, based on the formulation of Sellers et al. (1992) – (see Section 2.2.3). Both resistance terms aim to describe the resistance to evaporation exerted by a dry mulch soil layer. Similarly, the calculation of moisture limitation on stomatal conductance has changed. In the 2LAY version, moisture limitation depended on the dry soil depth of the upper layer; whereas in the 11LAY version, the limitation is based on plant water availability for root water uptake throughout the soil column. The 2LAY scheme was used in the previous CMIP5 runs that are still being used to understand ecosystem responses to changes in climate, whereas the 11LAY scheme is the default scheme in the current version of ORCHIDEE that is used in the ongoing Coupled Model Intercomparison Project (CMIP6) simulations (Ducharne et al., in prep).

In testing the ORCHIDEE against *in situ* semi-arid water stores and fluxes, a novel component of our study was to investigate whether some of the site-scale semi-arid LSM hydrology model discrepancies outlined above (e.g. underestimation of C4 grass T, weak dry season ET (and therefore low T/ET ratios), ET issues related to incorrect representation of leaf seasonality, and overestimation of surface runoff) are improved with recent ORCHIDEE model developments. Where the model does not capture observed patterns, we attempted to investigate which model processes and parameterizations are responsible for model-data discrepancies. First, we evaluated how changing from the conceptual 2LAY bucket model to the physically-based 11LAY soil hydrology scheme – and all associated modifications – influences the high temporal frequency and seasonal variability of semi-arid ecosystem soil moisture, ET (and its component fluxes), runoff, drainage, and snow mass/melt. Second, we evaluated the temporal dynamics of the 11LAY model against observations at three specific soil depths (shallow: ≤ 5cm; mid: 15-20cm; deep: ≥ 30cm) to assess whether the physically-based discretized scheme accurately captures moisture transport down the soil profile. Note that when evaluating the 11LAY model soil moisture against observations, our primary focus was on the temporal

https://doi.org/10.5194/hess-2019-598
Preprint. Discussion started: 4 December 2019


Hydrology and Earth System Sciences
Discussions

dynamics – rather than the absolute magnitude – given the difficulty of comparing absolute values of volumetric water content between the models and the data (see Section 2.3.2 for more detail). Therefore, in the model-data comparison, we scale the observations to the 11LAY model simulations via linear CDF matching. Finally, having evaluated the standard (default) 11LAY model, we investigated which model processes or mechanisms in the 11LAY scheme might be responsible for remaining model-data discrepancies in the abovementioned hydrological variables. In particular, we assessed the impact of a) decreasing the bare soil fraction; and b) including the optional bare soil resistance term into the 11LAY scheme (Ducharne et al., in prep.). Given the sparsely-vegetated nature of the low-elevation semi-arid grass- and shrub-dominated sites in our study it is possible that inclusion of this term may counter any dry season ET underestimate. Throughout, we explored if there are any discernible differences across sites due to elevation, soil type, or vegetation composition. Section 2 described the sites, data, models and methods used in this study; Section 3 details the results of the two-part model evaluation (as outlined above); and Section 4 discusses how future studies may resolve remaining model issues in order to improve LSM hydrology modeling in semi-arid regions.

2 Methods and Data

2.1 Southwestern US study sites

We used six semi-arid sites in the SW US that spanned a range of vegetation types and elevations (Biederman et al. 2017). The entire SW US is within the North American Monsoon region; therefore, these sites typically experience monsoon rainfall during July to October, preceded by a hot, dry period in May and June. Table 1 describes the dominant vegetation, species and soil texture characteristics at each site, together with the observation period. The four grass- and shrub-dominated sites (US-SRG, US-SRM, US-Whs and US-Wkg) are located at low-elevation (<1600m) in southern Arizona. These four sites are split into pairs of grass- and shrub-dominated systems: US-SRG and US-SRM are located at the Santa Rita Experimental Range ~60km south of Tucson, AZ, whilst US-Whs and US-Wkg are located at the Walnut Gulch Experimental Watershed ~120km to the southeast of Tucson, AZ. The US-Fuf (Flagstaff Unmanaged Forest) and US-Vcp (Valles Caldera Ponderosa) sites are at a higher elevation (2215m and 2501m, respectively) and are dominated by ponderosa pine. US-Fuf is located near the town of Flagstaff in northern AZ, whilst US-Vcp is located in the Valles Caldera National Preserve in the Jemez Mountains in north-central New Mexico. Flux tower instruments at all six sites collect half-hourly measurements of meteorological forcing data and eddy covariance measurements of net surface energy and carbon exchanges.

https://doi.org/10.5194/hess-2019-598
Preprint. Discussion started: 4 December 2019


Hydrology and Earth System Sciences
Discussions

2.2 ORCHIDEE Land Surface Model

2.2.1 General model description

The ORCHIDEE land surface model (LSM) forms the terrestrial component of the French IPSL ESM (Dufresne et al., 2013), which contributes climate projections to IPCC Assessment Reports. ORCHIDEE has undergone significant modification since the 'AR5' version (Krinner et al., 2005), which was used to run the CMIP5 (Coupled Model Inter-comparison Project) simulations included in the IPCC 5th Assessment Report (IPCC, 2013). Here, we use ORCHIDEE v2.0 that is used in the ongoing CMIP6 simulations (Peylin et al., in prep).

ORCHIDEE simulates fluxes of carbon, water and energy between the atmosphere and land surface (and within the sub-surface) on a half-hourly time step. In uncoupled mode, the model is forced with climatological fields derived either from climate reanalyses or site-based meteorological forcing data (2m air temperature, rainfall and snowfall, incoming long and shortwave radiation, wind speed, surface air pressure, and specific humidity). As in most LSMs, all vegetation is grouped into broad plant functional types (PFTs) based on physiology, phenology, and for trees, the biome in which they are located. In ORCHIDEE, by default, there are 12 vegetated PFTs plus a bare soil PFT. A prognostic leaf area is calculated based on phenology schemes originally described in Botta et al. (2000) and further detailed in MacBean et al. (2015 – Appendix A).

The albedo is calculated based on the average of the defined albedo coefficients for vegetation (one coefficient per PFT), soil (one value for each grid cell, referred to as background albedo) and snow weighted by their fractional cover. Snow albedo is also parameterized according to its age, which varies according to the underlying PFT. The albedo coefficients for each PFT and background albedo have recently been optimized within a Bayesian inversion system using the visible and near infrared MODIS white sky albedo product at 0.5x0.5° resolution for years 2000-2010 (Bastrikov et al., in prep). The prior background (bare soil) albedo values were retrieved from the MODIS data using the EU Joint Research Center Two Stream Inversion Package (JRC-TIP).

Evapotranspiration, ET, in the model is calculated as the sum of four components: 1) evaporation from bare soil; 2) evaporation, E, from water intercepted by the canopy; 3) transpiration, T, (controlled by stomatal conductance); and 4) snow sublimation (Guimberteau et al., 2012b). There are two soil hydrology models implemented in ORCHIDEE: one based on a 2-layer conceptual model, the other on a physically based representation of moisture redistribution across 11-layers. These two schemes were described and compared in the Amazon basin by Guimberteau et al. (2014), and are described further in the following sections. In this study, the depth of the soil for both schemes is set to 2m. Separate water budgets are calculated for each soil tile within a grid cell. In the 2-layer scheme, soil tiles correspond to PFTs; therefore, a separate water budget is calculated for each PFT within the grid cell. In the 11-layer scheme there are three soil tiles: one gathering all tree PFTs, one gathering grasses and crops, and the third as bare soil. The grid cell water budget is calculated by vegetation fraction weighted averaging across all soil tiles (Guimberteau et al., 2014). Soil texture classes and related parameters) are prescribed based on the percentage of sand, clay and loam.

https://doi.org/10.5194/hess-2019-598
Preprint. Discussion started: 4 December 2019


Hydrology and Earth System Sciences
Discussions

2.2.2 Soil Hydrology

2-layer conceptual soil hydrology model

In the 'AR5' version of ORCHIDEE used in the CMIP5 experiments, the soil hydrology scheme consisted of a conceptual 2-layer (hereafter, 2LAY) model based on Choisnel et al. (1995). The depth of the upper layer is variable up to 10 cm and changes with time depending on the balance between throughfall and snowmelt inputs, and outputs via three pathways: i) bare soil evaporation, limited by a soil resistance increasing with the dryness of the topmost soil layer; ii) root water extraction for transpiration, withdrawn from both layers proportionally to the root density profile; and iii) downward water flow (drainage) to the lower layer. If all moisture evaporates or if the entire soil saturates, the top layer can disappear entirely. Three empirical parameters govern the calculation of the drainage between the two layers, which depends on the water content of the upper layer and takes a non-linear form, so drainage from the upper layer increases considerably when the water content of the upper layer exceeds 75% of the maximum capacity (Ducharne et al., 1998). Transpiration is also withdrawn from the lower layer via water uptake by deep roots. Finally, runoff only occurs when the total soil water content exceeds the maximum capacity, set to 150 $kg.m^{-2}$ as in Manabe (1969). It is then arbitrarily partitioned into 5% surface runoff to feed the overland flow and 95% drainage to feed the groundwater flow of the routing scheme (Guimberteau et al., 2012b), which is not activated here.

11-layer mechanistic soil hydrology model

The 11LAY scheme was initially proposed by de Rosnay at al. (2002) and simulates vertical flow and retention of water in unsaturated soils based on a physical description of moisture diffusion (Richards, 1931). The scheme implemented in ORCHIDEE relies on a one-dimensional differential equation, combining the one-dimensional Darcy equation, which describes the rate of flow of a fluid (soil moisture) within a permeable medium, with the mass conservation equation (the complete formulation is more generally called the saturation-based Richards equation). The two main hydraulic parameters (hydraulic conductivity and diffusivity) depend on volumetric soil moisture content defined by the Mualem–van Genuchten model (Mualem, 1976; van Genuchten, 1980). The Richards equation is solved numerically using a finite-difference method, which requires the vertical discretisation of the 2m soil column. As described by de Rosnay at al. (2002), 11 layers are defined: the top layer is ~0.1mm thick and the thickness of each layer increases geometrically with depth. The fine vertical resolution near the surface aims at capturing strong vertical soil moisture gradients in response to high temporal frequency (sub-diurnal to few days) changes in precipitation or ET. The mechanistic representation of redistribution of moisture within the soil column also permits capillary rise, and a more mechanistic representation of surface runoff: The calculated soil hydraulic conductivity determines how much precipitation is partitioned between soil infiltration and runoff (d'Orgeval et al., 2008). Drainage is computed as free gravitational flow at the bottom of the soil (Guimberteau et al., 2014). The USDA soil texture classification, provided at 1/12-degree resolution by Reynolds et al. (2000), is combined with the look-up tables of Carsel and Parrish (1988) to derive the required soil hydrodynamic properties (saturated hydraulic conductivity Ks, porosity, Van Genuchten parameters,

https://doi.org/10.5194/hess-2019-598
Preprint. Discussion started: 4 December 2019


Hydrology and Earth System Sciences
Discussions

residual moisture), while field capacity and wilting point are deduced from the soil hydrodynamic properties listed above and the Van Genuchten equation for matric potential, by assuming they correspond to potentials of -3.3m and -150m respectively. The 11LAY soil hydrology scheme has been implemented in the ORCHIDEE Trunk since 2010, albeit with various modifications since as described in this study. Similar versions of the 11LAY scheme has been tested against a variety of hydrology-related observations in the Amazon Basin (Guimberteau et al., 2012a; Guimberteau et al., 2014), for predicting future changes in extreme runoff events (Guimberteau et al., 2013) and against a water storage and energy flux estimates as part of ALMIP in West Africa (as detailed in Section 1 – d'Orgeval et al., 2008; Boone et al., 2009; Grippa et al., 2011; Grippa et al., 2017).

2.2.3 Bare soil evaporation and additional resistance term

The computation of bare soil evaporation, E_g, in both versions is implicitly based on a supply and demand scheme. In the 2LAY version, E_g decreases when the upper layer gets drier, owing to a resistance term that depends on the height of the dry soil in the bare soil PFT column (Ducoudré et al., 1993). In the 11LAY version, E_g, proceeds at the potential rate E_{pot} unless the water supply via upward diffusion from the water column is limiting, in which case E_g is reduced to correspond to the situation in which the soil moisture of the upper 4 layers is at wilting point. However, since ORCHIDEE v2.0 (Ducharne et al., in prep.), E_g can also be reduced by including an optional bare soil evaporation resistance term, r_{soil}, which depends on the relative water content and is based on a parameterization fitted at the FIFE grassland experimental site at Konza Prairie Field Station in Kansas (Sellers et al., 1992):

$$r_{\mathrm{soil}} = \exp(8.206 - 4.255\, W_1) \quad (1)$$

where W_1 is the relative soil water content of the first four layers (2.2cm – Table S1). W_1 is calculated by dividing the mean soil moisture across these layers by the saturated water content. The calculation for E_g then becomes:

$$E_g = \min(E_{\mathrm{pot}}/(1 + r_{\mathrm{soil}}/r_a),\, Q) \quad (2)$$

where E_{pot} is the potential evaporation, r_a the aerodynamic resistance, Q the upward water supply from capillary diffusion through the soil, and r_{soil} the soil resistance to this upward exfiltration. In all simulations, the calculation of r_a includes a dynamic roughness height with variable LAI, based on a parameterization by Su et al. (2001). By default, in the 11LAY version there is no resistance ($r_{soil} = 0$).

2.2.4 Empirical plant water stress function, β

The soil moisture control on transpiration is defined by an empirical water stress function, called β. Whichever the soil hydrology model, it depends on soil moisture and on the root density profile $R(z) = \exp(-c_j z)$, where z is the soil depth and c_j (in m^{-1}) is the the root density decay factor. For a 2m soil profile, c_j is set to 4.0 for grasses, 1.0 for temperate needleleaved

https://doi.org/10.5194/hess-2019-598
Preprint. Discussion started: 4 December 2019


Hydrology and Earth System Sciences
Discussions

trees and 0.8 for temperate broadleaved trees. In 11 LAY, a related variable is $n_{root}(i)$, quantifying the mean relative root density of each soil layer *i*, so that $\sum n_{root}(i) = 1$.

In the 2LAY version, β is calculated as an exponential function of the root decay factor c_j and the dry soil height of the topmost soil layer (h_t^d):

$$\beta = exp^{(-c_j, h_t^d)} \quad (3)$$

In the 11LAY, β is rather based on the full soil moisture profile and is calculated for each PFT *v* and soil layer *i,* and then summed across all soil layers:

$$\beta = \sum_{i=2}^{11} n_{root}(i, v) \min\left(1, \max\left(0, \frac{(W_{i,v} - W_{wpt})}{(W_{\%} - W_{wpt})}\right)\right) \quad (4)$$

where *v* is the PFT index (starting at 2 given the 1st PFT is bare soil), *i* is the index for each soil layer (starting at the 2nd layer given no water stress in the 1st layer – a conservative condition that prevents T from inducing a negative soil moisture from this very thin soil layer), W_i is the soil moisture for that layer and soil tile in kgm^{-2}, W_{wpt} is the wilting point soil moisture, and $W_{\%}$ is the threshold above which T is maximum – i.e. above this threshold T is not limited by β. $W_{\%}$ is defined by:

$$W_{\%} = W_{wpt} + p_{\%}(W_{fc} - W_{wpt})$$

where $p_{\%}$ defines the threshold above which T is maximum. $p_{\%}$ is set to 0.8 and is constant for all PFTs.

This empirical water stress function equation means that, in the 11LAY, β varies linearly between 0 at the wilting point to 1 at $W_{\%}$, which is smaller or equal to the field capacity. LSMs typically apply β to limit photosynthesis (A) via the maximum carboxylation capacity parameter V_{cmax}, or to the stomatal conductance, g_s, via the g0 or g1 parameters of the A/gs relationship, or both (De Kauwe et al., 2013; 2015). In ORCHIDEE there is the option of applying β to limit either V_{cmax} or gs, or both. In the default configuration used in CMIP6, β is applied to both (based on results from Keenan et al., 2010; Zhou et al., 2013; Zhou et al., 2014); therefore, this is the configuration we use in this study.

2.2.5 Snow scheme

ORCHIDEE contains a multi-layer intermediate complexity snow scheme that is described in detail in Wang et al. (2013). The new scheme was introduced to overcome limitations of a single layer snow configuration. In a single layer scheme, the temperature and vertical density gradients through the snowpack, which affect the sensible, latent and radiative energy fluxes, are not calculated. The single layer snow scheme does not describe the insulating effect of the snow pack, nor the links between snow density and changes in snow albedo (due to aging) in a physically mechanistic way. In this explicit snow scheme, there are three layers that each have a specific thickness, density, temperature and liquid water and heat content: These variables are updated at each time step based on the snowfall and incoming surface energy fluxes, which are calculated from the surface energy balance equation. The model also accounts for sublimation, snow settling, water percolation and refreezing. Snow mass

https://doi.org/10.5194/hess-2019-598
Preprint. Discussion started: 4 December 2019


Hydrology and Earth System Sciences
Discussions

cannot exceed a threshold of 3000 kg.m^{-2}. Snow age is also calculated and is used to modify the snow albedo. Default snow albedo coefficients have been optimized using MODIS data per the method described in Section 2.2.1. Snow fraction is calculated at each time step according to snow mass and density following the parametrization proposed by Niu and Yang (2007).

2.3 Data

2.3.1 Site-level meteorological and eddy covariance data and processing

Meteorological forcing and eddy covariance flux data for each site were downloaded from the AmeriFlux data portal (http://ameriflux.lbl.gov). Meteorological forcing data included 2m air temperature and surface pressure, precipitation, incoming long and shortwave radiation, wind speed, and specific humidity. To run the ORCHIDEE model, we partitioned the *in situ* precipitation into rain and snowfall using a temperature threshold of 0°C. The meteorological forcing data were gap-filled following the approach of Vuichard and Papale (2015), which uses downscaled and corrected ERA-Interim data to fill gaps in the site-level data. Eddy covariance flux data were processed to provide ET from estimates of latent energy fluxes. ET gaps were filled using a modified look-up table approach based on Falge et al. (2001), with ET predicted from meteorological conditions within a 5-day moving window. Previous comparisons of annual sums of measured ET with site-level water balance measurements at a few of these sites show an average agreement within 3% of each other, but could differ by -10 to +17% in any given year (Scott and Biederman, 2019). T/ET ratio estimates were derived from Scott and Biederman (2017), who developed a new method to estimate average monthly T/ET during the summer growing season at the non-forested sites using eddy covariance fluxes.

2.3.2 Soil moisture data and processing

Daily mean volumetric soil moisture content (VWC, m^3m^{-3}) measurements at several depths were obtained directly from the site PIs. For each site, Table 2 details the depths at which soil moisture was measured. Soil moisture measurement uncertainty is highly site and instrument specific, but tests have shown that average errors are generally below 0.04 m^3 m^{-3} if site specific calibrations are made. We only used these measurements to evaluate the 11LAY model because, unlike the 2LAY model, with this version of the model we have model estimates of soil moisture at discrete soil depths. However, several factors mean that we cannot directly compare absolute values of measured versus modelled soil VWC, even though the 11LAY has discrete depths. First, site-specific values for soil saturated and residual water content were not available to parameterize the model; instead, these soil hydrology parameters are either fixed (in the 2LAY) or derived from prescribed soil texture properties (in the 11LAY – see Section 2.2.2). Therefore, we may expect a bias between the modelled and observed daily mean volumetric soil water content (θ). Second, while the soil moisture measurements are made with probes at specific depths, it is not precisely

https://doi.org/10.5194/hess-2019-598
Preprint. Discussion started: 4 December 2019


Hydrology and Earth System Sciences
Discussions

known over which depth ranges they are measuring VWC. Therefore, instead of focusing on absolute soil moisture values in the 11LAY model – data comparison, we specifically investigate how well the model captured the temporal dynamics at specific soil depths. To achieve this, we removed any model-data bias using a linear cumulative density function (CDF) matching function to re-scale and match the mean and standard deviation of soil moisture observations to that of the model for each layer where soil moisture is measured following the equation:

$$\theta_{Obs,CDF} = \frac{\sigma_{\theta,Mod}(\theta_{Obs}-\bar{\theta}_{Obs})}{\sigma_{\theta,Obs}} + \bar{\theta}_{Mod} \quad (5)$$

Raoult et al. (2018) found that linear CDF matching performed nearly as well as full CDF matching in capturing the main features of the soil moisture distributions; therefore, for this study we chose to simply use a linear CDF re-scaling function.

2.4 Simulation set-up and post-processing

All simulations were run for the period of available site data (including meteorological forcing and eddy covariance flux data – see Section 2.3.1 and Table 1). Table 1 also lists: i) the main species for each site and the fractional cover of each model PFT that corresponds to those species; iii) the maximum LAI for each PFT; and iii) the percent of each model soil texture class that corresponds to descriptions of soil characteristics for each site – all of which were derived from associated site literature detailed in references in Tables 1. The PFT fractional cover and the fraction of each soil texture class are defined in ORCHIDEE by the user. The maximum LAI has a default setting in ORCHIDEE that has not been used here; instead, values based on the site literature have been prescribed in the model (Table 1). Note that ORCHIDEE does not contain a PFT that specifically corresponds to shrub vegetation; therefore, the shrub cover fraction was prescribed to the forested PFTs (see Table 1).
At each site we ran five versions of the model: 1) 2LAY soil hydrology; 2) 11LAY soil hydrology with r_{soil} flag not set; 3) 11LAY soil hydrology with r_{soil} flag not set and with reduced bare soil fraction (increased C4 grass cover); 4) 11LAY soil hydrology with the r_{soil} flag set (therefore, Eqn. 2 activated) and 5) 11LAY soil hydrology with the r_{soil} flag set and with reduced bare soil fraction. A 400-year spinup was performed by cycling over the gap-filled forcing data for each site (see Table 1 for period of available site data) to ensure the water stores were at equilibrium. Following the spinup, transient simulations were run using the forcing data from each site. Daily outputs of all hydrological variables (soil moisture, ET and its component fluxes, snow pack, snow melt), the empirical water stress function, β, and LAI were saved for all years and summed or averaged to derive monthly values, where needed. For certain figures we show the 2009 daily time series because 2009 was the only year for which data from all sites overlapped and had a complete year of daily soil moisture observations.

3 Results

3.1 Differences between the 2LAY and 11LAY model versions for main hydrological stores and fluxes

Increasing the soil hydrology model complexity between the 2LAY and 11LAY model versions does not result in a uniform increase or decrease across sites in either the simulated upper layer (top 10cm) and total column (2m) soil moisture amount

https://doi.org/10.5194/hess-2019-598
Preprint. Discussion started: 4 December 2019


Hydrology and Earth System Sciences
Discussions

(kgm^{-2}) (Fig. 1 2nd and 3rd panel; also see Fig. S1 for complete daily time series for each site). The largest change between the 2LAY and 11LAY versions in the upper layer soil moisture are seen at the high-elevation ponderosa forest sites (US-Fuf and US-Vcp – Fig. 1 and Figs. S1 a and b). In the 2LAY simulations, the upper layer soil moisture is higher at the ponderosa sites; whereas, in the 11LAY simulations the difference between sites is reduced. At US-Fuf, both the upper layer and total column soil moisture increase in the 11LAY simulations compared to the 2LAY, which corresponds to an increase in ET (Fig. 1 top panel) and a decrease in total runoff (surface runoff plus drainage – Fig. 1 bottom panel). In contrast, at US-Vcp while there is an increase in the upper layer soil moisture, which results in a slight increase in mean ET (and ET variability) and a decrease in total runoff, there is hardly any change in the total column soil moisture. Note that changes in maximum soil water holding capacity are due to how soil hydrology parameters are defined. In the 2LAY, a maximum capacity is set to 150kgm^{-2} across all PFTs; whereas in the 11LAY, the capacity is based on soil texture properties and is therefore different for each site.

At the low-elevation shrub and grass sites the differences between the two model versions for both the upper layer and total column soil moisture are much smaller (Fig. 1). Correspondingly, the changes in mean ET and total runoff are also marginal (although the mean total runoff is lower at Walnut Gulch: US-Wkg and US-Whs). At Santa Rita (US-SRM and US-SRG), the 11LAY soil moisture is marginally lower than the 2LAY, whereas at the Walnut Gulch sites the 11LAY moisture is higher.

As described above, at all sites between the 2LAY and 11LAY simulations there is either no change (Santa Rita) or a decrease in total runoff (surface runoff plus drainage – Fig. 1 bottom panel). Across all sites, excess water is removed as drainage in the 2LAY simulations, with little to no runoff (Figs. S1 a-f 3rd panel); whereas in the 11LAY simulations excess water flows mostly as surface runoff, with more limited drainage (Figs. S1 a-f 2nd panel). This is explained by the fact that in the 2LAY scheme, the drainage is always set to 95% of the soil excess water (above saturation) and runoff can appear only when the total 2m soil is saturated. However, the 11LAY scheme also accounts for runoff that exceeds the infiltration capacity, which depends on the hydraulic conductivity function of soil moisture (Horton runoff).. This means that when the soil is dry, the conductivity is low and more runoff will be generated. In the 11LAY simulations, the temporal variability in total runoff (as represented by the error bars in Fig. 1) has also decreased. As just described, in the 11LAY the total runoff mostly corresponds to surface runoff (Figs. S1 a-f). The lower drainage flux in the 11LAY simulations corresponds well to the calculated water balance at US-SRM (Scott and Biederman, 2019). The 11LAY limited drainage is also likely to be the case at US-Fuf given that nearly all precipitation at the site is partitioned to ET (Dore et al., 2012).

Across all sites, the magnitude of the total column VWC temporal variability (Fig. 1 error bars) only increases slightly between the 2LAY and 11LAY model versions; although, the complete time series in Fig. S1 show that at all sites the 11LAY simulations appear to respond more dynamically to rain events. In the upper layer (top 10cm), the VWC temporal variability increases marginally between the 2LAY and 11LAY for the high-elevation forest sites; however, the magnitude of variability decreases considerably in the 11LAY model for the low-elevation shrub and grass sites (also see Fig. S2). Whereas the 2LAY upper layer soil moisture simulations at all sites fluctuate considerably between field capacity and zero throughout the year – including during dry periods with no rain – the temporal dynamics of the 11LAY upper layer moisture simulations correspond

https://doi.org/10.5194/hess-2019-598
Preprint. Discussion started: 4 December 2019


Hydrology and Earth System Sciences
Discussions

more directly to the timing of rainfall events (see Fig. 2 bottom panel for an example at 3 sites in 2009 and Fig. S2 for the complete time series for each site). Note that not only is the upper high frequency temporal variability therefore arguably more realistic in the 11LAY version, the finer scale discretization of the uppermost soil layer in this version will also allow a much easier comparison with satellite-derived soil moisture products that can only "sense" the upper few cm of the soil (Raoult et al., 2018).

A major and important consequence of the changes in the upper layer soil moisture high frequency temporal variability is a considerable improvement across all sites in the 11LAY simulated ET high frequency temporal dynamics (Fig. 2 2nd panel, which shows 2009 for three sites; Figs. S2 a-f shows the complete time series for all sites). As previously discussed, the increase in 11LAY model upper layer moisture content at the high-elevation forest sites (Fig. 1 2nd panel and Fig. 2 bottom panel) have resulted in an increase in E and T at those sites, which in turn results in a better match between the model and the observations (Figs. 2 and S2 2nd panel). At the low-elevation shrub and grass sites, the improvement in ET is also related to changes between the two versions in the calculation of the empirical water stress function, β (Figs. 2 and S2 5th panel), which acts to limit both photosynthesis and stomatal conductance (therefore, T) during periods of moisture stress (Section 2.2.4). With the new calculation in the latest version, we see a stronger, more rapid decrease in β (increased stress) during warm, dry periods that correspond to strong reductions in T (brown shaded zones in Fig. 2). Aside from T and E, the other ET components (interception and sublimation) are not changing much between the two hydrology schemes (results not shown); therefore, these terms are not contributing to improvements between the 2LAY and 11LAY versions.

The improvement in ET at high frequency timescales gives rise to a dramatic reduction in model-data misfit at monthly to seasonal timescales, particularly for low-elevation grass and shrub sites. Across all sites, the 11 LAY RMSE between daily modelled and observed ET has decreased in comparison to the 2LAY and the correlation has increased by a fraction of 0.3 to 0.4 (Table 3). With the exception of US-Vcp, the mean absolute daily model-observed ET bias has increased slightly between the 2LAY and 11LAY versions (Table 3), which is due to the fact that the 2LAY version both underestimates and overestimates ET in the spring and summer respectively (Fig. S3). However, the mean absolute biases are only a small fraction of the daily ET. The 11LAY model only slightly underestimates mean daily ET at most sites, except at US-Fuf.

The improvement in daily ET results in mean 11LAY monthly ET that is also well captured by the model throughout the year, including both the warm, dry May-June period followed by monsoon summer rains (Fig. 3 and Fig. S3). The improved, higher monthly ET in the 11LAY during the period of maximum productivity (i.e. the spring and summer for the high-elevation sites, and the summer monsoon for the low-elevation sites – Fig. 3) is likely due to the increase in plant available water (Fig. 1 – 2nd and 3rd panels and Fig. S1). Despite the improvement in the 11LAY temporal variability at the high-elevation forest sites, there is still a bias in the mean monthly ET magnitude between the 11LAY model and observations: At US-Fuf there is a distinct overestimation of ET during the spring (Fig. S3a), whereas at US-Vcp there is a noticeable underestimation of ET during the spring and monsoon periods (Fig. S3b). We will return to these remaining 11LAY ET model-data discrepancies in Section 3.3 after having evaluated the 11LAY soil moisture against observations.

https://doi.org/10.5194/hess-2019-598
Preprint. Discussion started: 4 December 2019


Hydrology and Earth System Sciences
Discussions

415

3.2 Comparison of 11LAY soil moisture against observations at different depths

Fig. 4 compares model versus observed daily volumetric soil water content time-series for 2009 at three different depths (see Fig. S4 for full time series at each site). The observations across the whole time series were re-scaled via linear CDF matching to remove model–observation biases (see Section 2.3.2); however, the linear CDF matching preserves the mean and standard
420 deviation of the temporal variability. As seen in Section 3.1 and Fig. 2 (bottom panels showing upper 10cm soil moisture), in Fig 4 the high frequency temporal variability of the 11LAY soil moisture in the uppermost layer almost perfectly matches the observed, particularly at the low-elevation shrub- and grass-dominated sites. At most of the low-elevation sites the soil moisture drying rates in the upper 20cm of soil are well captured by the model, with the small exception of the Santa Rita sites between January to March in which the model appears to dry down at a faster rate than observed (Fig. 4). In contrast, the
425 temporal mismatch between the observations and the model in the uppermost layer is higher at the forest sites (particularly US-Vcp). The US-Fuf and US-Vcp 11LAY simulations appear to compare reasonably well with observations in the upper 2cm of the soil from June through to the end of November (end of September in the case of US-Vcp) (Fig. 4). However, even when comparing with the re-scaled observations, at both sites the model appears to overestimate the VWC during the winter months, and underestimate the observed VWC during the spring months. Although US-Fuf and US-Vcp are semi-arid sites, their high-
430 elevation means that during winter precipitation falls as snow; therefore, this large apparent model bias may be related to either the ORCHIDEE snow scheme or incorrect snowfall meteorological forcing. During the winter period of peak snow fall the model soil moisture increases rapidly as the snow pack melts and is replenished by new snowfall, whereas the observed soil moisture response is slower (Fig. 5). Similarly, after the main period of snowfall the model again melts and sublimates snow quite rapidly; once all the snow has melted, the model soil moisture also declines. However, the observed soil moisture remains
435 high throughout the spring (Fig. 5).

To test the hypothesis that the model melts or sublimates snow too rapidly, we altered the model to artificially increase snow albedo and decrease the amount of sublimation; however, these tests had little impact on the rate of snow melt or the duration of snow cover (results not shown). Aside from model structural or parametric error, it is possible that there is an error in the meteorological forcing data. Although snow depth, mass or fractional cover data have not been collected at these sites, snow
440 typically remains on the ground until late spring after winters with heavy snowfall (T. Kolb pers. comm.) suggesting that the continued existence of a snow pack and snow melt replenishes soil moisture until late spring when all the snow melts. Therefore, it is important to note that the rain gauges almost certainly underestimate the actual snowfall amount (Rasmussen et al., 2012; Chubb et al., 2015) and/or that the method of assigning precipitation data as snowfall when $T < 0°C$ (Section 2.3.1) results in inaccurate estimates of the amount or duration of snow cover. To test this hypothesis, we artificially increased
445 the meteorological forcing snowfall amount. The result was an improvement in the modelled springtime soil moisture estimates, suggesting that inaccurate snowfall forcing estimates may be the main factor causing high-elevation spring soil VWC model-data misfits.

https://doi.org/10.5194/hess-2019-598
Preprint. Discussion started: 4 December 2019


Hydrology and Earth System Sciences
Discussions

Overall, there is a decrease in the model ability to capture both high frequency and seasonal variability with increasing soil depth. At all sites the temporal dynamics of the deepest observations are not well represented in the model (Fig. 4 bottom panels for each site). At the high-elevation forest sites (US-Fuf and US-Vcp), the model does not capture the response of observed soil moisture at the deepest layer to summer storm events. In contrast, at the low-elevation shrub and grass sites the 11LAY VWC is far too dynamic in the deepest layer. The poor model-data fit at lower depths may be related to the discretization of the soil column with a geometric increase of internode distance. Therefore, the soil layer thicknesses which increases substantially beyond the ~2-4cm (7^{th} and 8^{th} soil layers – Table S1). For the deeper soil moisture observations, it is therefore harder to match the depth of the observations with a specific soil layer. Alternatively, it is possible that soil texture changes with depth are altering hydraulic conductivity parameters; in the model however, hydraulic conductivity only changes exponentially with depth owing to soil compaction (Ducharne et al., in prep.).

3.3 Remaining discrepancies in ET and its component fluxes

Despite the improvement in seasonal ET temporal dynamics in the 11LAY model, particularly the timing of the reduction during the dry season, key model-data discrepancies in ET remain during spring (March-April) and monsoon (July-September) periods: i) At US-Fuf, the 11LAY observed ET is overestimated during the spring and early summer (Fig. S3a); ii) At US-Vcp, the model underestimates ET for much of the growing season (Fig. S3b); iii) at US-SRM the 11LAY model overestimates springtime ET (in contrast to other low-elevation monsoon sites) (Fig. S3c); and iv) the 11LAY model still slightly underestimates peak monsoon ET at the low-elevation shrub sites (US-SRM and US-Whs– Figs. S3 c-d) as seen in a previous semi-arid model evaluation study (Grippa et al., 2011).

Given the model tends to overestimate spring ET at US-Fuf and US-Vcp, this bias is likely related to the snowfall issues that are causing the model to underestimate spring soil moisture during the same period (Figs. 4 and 5 and see Section 3.2). The lack of a persistent snow pack in the model during this period could explain the positive bias in spring ET because in reality the presence of snow would suppress bare soil evaporation. To accurately diagnose this issue, we would need further information on snow mass or depth. Further support for the suggestion that modelled E is overestimated comes from examining the T/ET ratios. Although both E and T increase in the US-Fuf 11LAY simulations (compared to the 2LAY – Fig. S3a) – due to the increase in soil moisture (as previously described in Section 3.1 and Figs. 2 and S2a) – the larger increase in 11LAY E compared to T resulted in lower 11LAY T/ET ratios (Fig. S3a). The seasonal trajectory of T/ET ratios at US-Fuf appear to match data-derived estimates following the Zhou et al. (2016) method: the ratio peaks in the Spring before decreasing in July, with monsoon period T/ET values that are on average lower than the spring (Fig. 6). However, the magnitude of T/ET ratios are too low in all seasons given the 100% tree cover at this site with a LAI ~2.4. Whilst low spring 11LAY T/ET ratios may be due to overestimated E as a result of higher soil moisture and underestimated snow cover, the generally low bias in T/ET ratios may also be due to the fact there is no bare soil evaporation resistance term included in the default 11LAY version. We test this hypothesis in the following section.

https://doi.org/10.5194/hess-2019-598
Preprint. Discussion started: 4 December 2019


Hydrology and Earth System Sciences
Discussions

We can also glean some information on whether T or E (or both) are be responsible for the 11LAY overestimate of springtime ET at US-SRM by comparing modelled T/ET ratios against data-derived estimates. Observed T/ET ratios at the low-elevation sites were derived from independent eddy covariance data following the method of Scott and Biederman (2017) (Fig. 6). The observed spring T/ET at US-SRM is slightly underestimated by the model (Fig. 6). Given that T/ET ratios are underestimated by the model but ET is overestimated by the model, it is probable that spring E at this site is too high. Spring T could also be overestimated at US-SRM due potentially due to an overestimate in LAI (Fig. S5); however, the positive bias in E must be larger than the bias in T. If model LAI at US-SRM is too high during the spring, it is impossible to determine whether the shrub or grass LAI are inaccurate without independent, accurate estimates of seasonal leaf area for each vegetation type; however, in the field the spring C4 grass LAI is typically half that of its monsoon peak (R.L. Scott – pers. comm.) – a pattern not seen in the model (Fig. S6).

Data-derived T/ET ratios also help to diagnose why the 11LAY model underestimates monsoon ET at the low-elevation shrub sites (US-SRM and US-Whs– Figs. S3 c-d). Fig. 6 shows that the 11LAY model also underestimates monthly T/ET ratios, and furthermore, that the model does not capture the correct temporal trajectory (Fig. 6). Although the earlier summer drop in T/ET ratios in the 11LAY compared to the 2LAY simulations at grass and shrubland sites (Figs. S3 c-f) does result in a better match in ET between the model and the observations (Fig. 3), the 11LAY T/ET ratios are slightly out of phase. Observed T/ET ratios decline in June during the hottest, driest month, whereas model values decrease one month later in July (Fig. 6). Furthermore, the ratios do not increase as rapidly as observed during the wet monsoon period (July – September).

The underestimate in modelled monsoon T/ET ratios across all grassland and shrubland sites (and likely at US-Fuf and US-Vcp) suggests either that transpiration is too low or bare soil evaporation is too high. At the shrubland sites (US-SRM and US-Whs), both monsoon ET and T/ET are underestimated; therefore, for these sites it is plausible that the dominant cause is a lack of transpiring leaf area. Certainly, monsoon model-data ET biases are better correlated with LAI at shrubland sites compared to grassland sites (Fig. S7). The underestimate in modelled monsoon period leaf area could either be: i) an underestimate of maximum LAI for either grasses or shrubs; or ii) due to the fact the static vegetation fractions prescribed in the model do not allow for an increase in vegetation cover during the wet season (e.g. the lack grass growth in the model in interstitial bare soil areas). In contrast, at the grassland sites (US-SRG and US-Wkg) monsoon ET is well approximated by the 11LAY model; thus, the underestimate in T/ET ratios suggests that both the transpiration is too low and the bare soil evaporation too high. We attempted to address both of these possibilities with two further tests. Given there is a >45% bare soil fraction at all four grass and shrub dominated sites, in the final set of simulations in this study we tested a) the reduction of the bare soil fraction by increasing C4 grass cover; and b) the inclusion of this bare soil resistance term (Section 2.2.3). The results of these simulations are described in the following section.

https://doi.org/10.5194/hess-2019-598
Preprint. Discussion started: 4 December 2019


Hydrology and Earth System Sciences
Discussions

3.4 Testing decreased bare soil cover and the addition of the 11LAY bare soil resistance term

To further investigate the possibility that summer ET and T/ET ratios are underestimated at low-elevation sites because of a lack of transpiring leaf area, we increased C4 grass fraction to the maximum observed C4 grass cover under the most productive conditions at the expense of bare soil fraction (R. L. Scott pers. comm.). This decrease in bare soil fraction increased ET and T/ET ratios during the monsoon period at all sites (Fig. 7) and also increased ET during spring at the Santa Rita sites (Fig. S8). However, although the T/ET ratios better matched the data-derived values, the model now overestimated ET (Fig. 7).

As described in Section 3.3, the remaining model ET issues (and its component fluxes) in both high-elevation forest sites and low-elevation shrub- and grass-dominated sites could also be due to the fact the model simulates too much bare soil evaporation. The 11LAY version has an optional bare soil evaporation resistance term that is not activated in the default version; therefore, the 11LAY simulations presented thus far have not included any such resistance term. We tested the application of the bare soil resistance term across all sites. The lack of bare soil at the high-elevation forested sites resulted in a higher sensitivity of the TeNE forest PFT T to the addition of the bare soil evaporation resistance term (Fig. 8 – left column). The reduction in E during the winter allowed for higher soil moisture content (Figs. S9a and b) and therefore a greater T (and E) during the spring and summer. As a result, T/ET ratios were increased with the addition of the bare soil evaporation term, thus resolving the issue of negatively biased T/ET ratio issue seen in the default 11LAY simulations (see Section 3.3). However, the dramatic increase in T resulted in a simulated ET that strongly overestimated the observations (Fig. 8 and Figs. S9a and b); therefore, overall the addition of the bare soil evaporation resistance term did not improve the ET model-data fit at these sites. However, as discussed in Section 3.3, Spring ET may be overestimated at these sites due to the lack of a persistent snowpack. The increase in plant available moisture with the addition of the resistance term does lead to a strong increase in LAI at US-Vcp from a mean ~0.5 to a mean around 2.1 (Fig. S9b), which is much closer to the observed LAI for the site.

At all the low-elevation grass and shrub sites the addition of the bare soil resistance term resulted in a strong decrease in soil evaporation during the monsoon season, and a lesser, but non-negligible, decrease to almost zero evaporation during the winter (Fig. 8 – right column). Bare soil evaporation remained much the same during the spring and the hot, dry season months of May and June. As seen for the forest sites, the decline in bare soil evaporation during the monsoon period results in a slightly higher moisture storage (Figs. S9c-f), which in turn fractionally increases transpiration throughout the year (Fig. 8). The net effect is a reduction in ET during summer and winter and an increase in spring and dry season ET (Fig. 8). However, as for the forested sites, this net effect in the simulated ET produces a worse fit to ET. Therefore, the addition of this term does not resolve the ET issues documented in Section 3.3: A further positive bias in spring ET estimates is observed at US-SRM, and the underestimate in monsoon ET at US-SRM and US-Whs is further exacerbated. Furthermore, the near zero evaporation in the winter months with the introduction of the bare soil resistance term results in an increase in winter T/ET ratios. Therefore, at the low-elevation sites the monthly seasonality of T/ET differs quite considerably from the default 11LAY model runs (Figs. S9c-f) and does not follow the summer temporal trajectory observed by Scott et al. (2017).

https://doi.org/10.5194/hess-2019-598
Preprint. Discussion started: 4 December 2019


Hydrology and Earth System Sciences
Discussions

In a final test, we combined both the decrease in bare soil fraction with the addition of the bare soil resistance term. The addition of the bare soil resistance term reduced the positive bias seen with the increase of C4 grass (decrease of the bare soil fraction) (Fig. S10). However, as seen in the bare soil resistance tests with the original vegetation and bare soil fractions, the addition of the resistance term increased spring T due to the higher spring soil moisture – thus exacerbating the positive bias in ET. It is clear that neither of these tests fully deal with remaining ET model-data biases in the 11LAY version; however, the fit to the observations is still greatly improved in the 11LAY version compared to the 2LAY and the ET seasonal temporal dynamics remain much the same in all tests. It is therefore possible that some combination of these two configurations plus further calibration of hydrology, phenology, and stomatal conductance parameters would result in a near perfect fit to the data at these sites. This is beyond the scope of this study, but the options are discussed more in Section 4.

4 Discussion

This study showed that in comparison to a simple bucket model (Manabe, 1969), a discretized soil hydrology scheme based on the Richards equation – and associated model developments – results in considerable improvements in simulated semi-arid site soil moisture temporal dynamics that exhibit a more realistic response to rainfall events (contrary to the model-data comparison of Lohou et al., 2014). As a result, we see dramatic improvements in high temporal frequency to seasonal ET simulation. In particular, there is a dramatic improvement in the model's ability to capture the decline in ET during the hot, dry May-June period. Total runoff also decreased at forested sites owing to change in the calculation of soil moisture infiltration and partitioning between surface runoff and drainage. Such improvements might counter previous work highlighting that models tend to overestimate runoff (Grippa et al., 2017). Previous studies have also demonstrated that the more mechanistic descriptions of soil hydrology included in the latest LSM versions have resulted in improvements to surface latent and sensible heat fluxes (de Rosnay et al., 2002; Best et al., 2015); yet, few studies have specifically compared these two model versions across a range semi-arid ecosystems as we have attempted in this study. However, there remain a number of missing hydrological processes that have not yet been incorporated into LSMs, and/or inadequate existing processes, which will clearly have an impact on semi-arid hydrological modeling (Boone et al., 2009; Grippa et al., 2017). We highlight these in the sections below.

Issues with modelling vegetation dynamics in semi-arid ecosystems

Our analysis has suggested that in order to obtain realistic estimates of ET and its component fluxes, it is important that the model can accurately simulate changes in leaf area and in vegetation versus bare soil fractional cover. We have inferred that biases at low-elevation shrub and grassland site ET might be due to incorrect LAI simulations. The connection between vegetation fractional cover and LAI is also a particular issue in sparsely vegetated regions when low LAI effectively means more bare soil is coupled with the atmosphere. To account for this in ORCHIDEE, the bare soil fraction is slightly increased

https://doi.org/10.5194/hess-2019-598
Preprint. Discussion started: 4 December 2019


Hydrology and Earth System Sciences
Discussions

when LAI < 1, which is often the case at these sites; however, there are only limited observations to support this model specification. Similarly, there are not many LAI measurements for grasses and shrubs in these ecosystems; therefore, we have relied on estimating the LAI_{max} parameter from MODIS LAI data. While different satellite LAI products often correspond well to each other in terms of temporal variability, there is often a considerable spread in their absolute LAI values (Garrigues et al., 2008; Fan et al., 2013); therefore, the MODIS LAI data may not be accurate for these ecosystems. In any case, the satellite LAI values represent a mix of different vegetation types and unlike satellite reflectance data it is not possible to linearly unmix the satellite LAI estimates based on fractional cover. More field LAI measurements are needed from different vegetation types (especially grasses and shrubs) to verify what the likely maximum LAI is for each PFT. While not tested in this study, it is also possible that LSMs contain an inaccurate representation of different semi-arid vegetation phenological strategies, including drought-deciduous shrubs and annual versus perennial C4 grasses (Smith et al., 1997). Again, seasonal LAI measurements of different vegetation types would significantly help to improve or further develop semi-arid phenology models.

Alternatively, it may be that other model parameters and processes involved in leaf growth – for example photosynthesis-related parameters – are inaccurate and in need of calibration. Incorrect representations of how we model low temperature and high VPD constraints on stomatal conductance may also factor. At the high-elevation sites, we assumed the ponderosa pine trees should be modelled as temperate needleleaved evergreen PFT. The default model parameters assigned to this PFT may not be appropriate for modelling this plant functional type in water-limited semi-arid environments. Another likely issue for modelling low elevation sparsely vegetated semi-arid ecosystems with ORCHIEE is that there is no specific shrub PFT, although a recent ORCHIDEE version includes shrub PFTs for high latitude tundra ecosystems (Druel et al., 2017): In a future study we will adapt similar shrub parameterizations for semi-arid environments.

The importance of vegetation cover and seasonal changes in leaf area for modeling hydrological fluxes – particularly T – is not a new observation (e.g. Ha et al., 2015; Grippa et al., 2017). Baldocchi et al. (2010) found that LAI was important at five Mediterranean sites in California and Europe for determining how much carbon is assimilated and how much water is lost. Hogue et al. (2005) also found the Noah LSM was not able to replicate monsoon period LE increases at the Walnut Gulch sites, which they suggested may be related to inaccuracies in the satellite greenness fraction estimates that are used to run the model. Whitley et al. (2016 and 2017) also proposed that any improvements needed for terrestrial biosphere modelling of savanna ecosystems should include modifications to the phenology scheme and the split between fractional cover of trees and grasses.

ET partitioning (T/ET ratio)

ORCHIDEE underestimated T/ET ratios at the low-elevation sites (Fig. 6). The partitioning of ET between evaporation, E, and transpiration, T, is strongly related to presence of vegetative cover and plant physiological functioning. Whereas E results from physical processes and is linked to bare soil surface moisture and interception by the plant canopy, T is the result of leaf

https://doi.org/10.5194/hess-2019-598
Preprint. Discussion started: 4 December 2019


Hydrology and Earth System Sciences
Discussions

level biological processes regulated by LAI and stomatal conductance. Stomatal conductance is in turn modulated by root zone soil moisture availability, vapor pressure deficit (VPD) and atmospheric CO_2 concentrations (Novick et al., 2016). Both E and T are therefore governed by the fractional cover of vegetation, vegetation type, and changes in leaf area (i.e. phenology). Diagnosing and addressing model-data discrepancies in T/ET is important, specifically for dryland ecosystems where increases in vegetation productivity and/or cover in response to rising atmospheric CO_2 appears to be driving higher observed T/ET rates (Lian et al., 2018).

In agreement with this study, Lian et al. (2018) also show that CMIP5 models vastly underestimate T/ET ratios. They estimated a new global T/ET ratio of 0.62 ± 0.06, which is similar to the upscaled estimate of 0.57 +/= 7% of Wei et al. (2017), and suggest that model underestimates could be caused by misrepresentation of vegetation structure impacts on canopy light use, interception loss and root water uptake. Their conclusions lend further weight to our suggestion that further improvements in T/ET ratios may result from more accurate simulations of seasonal phenology and fractional vegetation cover (see previous section). Alternatively, Chang et al., (2018) have suggested that neglecting to account for lateral redistribution of moisture is responsible for model inability to capture T/ET partitioning. However, they also mention other LSM issues that might be affecting the T/ET ratio, such as the lack of root dynamics, vegetation shading, topographic effects and the representation of bare soil evaporation. In order to properly diagnose if discrepancies in modelled T/ET estimates are caused by inaccurate representation of lateral moisture redistribution, we need to perform a comparison of a spatially distributed model simulation with a high-density network of hydrological observations. Nevertheless, in spatially heterogeneous mixed shrub-grass ecosystems it seems likely that missing model processes will need to be accounted for before accurate simulations of T/ET ratios are achieved.

Bare soil evaporation

Following the results of our analysis, we surmised that bare soil evaporation might be overestimated, particularly at grassland sites that showed a good match between modelled and observed ET, but negatively biased T/ET ratios. The addition of a term that simulates bare soil evaporation resistance to dry soil served to reduce E in the summer and winter and thus to increase (and improve) the T/ET ratios. However, resulting changes in modelled ET provided a worse fit to the observations. It is possible that the bare soil resistance is only part of the solution, and that the simulation of ET and its component fluxes could be fixed with both a more realistic representation of semi-arid phenology or vegetation fractional cover at both grass and shrub dominated sites (as discussed above) and a re-calibration of relevant vegetation, soil and hydraulic parameters. Alternatively, the relatively simple implementation of a bare soil resistance term (Eq. 2 – Section 2.2.3) might need to be adapted to include bare soil evaporation resistance across a litter or biocrust layer. Decker et al. (2017) noted a reduction in positive biases in evaporation after updating the soil evaporation scheme in the CABLE LSM to limit bare soil evaporation based on the physics of evaporation from porous media across a viscous sublayer (analogous to a dry matter litter layer). Similarly, Swenson and Lawrence (2014) accounted for estimated positive biases in semi-arid bare soil evaporation fluxes (and too high fluctuations in modelled ET) in the Community Land Model (CLM v4.5) by replacing the existing empirical evaporation resistance function

https://doi.org/10.5194/hess-2019-598
Preprint. Discussion started: 4 December 2019

with a more mechanistic scheme. In the more mechanistic scheme, bare soil evaporation is formulated as a diffusion of moisture through a dry surface layer (again, approximating a litter layer).

Despite the fact that site-based soil textural properties were set for all simulations, it is also possible that modelled bare soil evaporation is too high because the model lacks vertical soil texture variability. The low-elevation sites typically have a very cobbly, rocky soil surface that is not accounted for in ORCHIDEE. Including soil texture variability with different soil horizons could further improve ORCHIDEE's capability to capture the correct E, ET and T/ET ratios. Furthermore, at the sparsely vegetated grass and shrub dominated sites in southern Arizona litter is barely present; instead, biological soil crusts (biocrusts) composed of assemblages of lichens, bryophytes, cyanobacteria, algae and microbes form across much of the bare soil surface (Belnap et al., 2016). Biocrust layers may significantly alter bare soil evaporation (and other aspects of ecosystem ecology and functioning – Ferrenberg et al., 2017) in sparsely vegetated regions in ways that have not yet been considered in any LSM bare soil evaporation scheme. Therefore, it is possible that in addition to a more mechanistically-based formulation of bare soil evaporation through a litter layer (as per Swenson and Lawrence, 2014 or Decker et al., 2017), separate formulations of evaporation through biocrust/mulch layers may need to be developed (e.g. Saux-Picard et al., 2009).

High-elevation model snowpack and snow melt predictions

The model also needs to be tested at other high-elevation semi-arid mountainous sites (such as the Sierra Nevada mountains in California) for which spring snowmelt is the predominant (and controlling) source of moisture. More specifically, more information on snow cover, depth or mass would be useful to test if the precipitation data measured by the meteorological stations accurately captures the right amount of snowfall. It is crucial that LSMs accurately capture semi-arid high-elevation snowfall temporal dynamics if we are to have unbiased projections in future moisture availability and productivity for these regions.

5 Conclusions

These results strongly suggest that a more complex, process-based hydrology model – in particular one which contains fine scale discretization of the upper soil moisture layers and associated improvements in bare soil evaporation and water limitation of stomatal conductance – improves high frequency to seasonal predictions of the root-zone soil moisture dynamics and ET (as seen in de Rosnay et al., 2002). Associated changes in the calculations of runoff, soil moisture infiltration, and bottom layer drainage also resulted in more realistic estimates of surface runoff. ORCHIDEE CMIP5 simulations used the 2-layer conceptual bucket scheme of Manabe (1969); therefore, ORCHIDEE CMIP5 predictions of semi-arid water availability and consequent impacts on ecosystem functioning and feedback to climate were likely inaccurate. Despite the appeal of simplicity and low calculation costs, 2-layer simple bucket hydrology models are likely unsuitable for accurate water flux simulations in

https://doi.org/10.5194/hess-2019-598
Preprint. Discussion started: 4 December 2019


Hydrology and Earth System Sciences
Discussions

the semi-arid SW US. The forthcoming ORCHIDEE CMIP6 simulations of semi-arid soil moisture availability will likely provide more accurate and reliable results. Remaining discrepancies in underestimated soil moisture at high-elevation semi-arid forested sites might be related to inaccurate snowfall forcing data, with consequent implications of for predictions of water availability in regions that rely on springtime snowmelt. The addition of an empirical bare soil resistance term by itself did not improve estimates of ET in these systems, although T/ET ratios were increased and improved with the addition of this term. Our analysis shows that remaining discrepancies in sparsely vegetated grass and shrub dominated semi-arid site ET simulations (and its constituent fluxes) might therefore be related to a combination of factors impacting the amount of transpiring leaf area and resistance to bare soil evaporation. We recommend that future work on improving LSM semi-arid hydrological predictions focuses not only on issues highlighted in previous studies such as dynamic root zone moisture uptake, inclusion of ground water, lateral and vertical redistribution of moisture (e.g. Whitley et al., 2016; 2017; Grippa et al., 2017) but also on: i) multi-variable calibration of hydrology-related parameters at fully documented sites; ii) better estimating the seasonal trajectory and amount of vegetation fractional cover; and iii) more mechanistic descriptions of resistance to bare soil evaporation.

Code availability

The ORCHIDEE model code and documentation are publicly available via the ORCHIDEE wiki page (https://forge.ipsl.jussieu.fr/orchidee/browser) under the CeCILL license (http://www.cecill.info/index.en.html). ORCHIDEE model code is written in Fortran 90 and is maintained and developed under an SVN version control system at the Institute Pierre Simon Laplace (IPSL) in France.

Data availability

Meteorological forcing and evapotranspiration data for each are available from the Ameriflux site: https://ameriflux.lbl.gov. Soil moisture was obtained directly from site PIs. Vegetation and soil texture characteristics were derived from the published literature, as specified in Table 1, and from site PIs.

Appendices

Author contribution

NM, RLS, JAB and DJPM designed the overall study. NM carried out the model simulations, post-simulation analysis and figure plotting. CO, NV and AD provided detailed inputs on model description/code and recommendations for further tests to diagnose model-data deficiencies. NV provided scripts to gap-fill the meteorological data. JAB gap-filled the ET data. RLS,

https://doi.org/10.5194/hess-2019-598
Preprint. Discussion started: 4 December 2019

JAB, TK and ML provided gap-filled soil moisture data and information on site characteristics and typical behaviour of seasonal vegetation cover, LAI, and snowfall. NM wrote the manuscript. All co-authors provided detailed comments, suggestions, and edits on the first and second drafts of the manuscript.

Competing Interests

The authors declare that they have no conflict of interest.

Acknowledgements

Funding for AmeriFlux data resources and data collection at US-SRM, US-SRG, US-Wkg, and US-Whs was provided by the U.S. Department of Energy's Office of Science and the USDA. Data collection at US-Fuf supported by grants from the North American Carbon Program/USDA CREES NRI (2004-3511115057), the U.S. National Science Foundation MRI Program, Science Foundation Arizona (CAA 0-203-08), the Arizona Water Institute, and the Mission Research Program, School of Forestry, Northern Arizona University (McIntire-Stennis/ Arizona Bureau of Forestry). The US-Vcp site funded by U.S. DOE Office of Science through the AmeriFlux Management Project (AMP) at Lawrence Berkeley National Laboratory (Award #7074628) and Catalina-Jemez Critical Zone Observatory (NSF EAR 1331408). NM was funded by the US National Science Foundation Award Numbers 1065790 (Emerging Frontiers Program) and 1754430 (Division of Environmental Biology Ecosystems Program). We would like to thank the ORCHIDEE team for development and maintenance of the ORCHIDEE code and for providing the ORCHIDEE version used in this study.

References

Abramowitz, G., Leuning, R., Clark, M., & Pitman, A.: Evaluating the performance of land surface models. *Journal of Climate*, *21*(21), 5468-5481, 2008.

Ahlström, A., Raupach, M.R., Schurgers, G., Smith, B., Arneth, A., Jung, M., Reichstein, M., Canadell, J.G., Friedlingstein, P., Jain, A.K. and Kato, E.: The dominant role of semi-arid ecosystems in the trend and variability of the land CO2 sink. Science, 348(6237), pp.895-899, 2015

Allen CD (2016) Chapter 4 - Forest ecosystem reorganization underway in the Southwestern US: A preview of widespread forest changes in the Anthropocene? In: RP Bixler and C Miller (eds) Forest Conservation and Management in the Anthropocene: Adaptation of Science, Policy and Practice. University Press of Colorado, Boulder, Colorado, pp 57-79

Ault, T. R., Mankin, J. S., Cook, B. I. and Smerdon, J. E.: Relative impacts of mitigation, temperature, and precipitation on 21st-century megadrought risk in the American Southwest, Science Advances, 2(10), doi:10.1126/sciadv.1600873, 2016.

Baldocchi, D. D., Ma, S., Rambal, S., Misson, L., Ourcival, J.-M., Limousin, J.-M., Pereira, J. and Papale, D.: On the differential advantages of evergreenness and deciduousness in mediterranean oak woodlands: a flux perspective, Ecological Applications, 20(6), 1583–1597, doi:10.1890/08-2047.1, 2010.

Bastrikov, V., Peylin, P. and Ottlé, C: Optimizing albedo computation in ORCHIDEE land surface model by assimilating MODIS satellite observation data. In prep.

Belnap, J., Weber, B., and Büdel B.: Biological soil crusts: an organizing principle in drylands, Springer., 2016.

Best, M. J., Abramowitz, G., Johnson, H. R., Pitman, A. J., Balsamo, G., Boone, A., Cuntz, M., Decharme, B., Dirmeyer, P. A., Dong, J., Ek, M., Guo, Z., Haverd, V., B. J. J. Van Den Hurk, Nearing, G. S., Pak, B., Peters-Lidard, C., Santanello, J. A., Stevens, L. and Vuichard, N.: The Plumbing of Land Surface Models: Benchmarking Model Performance, Journal of Hydrometeorology, 16(3), 1425–1442, doi:10.1175/jhm-d-14-0158.1, 2015.

Berg, A., Findell, K., Lintner, B., Giannini, A., Seneviratne, S., van den Hurk, B., Lorenz, R., Pitman, A., Hagemann, S., Meier, A., Cheruy, F., Ducharne, A., Malyshev, S., and Milly, P.C.D.: Land-atmosphere feedbacks amplify aridity increase over land under global warning, Nature Climate Change, 6, 869–874, doi:10.1038/nclimate3029, 2016.

Biederman, J. A., Scott, R. L., Goulden, M. L., Vargas, R., Litvak, M. E., Kolb, T. E., Yepez, E. A., Oechel, W. C., Blanken, P. D., Bell, T. W., Garatuza-Payan, J., Maurer, G. E., Dore, S. and Burns, S. P.: Terrestrial carbon balance in a drier world: the effects of water availability in southwestern North America, Global Change Biology, 22(5), 1867–1879, doi:10.1111/gcb.13222, 2016.

Boone, A., de Rosnay, P. D., Balsamo, G., Beljaars, A., Chopin, F., Decharme, B., Delire, C., Ducharne, A., Gascoin, S., Grippa, M., Guichard, F., Gusev, Y., Harris, P., Jarlan, L., Kergoat, L., Mougin, E., Nasonova, O., Norgaard, A., Orgeval, T., Ottlé, C., Poccard-Leclercq, I., Polcher, J., Sandholt, I., Saux-Picart, S., Taylor, C. and Xue, Y.: The AMMA Land Surface Model Intercomparison Project (ALMIP), Bulletin of the American Meteorological Society, 90(12), 1865–1880, doi:10.1175/2009bams2786.1, 2009.

Botta, A., Viovy, N., Ciais, P., Friedlingstein, P. and Monfray, P.: A global prognostic scheme of leaf onset using satellite data, Global Change Biology, 6(7), 709–725, doi:10.1046/j.1365-2486.2000.00362.x, 2000.

Campoy, A., Ducharne, A., Cheruy, F., Hourdin, F., Polcher, J. and Dupont, J. C.: Response of land surface fluxes and precipitation to different soil bottom hydrological conditions in a general circulation model, Journal of Geophysical Research: Atmospheres, 118(19), doi:10.1002/jgrd.50627, 2013.

Chang, L.-L., Dwivedi, R., Knowles, J. F., Fang, Y.-H., Niu, G.-Y., Pelletier, J. D., Rasmussen, C., Durcik, M., Barron-Gafford, G. A. and Meixner, T.: Why Do Large-Scale Land Surface Models Produce a Low Ratio of Transpiration to Evapotranspiration?, Journal of Geophysical Research: Atmospheres, 123(17), 9109–9130, doi:10.1029/2018jd029159, 2018.

Clark, M. P., Fan, Y., Lawrence, D. M., Adam, J. C., Bolster, D., Gochis, D. J., Hooper, R. P., Kumar, M., Leung, L. R., Mackay, D. S., Maxwell, R. M., Shen, C., Swenson, S. C. and Zeng, X.: Improving the representation of hydrologic processes in Earth System Models, Water Resources Research, 51(8), 5929–5956, doi:10.1002/2015wr017096, 2015.

Cook, B. I., Ault, T. R. and Smerdon, J. E.: Unprecedented 21st century drought risk in the American Southwest and Central Plains, Science Advances, 1(1), doi:10.1126/sciadv.1400082, 2015.

De Kauwe, M. G., Zhou, S.-X., Medlyn, B. E., Pitman, A. J., Wang, Y.-P., Duursma, R. A. and Prentice, I. C.: Do land surface models need to include differential plant species responses to drought? Examining model predictions across a mesic-xeric gradient in Europe, Biogeosciences, 12(24), 7503–7518, doi:10.5194/bg-12-7503-2015, 2015.

De Kauwe, M. G., Taylor, C. M., Harris, P. P., Weedon, G. P. and Ellis, R. J.: Quantifying Land Surface Temperature Variability for Two Sahelian Mesoscale Regions during the Wet Season, Journal of Hydrometeorology, 14(5), 1605–1619, doi:10.1175/jhm-d-12-0141.1, 2013.

Decker, M., Or, D., Pitman, A. and Ukkola, A.: New turbulent resistance parameterization for soil evaporation based on a pore-scale model: Impact on surface fluxes in CABLE, Journal of Advances in Modeling Earth Systems, 9(1), 220–238, doi:10.1002/2016ms000832, 2017.

de Rosnay, P. and Polcher, J.: Modelling root water uptake in a complex land surface scheme coupled to a GCM, Hydrol. Earth Syst. Sci., 2, 239–255, doi:10.5194/hess-2-239-1998, 1998.

de Rosnay, P. D., Polcher, J., Bruen, M. and Laval, K.: Impact of a physically based soil water flow and soil-plant interaction representation for modeling large-scale land surface processes, Journal of Geophysical Research: Atmospheres, 107(D11), doi:10.1029/2001jd000634, 2002.

Diffenbaugh, N. S., Giorgi, F. and Pal, J. S.: Climate change hotspots in the United States, Geophysical Research Letters, 35(16), doi:10.1029/2008gl035075, 2008.

Dirmeyer, P. A.: A History and Review of the Global Soil Wetness Project (GSWP), Journal of Hydrometeorology, 110404091221083, doi:10.1175/jhm-d-10-05010, 2011.

Donat, M. G., Lowry, A. L., Alexander, L. V., O'Gorman, P. A. and Maher, N.: More extreme precipitation in the world's dry and wet regions, Nature Climate Change, 6(5), 508–513, doi:10.1038/nclimate2941, 2016.

Dore, S., Kolb, T. E., Montes-Helu, M., Eckert, S. E., Sullivan, B. W., Hungate, B. A., Kaye, J. P., Hart, S. C., Koch, G. W. and Finkral, A.: Carbon and water fluxes from ponderosa pine forests disturbed by wildfire and thinning, Ecological Applications, 20(3), 663–683, doi:10.1890/09-0934.1, 2010.

Dore, S., Montes-Helu, M., Hart, S. C., Hungate, B. A., Koch, G. W., Moon, J. B., Finkral, A. J. and Kolb, T. E.: Recovery of ponderosa pine ecosystem carbon and water fluxes from thinning and stand-replacing fire, Global Change Biology, 18(10), 3171–3185, doi:10.1111/j.1365-2486.2012.02775.x, 2012.

d'Orgeval, T., Polcher, J., and de Rosnay, P.: Sensitivity of the West African hydrological cycle in ORCHIDEE to infiltration processes, Hydrol. Earth Syst. Sci, 12, 1387–1401, 2008.

Druel, A., Peylin, P., Krinner, G., Ciais, P., Viovy, N., Peregon, A., Bastrikov, V., Kosykh, N., and Mironycheva-Tokareva, N.: Towards a more detailed representation of high-latitude vegetation in the global land surface model ORCHIDEE (ORC-HL-VEGv1.0), Geosci. Model Dev., 10, 4693–4722.

Ducharne, A., Laval, K. and Polcher, J.: Sensitivity of the hydrological cycle to the parametrization of soil hydrology in a GCM, Climate Dynamics, 14(5), 307–327, doi:10.1007/s003820050226, 1998.

Ducharne, A., Ghattas, J., Maignan, F., Ottlé, C., Vuichard, N., Guimberteau, M., Krinner, G., Polcher, J., Tafasca, S., Bastrikov, V., Cheruy, F., Guénet, B., Mizuochi, H., Peylin, P., Tootchi, A. and Wang, F.: Soil water processes in the ORCHIDEE-2.0 land surface model: state of the art for CMIP6, In prep. for Geosci. Model Dev.

Ducoudré, N. I., Laval, K. and Perrier, A.: SECHIBA, a New Set of Parameterizations of the Hydrologic Exchanges at the Land-Atmosphere Interface within the LMD Atmospheric General Circulation Model, Journal of Climate, 6(2), 248–273, doi:10.1175/1520-0442(1993)006<0248:sansop>2.0.co;2, 1993.

Dufresne, J.-L., Foujols, M.-A., Denvil, S., Caubel, A., Marti, O., Aumont, O., Balkanski, Y., Bekki, S., Bellenger, H., Benshila, R., Bony, S., Bopp, L., Braconnot, P., Brockmann, P., Cadule, P., Cheruy, F., Codron, F., Cozic, A., Cugnet, D., Noblet, N. D., Duvel, J.-P., Ethé, C., Fairhead, L., Fichefet, T., Flavoni, S., Friedlingstein, P., Grandpeix, J.-Y., Guez, L., Guilyardi, E., Hauglustaine, D., Hourdin, F., Idelkadi, A., Ghattas, J., Joussaume, S., Kageyama, M., Krinner, G., Labetoulle, S., Lahellec, A., Lefebvre, M.-P., Lefevre, F., Levy, C., Li, Z. X., Lloyd, J., Lott, F., Madec, G., Mancip, M., Marchand, M., Masson, S., Meurdesoif, Y., Mignot, J., Musat, I., Parouty, S., Polcher, J., Rio, C., Schulz, M., Swingedouw, D., Szopa, S., Talandier, C., Terray, P., Viovy, N. and Vuichard, N.: Climate change projections using the IPSL-CM5 Earth System Model: from CMIP3 to CMIP5, Climate Dynamics, 40(9-10), 2123–2165, doi:10.1007/s00382-012-1636-1, 2013.

Fang, H., Jiang, C., Li, W., Wei, S., Baret, F., Chen, J.M., Garcia-Haro, J., Liang, S., Liu, R., Myneni, R.B. and Pinty, B.: Characterization and intercomparison of global moderate resolution leaf area index (LAI) products: Analysis of climatologies and theoretical uncertainties. Journal of Geophysical Research: Biogeosciences, 118(2), 529-548, 2013.

https://doi.org/10.5194/hess-2019-598
Preprint. Discussion started: 4 December 2019

Ferrenberg, S. and Reed, S. C.: Biocrust ecology: unifying micro- and macro-scales to confront global change, New Phytologist, 216(3), 643–646, doi:10.1111/nph.14826, 2017.

Garrigues, S., Lacaze, R., Baret, F.J.T.M., Morisette, J.T., Weiss, M., Nickeson, J.E., Fernandes, R., Plummer, S., Shabanov, N.V., Myneni, R.B. and Knyazikhin, Y.: Validation and intercomparison of global Leaf Area Index products derived from remote sensing data. Journal of Geophysical Research: Biogeosciences, 113(G2), 2008.

Genuchten, M. T. V.: A Closed-form Equation for Predicting the Hydraulic Conductivity of Unsaturated Soils1, Soil Science Society of America Journal, 44(5), 892, doi:10.2136/sssaj1980.03615995004400050002x, 1980.

Gremer, J. R., Bradford, J. B., Munson, S. M. and Duniway, M. C.: Desert grassland responses to climate and soil moisture suggest divergent vulnerabilities across the southwestern United States, Global Change Biology, 21(11), 4049–4062, doi:10.1111/gcb.13043, 2015.

Grippa, M., Kergoat, L., Frappart, F., Araud, Q., Boone, A., de Rosnay, P. D., Lemoine, J.-M., Gascoin, S., Balsamo, G., Ottlé, C., Decharme, B., Saux-Picart, S. and Ramillien, G.: Land water storage variability over West Africa estimated by Gravity Recovery and Climate Experiment (GRACE) and land surface models, Water Resources Research, 47(5), doi:10.1029/2009wr008856, 2011.

Grippa, M., Kergoat, L., Boone, A., Peugeot, C., Demarty, J., Cappelaere, B., Gal, L., Hiernaux, P., Mougin, E., Ducharne, A., Dutra, E., Anderson, M. and Hain, C.: Modeling Surface Runoff and Water Fluxes over Contrasted Soils in the Pastoral Sahel: Evaluation of the ALMIP2 Land Surface Models over the Gourma Region in Mali, Journal of Hydrometeorology, 18(7), 1847–1866, doi:10.1175/jhm-d-16-0170.1, 2017.

Guimberteau, M., Drapeau, G., Ronchail, J., Sultan, B., Polcher, J., Martinez, J.-M., Prigent, C., Guyot, J.-L., Cochonneau, G., Espinoza, J. C., Filizola, N., Fraizy, P., Lavado, W., Oliveira, E. D., Pombosa, R., Noriega, L. and Vauchel, P.: Discharge simulation in the sub-basins of the Amazon using ORCHIDEE forced by new datasets, Hydrology and Earth System Sciences, 16(3), 911–935, doi:10.5194/hess-16-911-2012, 2012a.

Guimberteau, M., Ronchail, J., Espinoza, J. C., Lengaigne, M., Sultan, B., Polcher, J., Drapeau, G., Guyot, J.-L., Ducharne, A. and Ciais, P.: Future changes in precipitation and impacts on extreme streamflow over Amazonian sub-basins, Environmental Research Letters, 8(1), 014035, doi:10.1088/1748-9326/8/1/014035, 2013.

Guimberteau, M., Ducharne, A., Ciais, P., Boisier, J. P., Peng, S., Weirdt, M. D. and Verbeeck, H.: Testing conceptual and physically based soil hydrology schemes against observations for the Amazon Basin, Geoscientific Model Development, 7(3), 1115–1136, doi:10.5194/gmd-7-1115-2014, 2014.

Guimberteau, M., Perrier, A., Laval, K. and Polcher, J.: A comprehensive approach to analyze discrepancies between land surface models and in-situ measurements: a case study over the US and Illinois with SECHIBA forced by NLDAS, Hydrology and Earth System Sciences, 16(11), 3973–3988, doi:10.5194/hess-16-3973-2012, 2012b.

Haverd, V., Ahlström, A., Smith, B. and Canadell, J. G.: Carbon cycle responses of semi-arid ecosystems to positive asymmetry in rainfall, Global Change Biology, 23(2), 793–800, doi:10.1111/gcb.13412, 2016.

Hogue, T. S., Bastidas, L., Gupta, H., Sorooshian, S., Mitchell, K. and Emmerich, W.: Evaluation and Transferability of the Noah Land Surface Model in Semiarid Environments, Journal of Hydrometeorology, 6(1), 68–84, doi:10.1175/jhm-402.1, 2005.

Huang, J., Yu, H., Dai, A., Wei, Y. and Kang, L.: Drylands face potential threat under 2 °C global warming target, Nature Climate Change, 7(6), 417–422, doi:10.1038/nclimate3275, 2017.

IPCC, Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovern- mental Panel on Climate Change [Stocker, T.F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 1535 pp., 2013.

Koster, R. D.: Regions of Strong Coupling Between Soil Moisture and Precipitation, Science, 305(5687), 1138–1140, doi:10.1126/science.1100217, 2004.

Krinner, G., Viovy, N., Noblet-Ducoudré, N. D., Ogée, J., Polcher, J., Friedlingstein, P., Ciais, P., Sitch, S. and Prentice, I. C.: A dynamic global vegetation model for studies of the coupled atmosphere-biosphere system, Global Biogeochemical Cycles, 19(1), doi:10.1029/2003gb002199, 2005.

Lian, X., Piao, S., Huntingford, C., Li, Y., Zeng, Z., Wang, X., Ciais, P., Mcvicar, T. R., Peng, S., Ottlé, C., Yang, H., Yang, Y., Zhang, Y. and Wang, T.: Partitioning global land evapotranspiration using CMIP5 models constrained by observations, Nature Climate Change, 8(7), 640–646, doi:10.1038/s41558-018-0207-9, 2018.

Lohou, F., Kergoat, L., Guichard, F., Boone, A., Cappelaere, B., Cohard, J.-M., Demarty, J., Galle, S., Grippa, M., Peugeot, C., Ramier, D., Taylor, C. M. and Timouk, F.: Surface response to rain events throughout the West African monsoon, Atmospheric Chemistry and Physics, 14(8), 3883–3898, doi:10.5194/acp-14-3883-2014, 2014.

MacBean, N., Maignan, F., Peylin, P., Bacour, C., Bréon, F.-M. and Ciais, P.: Using satellite data to improve the leaf phenology of a global terrestrial biosphere model, Biogeosciences, 12(23), 7185–7208, doi:10.5194/bg-12-7185-2015, 2015.

Maestre, F. T., Salguero-Gomez, R. and Quero, J. L.: It is getting hotter in here: determining and projecting the impacts of global environmental change on drylands, Philosophical Transactions of the Royal Society B: Biological Sciences, 367(1606), 3062–3075, doi:10.1098/rstb.2011.0323, 2012.

Manabe, S.: Climate And The Ocean Circulation1, Monthly Weather Review, 97(11), 739–774, doi:10.1175/1520-0493(1969)097<0739:catoc>2.3.co;2, 1969.

Mualem, Y.: A new model for predicting the hydraulic conductivity of unsaturated porous media, Water Resources Research, 12(3), 513–522, doi:10.1029/wr012i003p00513, 1976.

Mueller, B. and Seneviratne, S. I.: Systematic land climate and evapotranspiration biases in CMIP5 simulations, Geophysical Research Letters, 41(1), 128–134, doi:10.1002/2013gl058055, 2014.

Novick, K. A., Ficklin, D. L., Stoy, P. C., Williams, C. A., Bohrer, G., Oishi, A. C., Papuga, S. A., Blanken, P. D., Noormets, A., Sulman, B. N., Scott, R. L., Wang, L. and Phillips, R. P.: The increasing importance of atmospheric demand for ecosystem water and carbon fluxes, Nature Climate Change, 6(11), 1023–1027, doi:10.1038/nclimate3114, 2016.

Peylin, P., et al.: The ORCHIDEE global land surface model version v2.0: description and evaluation. Geoscientific Model Development, in prep.

Poulter, B., Frank, D., Ciais, P., Myneni, R. B., Andela, N., Bi, J., Broquet, G., Canadell, J. G., Chevallier, F., Liu, Y. Y., Running, S. W., Sitch, S. and Werf, G. R. V. D.: Contribution of semi-arid ecosystems to interannual variability of the global carbon cycle, Nature, 509(7502), 600–603, doi:10.1038/nature13376, 2014.

Raoult, N., Delorme, B., Ottlé, C., Peylin, P., Bastrikov, V., Maugis, P. and Polcher, J.: Confronting Soil Moisture Dynamics from the ORCHIDEE Land Surface Model With the ESA-CCI Product: Perspectives for Data Assimilation, Remote Sensing, 10(11), 1786, doi:10.3390/rs10111786, 2018.

Reynolds, C. A., Jackson, T. J. and Rawls, W. J.: Estimating soil water-holding capacities by linking the Food and Agriculture Organization Soil map of the world with global pedon databases and continuous pedotransfer functions, Water Resources Research, 36(12), 3653–3662, doi:10.1029/2000wr900130, 2000.

Richards, L. A.: Capillary Conduction Of Liquids Through Porous Mediums, Physics, 1(5), 318–333, doi:10.1063/1.1745010, 1931.

Saux-Picart, S., Ottlé, C., Perrier, A., Decharme, B., Coudert, B., Zribi, M., ... & Ramier, D.: SEtHyS_Savannah: A multiple source land surface model applied to Sahelian landscapes. Agricultural and Forest Meteorology, 149(9), 1421-1432, 2009.

Scanlon, B. R., Zhang, Z., Rateb, A., Sun, A., Wiese, D., Save, H., Beaudoing, H., Lo, M. H., Müller-Schmied, H., Döll, P., Beek, R., Swenson, S., Lawrence, D., Croteau, M. and Reedy, R. C.: Tracking Seasonal Fluctuations in Land Water Storage Using Global Models and GRACE Satellites, Geophysical Research Letters, doi:10.1029/2018gl081836, 2019.

Scanlon, B. R., Zhang, Z., Save, H., Sun, A. Y., Schmied, H. M., Beek, L. P. H. V., Wiese, D. N., Wada, Y., Long, D., Reedy, R. C., Longuevergne, L., Döll, P. and Bierkens, M. F. P.: Global models underestimate large decadal declining and rising water storage trends relative to GRACE satellite data, Proceedings of the National Academy of Sciences, 115(6), doi:10.1073/pnas.1704665115, 2018.

Scott, R. L. and Biederman, J. A.: Critical Zone Water Balance Over 13 Years in a Semiarid Savanna, Water Resources Research, 55(1), 574–588, doi:10.1029/2018wr023477, 2019.

Scott, R. L. and Biederman, J. A.: Partitioning evapotranspiration using long-term carbon dioxide and water vapor fluxes, Geophysical Research Letters, 44(13), 6833–6840, doi:10.1002/2017gl074324, 2017.

Scott, R. L., Biederman, J. A., Hamerlynck, E. P. and Barron-Gafford, G. A.: The carbon balance pivot point of southwestern U.S. semiarid ecosystems: Insights from the 21st century drought, Journal of Geophysical Research: Biogeosciences, 120(12), 2612–2624, doi:10.1002/2015jg003181, 2015.

Scott, R. L., Biederman, J. A., Hamerlynck, E. P. and Barron-Gafford, G. A.: The carbon balance pivot point of southwestern U.S. semiarid ecosystems: Insights from the 21st century drought, Journal of Geophysical Research: Biogeosciences, 120(12), 2612–2624, doi:10.1002/2015jg003181, 2015.

Seager, R., Ting, M., Held, I., Kushnir, Y., Lu, J., Vecchi, G., Huang, H.-P., Harnik, N., Leetmaa, A., Lau, N.-C., Li, C., Velez, J. and Naik, N.: Model Projections of an Imminent Transition to a More Arid Climate in Southwestern North America, Science, 316(5828), 1181–1184, doi:10.1126/science.1139601, 2007.

Sellers, P. J., Heiser, M. D. and Hall, F. G.: Relations between surface conductance and spectral vegetation indices at intermediate (100 m2to 15 km2) length scales, Journal of Geophysical Research, 97(D17), 19033, doi:10.1029/92jd01096, 1992.

Seneviratne, S. I., Wilhelm, M., Stanelle, T., Hurk, B., Hagemann, S., Berg, A., Cheruy, F., Higgins, M. E., Meier, A., Brovkin, V., Claussen, M., Ducharne, A., Dufresne, J. L., Findell, K. L., Ghattas, J., Lawrence, D. M., Malyshev, S., Rummukainen, M. and Smith, B.: Impact of soil moisture-climate feedbacks on CMIP5 projections: First results from the GLACE-CMIP5 experiment, Geophysical Research Letters, 40(19), 5212–5217, doi:10.1002/grl.50956, 2013.

Sippel, S., Zscheischler, J., Heimann, M., Lange, H., Mahecha, M. D., Oldenborgh, G. J. V., Otto, F. E. L. and Reichstein, M.: Have precipitation extremes and annual totals been increasing in the world's dry regions over the last 60 years?, Hydrology and Earth System Sciences, 21(1), 441–458, doi:10.5194/hess-21-441-2017, 2017.

Smith, S.D., Monson, R.K. and Anderson, J.E.: Physiological Ecology of North American Desert Plants, Springer-Verlag Berlin Heidelberg, 1997.

Swenson, S. C. and Lawrence, D. M.: Assessing a dry surface layer-based soil resistance parameterization for the Community Land Model using GRACE and FLUXNET-MTE data, Journal of Geophysical Research: Atmospheres, 119(17), doi:10.1002/2014jd022314, 2014.

Tietjen, B., Jeltsch, F., Zehe, E., Classen, N., Groengroeft, A., Schiffers, K. and Oldeland, J.: Effects of climate change on the coupled dynamics of water and vegetation in drylands, Ecohydrology, doi:10.1002/eco.70, 2009.

Ukkola, A. M., Kauwe, M. G. D., Pitman, A. J., Best, M. J., Abramowitz, G., Haverd, V., Decker, M. and Haughton, N.: Land surface models systematically overestimate the intensity, duration and magnitude of seasonal-scale evaporative droughts, Environmental Research Letters, 11(10), 104012, doi:10.1088/1748-9326/11/10/104012, 2016a.

Ukkola, A. M., Pitman, A. J., Decker, M., Kauwe, M. G. D., Abramowitz, G., Kala, J. and Wang, Y.-P.: Modelling evapotranspiration during precipitation deficits: identifying critical processes in a land surface model, Hydrology and Earth System Sciences, 20(6), 2403–2419, doi:10.5194/hess-20-2403-2016, 2016b.

Villarreal, S., Vargas, R., Yepez, E. A., Acosta, J. S., Castro, A., Escoto-Rodriguez, M., Lopez, E., Martínez-Osuna, J., Rodriguez, J. C., Smith, S. V., Vivoni, E. R. and Watts, C. J.: Contrasting precipitation seasonality influences evapotranspiration dynamics in water-limited shrublands, Journal of Geophysical Research: Biogeosciences, 121(2), 494–508, doi:10.1002/2015jg003169, 2016.

Vuichard, N. and Papale, D.: Filling the gaps in meteorological continuous data measured at FLUXNET sites with ERA-Interim reanalysis, Earth System Science Data, 7(2), 157–171, doi:10.5194/essd-7-157-2015, 2015.Wang, T., Ottlé, C., Boone, A., Ciais, P., Brun, E., Morin, S., Krinner, G., Piao, S. and Peng, S.: Evaluation of an improved intermediate complexity snow scheme in the ORCHIDEE land surface model, Journal of Geophysical Research: Atmospheres, 118(12), 6064–6079, doi:10.1002/jgrd.50395, 2013.

Whitley, R., Beringer, J., Hutley, L. B., Abramowitz, G., Kauwe, M. G. D., Evans, B., Haverd, V., Li, L., Moore, C., Ryu, Y., Scheiter, S., Schymanski, S. J., Smith, B., Wang, Y.-P., Williams, M. and Yu, Q.: Challenges and opportunities in land surface modelling of savanna ecosystems, Biogeosciences, 14(20), 4711–4732, doi:10.5194/bg-14-4711-2017, 2017.

Whitley, R., Beringer, J., Hutley, L. B., Abramowitz, G., Kauwe, M. G. D., Duursma, R., Evans, B., Haverd, V., Li, L., Ryu, Y., Smith, B., Wang, Y.-P., Williams, M. and Yu, Q.: A model inter-comparison study to examine limiting factors in modelling Australian tropical savannas, Biogeosciences, 13(11), 3245–3265, doi:10.5194/bg-13-3245-2016, 2016.

https://doi.org/10.5194/hess-2019-598
Preprint. Discussion started: 4 December 2019

Hydrology and Earth System Sciences
Discussions

965

970

https://doi.org/10.5194/hess-2019-598
Preprint. Discussion started: 4 December 2019

Hydrology and Earth System Sciences
Discussions

Table 1: Site descriptions, period of available site data, and associated ORCHIDEE model parameters, including vegetation plant functional type (PFT), soil texture fractions and maximum LAI used in ORCHIDEE model simulations (also see Table 1 for general site descriptions). Simulation period correspond to the period of available site data. PFT fractional cover and the fraction of each soil texture class are defined in ORCHIDEE by the user. Note that ORCHIDEE does not contain and explicit representation of shrub PFTs; therefore, shrubs were included in the forest PFTs. The maximum LAI has a default setting in ORCHIDEE that has not been used here; instead, values based on the site literature have been prescribed in the model. The USDA soil texture classification (12 classes – see section 2.3.3 for description) is used to define hydraulic parameters in the ll-layer mechanistic hydrology scheme (see Section 2.2.2 and 2.2.3 for a description). For some sites, soil texture fractions are taken from the ancillary Ameriflux BADM (Biological, Ancillary, Disturbance and Metadata) Data Product BIF (BADM Interchange Format) files (see https://ameriflux.lbl.gov/data/aboutdata/badm-data-product/) that are downloaded with the site data. PFT acronyms: BS = Bare soil; TeNE = Temperate Needleleaved Evergreen forest; TeBE = Temperate Broadleaved Evergreen forest; TeBD = Temperate Broadleaved Deciduous forest; C3G = C3 grass; C4G = C4 grass.

Site ID	Description	Dominant Species	Soil texture	Period of site data	PFT fractions	Soil texture class fractions	Maximum LAI	Reference
US-SRM	Shrub encroached C4 grassland / savanna	*Prosopis velutina, Eragrostis lehmanniana*	Deep loamy sands	2004-2015	50% BS; 35% TeBD; 15% C4G	USDA: Loamy sand	0.85 (TeBD & C4G)	Scott et al. (2015); Ameriflux BADM.
US-SRG	C4 grassland	*Eragrostis lehmanniana*	Deep loamy sands	2008-2015	45% BS; 11% TeBD; 44% C4G	USDA: Loamy sand	1.0 (C4G)	Scott et al. (2015)
US-Whs	Shrub-dominated shrubland	*Larrea tridentata, Parthenium incanum, Acacia constricta, Rhus microphylla*	Gravelly sandy loams	2007-2015	57% BS; 40% TeBE; 3% C4G	USDA: Sandy loam	0.6 (TeBE & C4G)	Scott et al. (2015)
US-Wkg	C4 grassland	*Eragrostis lehmanniana, Bouteloua* spp. *Calliandra eriophylla*	Very gravelly, sandy to fine sandy, and clayey loams	2004-2015	60% BS; 3% TeBE; 37% C4G	USDA: Sandy loam	0.85 (C4G)	Scott et al. (2015); Ameriflux BADM.
US-Fuf	Unmanaged ponderosa pine forest	*Pinus ponderosa*	Clay loam	2005-2010	100% TeNE	USDA: Clay loam	2.4	Dore et al. (2010; 2012); Ameriflux BADM.

https://doi.org/10.5194/hess-2019-598
Preprint. Discussion started: 4 December 2019

Hydrology and Earth System Sciences
Discussions
Open Access
EGU

US-Vcp	Unmanaged ponderosa pine forest	*Pinus ponderosa*	Silt loam	2007-2014	100% TeNE	USDA: 100% silt loam	2.4	Anderson-Teixeira et al. (2011)

https://doi.org/10.5194/hess-2019-598
Preprint. Discussion started: 4 December 2019

Hydrology and Earth System Sciences
Discussions

Table 2: Soil moisture measurement depths (and corresponding model layer in brackets – see Table S1).

	US-SRM	US-SRG	US-Whs	US-Wkg	US-Fuf	US-Vcp
Soil moisture depths	2.5-5cm (5) 15-20cm (7) 60-70cm (9)	2.5-5cm (5) 15-20cm (7) 75cm (9)	5cm (6) 15cm (7) 30cm (8)	5cm (6) 15cm (7) 30cm (8)	2cm (4) 20cm (7) 50cm (9)	5cm (6) 20cm (7) 50cm (9)

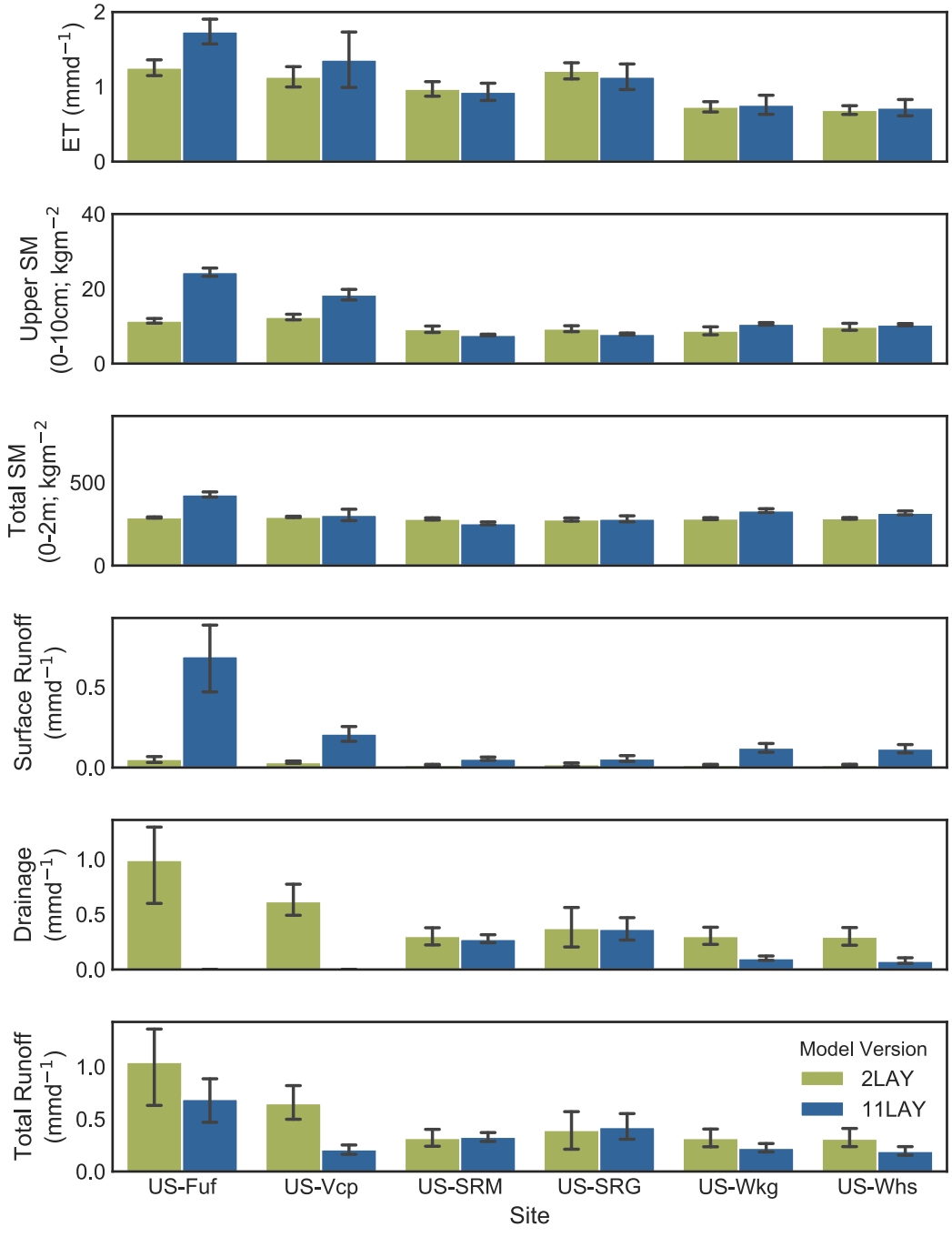
Table 3: Comparison of the 2LAY and 11LAY daily ET model-data evaluation metrics across the whole timeseries (where data present – see Fig. S3). Mean absolute bias = model – observations; therefore, a negative value represents a mean model underestimation of observed ET.

Site	Model Version	Mean Bias (mmd^{-1})	RMSE	R
US-Fuf	2LAY	-0.08	1.04	0.36
	11LAY	0.38	0.86	0.76
US-Vcp	2LAY	-0.54	1.39	0.26
	11LAY	-0.27	1.02	0.59
US-SRM	2LAY	-0.03	0.84	0.53
	11LAY	-0.07	0.53	0.84
US-Whs	2LAY	-0.03	0.68	0.54
	11LAY	-0.02	0.43	0.85
US-SRG	2LAY	0.01	1.02	0.52
	11LAY	-0.11	0.57	0.88
US-Wkg	2LAY	0	0.63	0.62
	11LAY	-0.01	0.37	0.9

https://doi.org/10.5194/hess-2019-598
Preprint. Discussion started: 4 December 2019

Hydrology and Earth System Sciences
Discussions
Open Access
EGU

Figure 1: Comparison of the 2LAY versus 11LAY mean daily hydrological stores and fluxes: i) evapotranspiration (ET – top panel); ii) soil moisture (SM, kgm^{-2}) in the upper 10cm of the soil (2nd panel); iii) total column (0-2m) SM (3rd panel); iv) surface runoff (4th panel); v) drainage (5th panel); and vi) total runoff (surface runoff plus drainage – bottom panel). Error bars show the standard deviation for ET and SM, and 95% confidence interval for runoff and drainage.

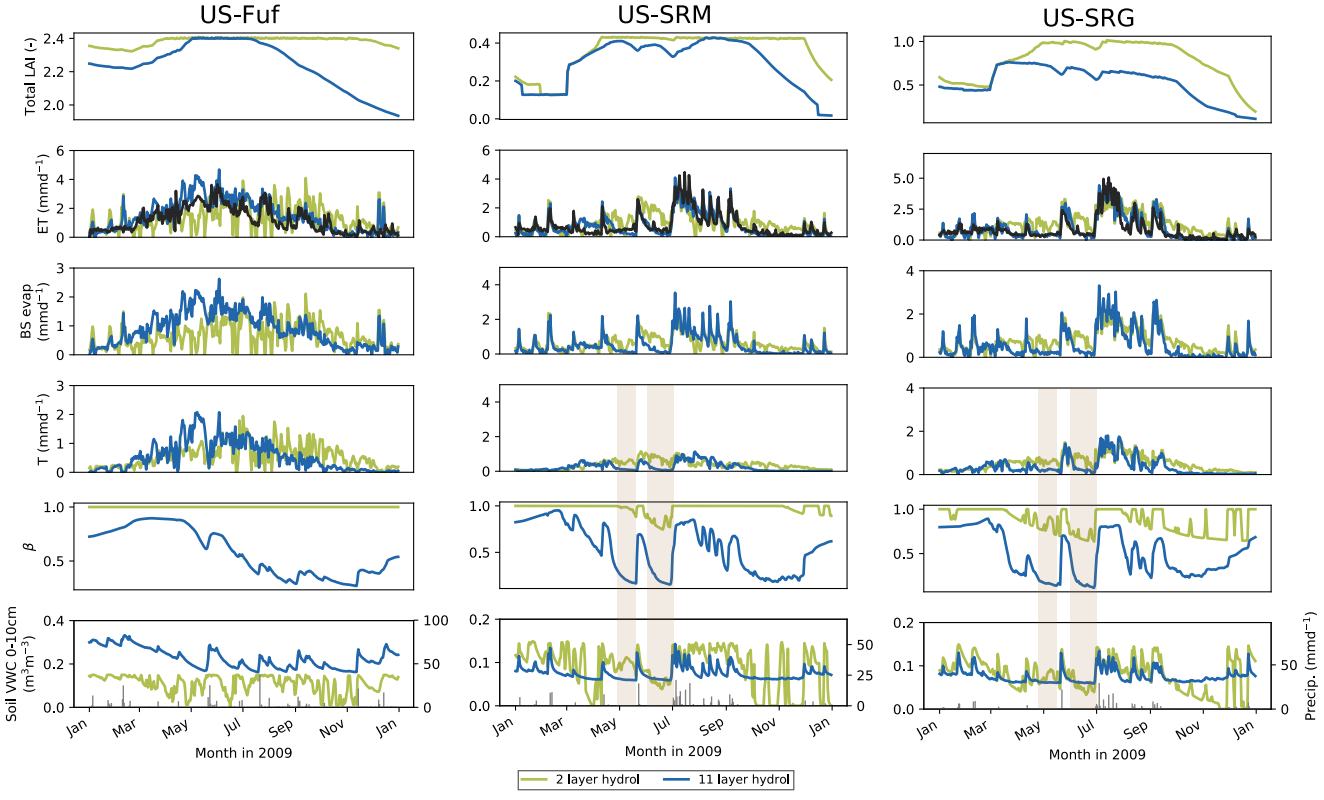


5

https://doi.org/10.5194/hess-2019-598
Preprint. Discussion started: 4 December 2019

Hydrology and Earth System Sciences
Discussions
Open Access
EGU

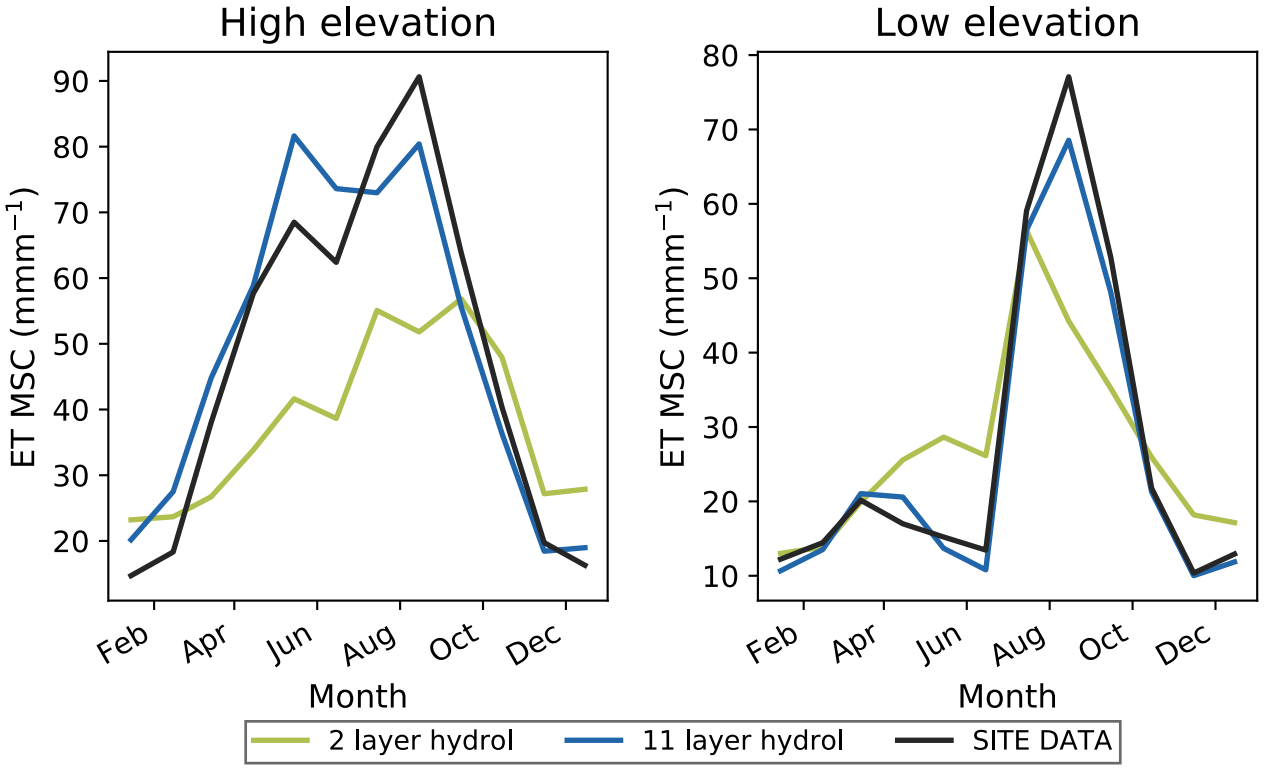
Figure 2: Daily time series (for 2009) of variables influencing changes in ET between the 2LAY (green curve) and 11LAY (blue curve) simulations. Changes are shown for three sites representing the main vegetation types: left column = high-elevation tree-dominated site (US-Fuf); middle column = low-elevation mesquite shrub-dominated site (US-SRM); right column = low-elevation C4 grass site (US-SRG). At each site, top panel: LAI; 2nd panel: ET compared to observations (black curve); 3rd panel: transpiration; 4th panel: bare soil evaporation; 5th panel: empirical water limitation function (β) that scales photosynthesis and stomatal conductance; bottom panel: soil moisture expressed as volumetric water content (VWC) in the uppermost 10cm of the soil. Precipitation is shown in the grey lines in the bottom panel for each site. (Note: full time series across all years are shown for all site in Figs. S2a-f).



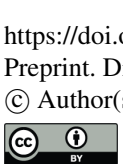
https://doi.org/10.5194/hess-2019-598
Preprint. Discussion started: 4 December 2019

Hydrology and Earth System Sciences
Discussions
Open Access
EGU

Figure 3: Evapotranspiration (ET) monthly mean seasonal cycle comparing the 2LAY (green curve) and 11LAY (blue curve) simulations with observations (black curve). Individual site simulations have been averaged over the high-elevation tree dominated sites (left panel) and across all the low-elevation grass- and shrub-dominated sites (right panel).

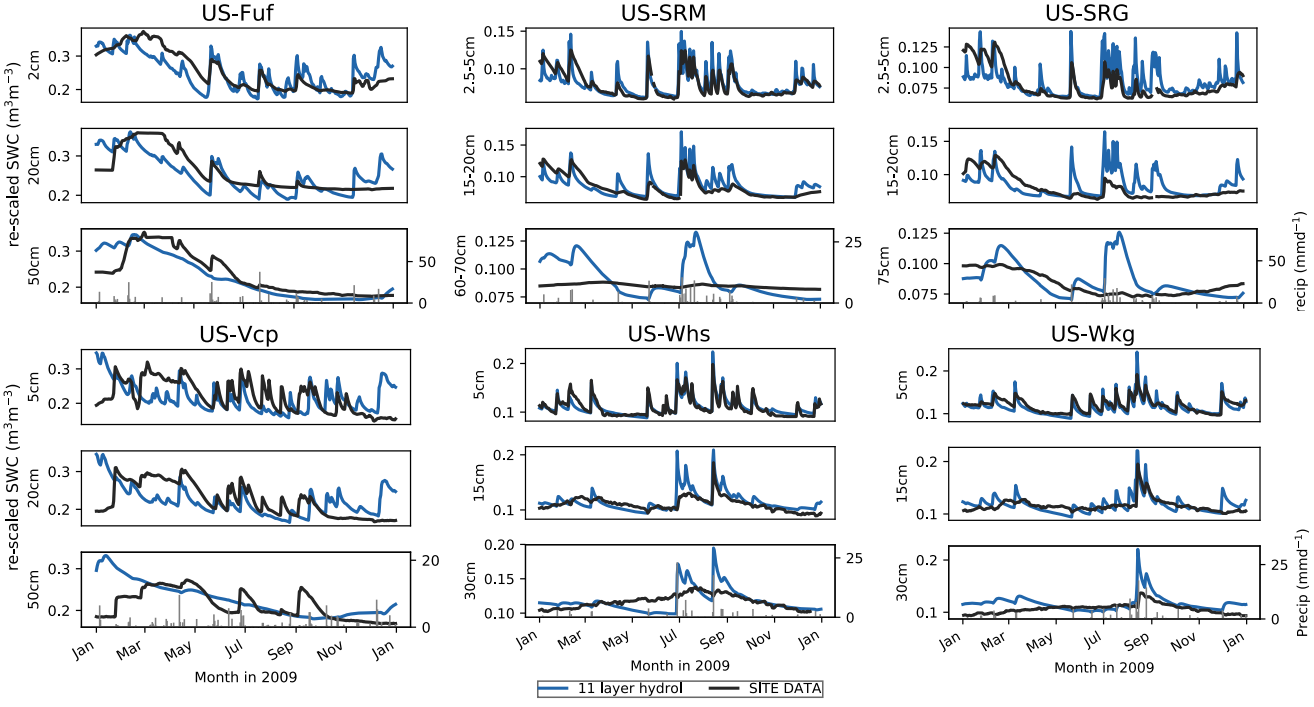


https://doi.org/10.5194/hess-2019-598
Preprint. Discussion started: 4 December 2019



Hydrology and Earth System Sciences
Discussions

Figure 4: Daily simulated volumetric soil water content (VWC – m^3m^{-3}) in 2009 compared to re-scaled (via linear CDF matching) observations at each site for three depths (upper, middle, lower) in the soil profile. The soil depths and their corresponding model layers are given in Table 2. Precipitation is shown in the grey lines in the bottom panel for each site.

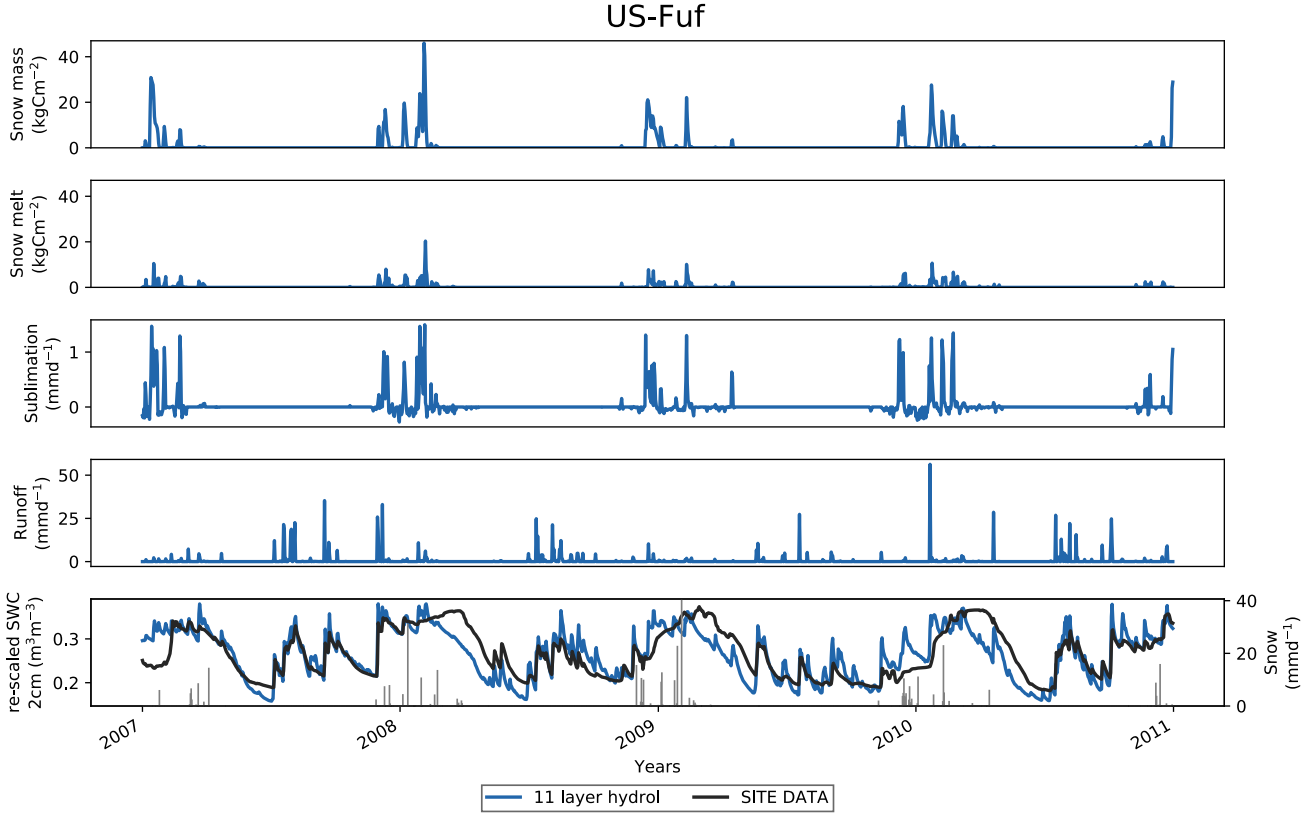


https://doi.org/10.5194/hess-2019-598
Preprint. Discussion started: 4 December 2019

Hydrology and Earth System Sciences
Discussions

Figure 5: a) US-Fuf and b) US-Vcp 11LAY (blue curve) daily time series (2007-2010) of model versus re-scaled (via linear CDF matching) observed volumetric soil water content (SWC – m^3m^{-3}) (black curve), compared to simulated runoff, sublimation, snow melt and snow mass. Snowfall is also shown as grey lines in the bottom panel.

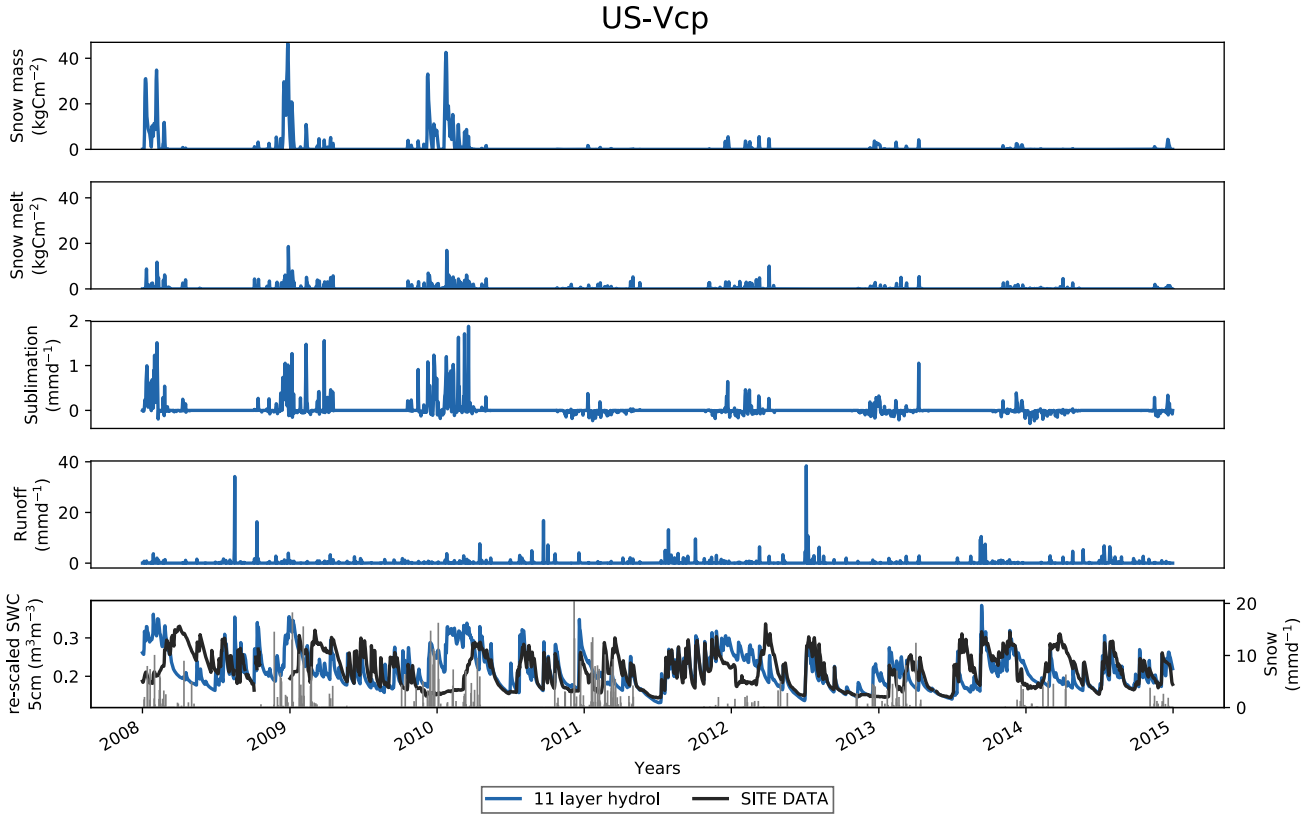
a) US-Fuf



https://doi.org/10.5194/hess-2019-598
Preprint. Discussion started: 4 December 2019

Hydrology and Earth System Sciences
Discussions
Open Access
EGU

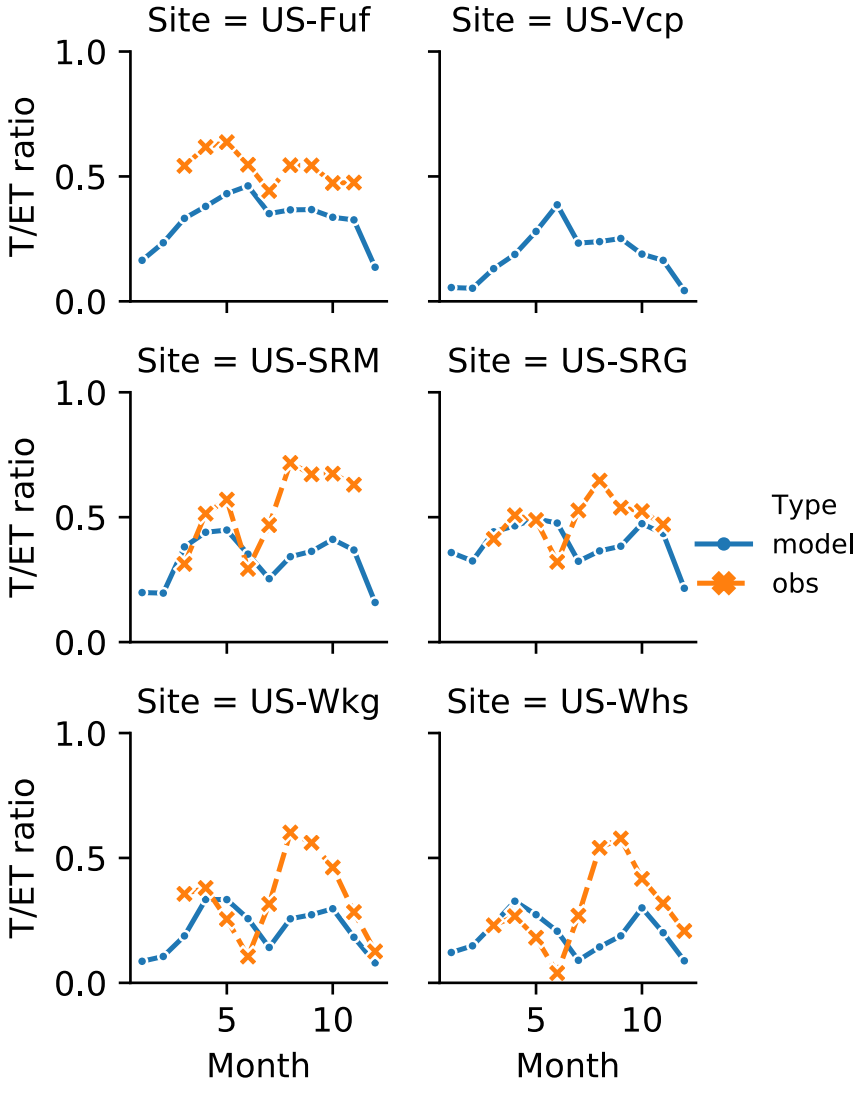
b) US-Vcp



https://doi.org/10.5194/hess-2019-598
Preprint. Discussion started: 4 December 2019

Hydrology and Earth System Sciences
Discussions

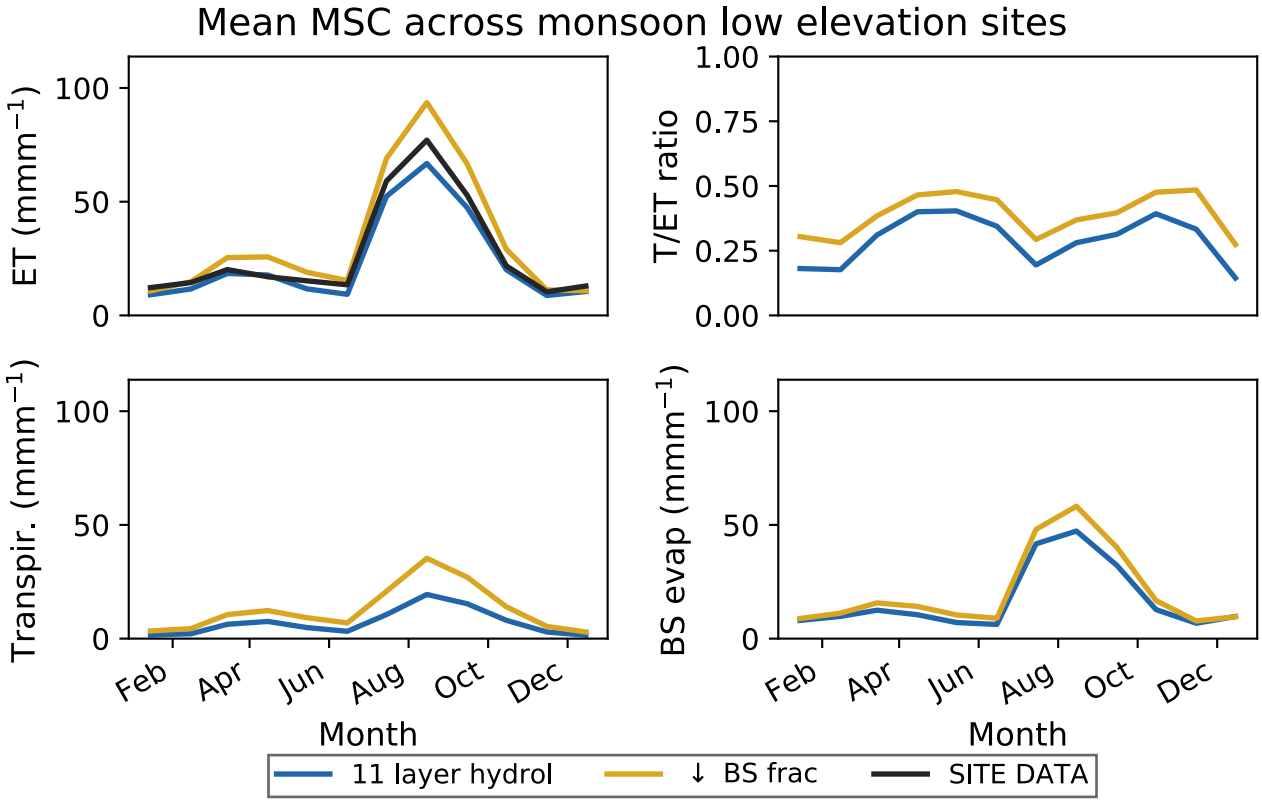
Figure 6: Mean monthly T/ET ratios for each site. Blue curves show the model estimates at each site; orange curves show T/ET ratios at sites/months for which data-driven estimates are available. US-Fuf T/ET ratios are derived using the method of Zhou et al. (2016). Monsoon low-elevation grass- and shrub-dominated site T/ET ratios are derived following Scott et al. (2017).



https://doi.org/10.5194/hess-2019-598
Preprint. Discussion started: 4 December 2019

Hydrology and Earth System Sciences
Discussions

Figure 7: Monthly mean seasonal cycle for ET, T/ET ratios, T and E averaged across all low-elevation grass- and shrub-dominated sites comparing the default 11LAY simulations (blue curve) with a simulation in which bare soil fraction is decreased (C4 grass cover increased (yellow curve). ET is compared to observations (black curve). Units in mmmonth^{-1}.



https://doi.org/10.5194/hess-2019-598
Preprint. Discussion started: 4 December 2019

Hydrology and Earth System Sciences
Discussions

Figure 8: Monthly mean seasonal cycle for evapotranspiration (ET), transpiration and bare soil evaporation averaged across all high-elevation forest sites (left column) and low-elevation monsoon grass- and shrub-dominated sites (right column) comparing the default 11LAY simulations (blue curve) with a simulation that included an additional bare soil evaporation resistance term (red curve). ET is compared to observations (black curve). Units in mmmonth^{-1}.