Multi-variable evaluation of land surface processes in forced and coupled modes reveals new error sources to the simulated water cycle in the IPSL climate model

Hiroki Mizuochi[1,2], Agnès Ducharne[2,3], Frédérique Cheruy[3,4,5], Josefine Ghattas[3], Amen Al-Yaari[2,3,6], Jean-Pierre Wigneron[6], Vladislav Bastrikov[7,a], Philippe Peylin[3,7], Fabienne Maignan[3,7], Nicolas Vuichard[3,7]

[1]National Institute of Advanced Industrial Science and Technology (AIST), Geological Survey of Japan, Tsukuba, 305-8567, Japan
[2]UMR METIS (Milieux environnementaux, transferts et interactions dans les hydrosystèmes et les sols), Sorbonne Université, CNRS, EPHE, Paris, France
[3]IPSL (Institut Pierre Simon Laplace), Sorbonne Université, CNRS, Paris France
[4]LMD (Laboratoire de Météorologie Dynamique), Sorbonne Université, ENS, PSL Université, École polytechnique, Institut Polytechnique de Paris, CNRS, Paris France
[5]ISMAR/CNR Via del Fosso del Cavaliere, 100 00133 ROMA Italy
[6]INRA, UMR 1391 ISPA, Villenave d'Ornon, France
[7]LSCE (Laboratoire des Sciences du Climat et de l'Environnement), UMR 8212 CEA-CNRS-UVSQ, 91191 Gif-sur-Yvette CEDEX, France
[a]Now at : Science Partners, Paris, 75010, France

Correspondence to: Hiroki Mizuochi (mizuochi.hiroki@aist.go.jp)

書式を変更: 英語 (米国)

Abstract. Evaluating land surface models (LSMs) using available observations is important to understand the potential and limitations of current Earth system models in simulating water- and carbon-related variables. To reveal the error sources of a ~~land surface model (~~LSM~~)~~, f~~i~~ve~~our~~ essential climate variables have been evaluated in this paper (i.e., surface soil moisture, evapotranspiration, leaf area index, ~~and~~ surface albedo, and precipitation) via simulations with IPSL LSM ORCHIDEE (Organizing Carbon and Hydrology in Dynamic Ecosystems), particularly focusing on the difference between (i) forced simulations with atmospheric forcing data (WATCH-Forcing-DATA-ERA-Interim: WFDEI) and (ii) coupled simulations with the IPSL atmospheric general circulation model. Results from statistical evaluation using satellite- and ground-based reference data show that ORCHIDEE is well equipped to represent spatiotemporal patterns of all variables in general. However, further analysis against various

landscape/meteorological factors (e.g., plant functional type, slope, precipitation, and irrigation) suggests potential uncertainty relating to freezing/snowmelt, temperate plant phenology, irrigation, as well as contrasted responses between forced and coupled mode simulations. The biases in the simulated variables

35   are amplified in coupled mode via surface–atmosphere interactions, indicating a strong link between irrigation–precipitation and a relatively complex link between precipitation–evapotranspiration that reflects the hydrometeorological regime of the region (energy-limited or water-limited) and snow albedo feedback in mountainous and boreal regions. The different results between forced and coupled modes imply the importance of model evaluation under both modes to isolate potential sources of uncertainty in

40   the model.

1 Introduction

Land surface models (LSMs) are essential to understand large-scale exchange of energy, water, and carbon between the land surface and the atmosphere. LSMs coupled with atmospheric general circulation

45   models (GCMs) have been used to simulate global climate and climate change under international frameworks such as the Coupled Model Intercomparison Project (CMIP) (Taylor et al., 2012; Eyring et al., 2016a), contributing to Earth sciences and policy making for mitigating and adapting to climate change. To understand the potential and limitations of climate change simulations, evaluating outputs of LSMs with available observations is important (Flato et al., 2013). Uncertainties associated with LSMs

50   can arise from a deficiency of model physics and parameterization (Liu et al., 2003), errors in atmospheric forcing data (Guo et al., 2006; Nasonova et al., 2011; Yin et al., 2018), boundary conditions including vegetation and land use changes (Guimberteau et al., 2017; Boisier et al., 2014), and/or error propagation through land-atmosphere coupling (so-called "climate drift") (Dirmeyer, 2001). Recently, convenient tools for systematic model evaluation have been developed (e.g., Eyring et al., 2016c; Gleckler et al.,

55   2016; Best et al., 2015); however, further in-depth model evaluation is required to reveal the underlying processes and sources that lead to uncertainties in simulations (Eyring et al. 2016b).

Notably, focusing on the differences between LSM simulations with and without GCM coupling would provide novel knowledge about LSM evaluation (Liu et al., 2003; Zabel et al., 2012; Wang, T. et

2

al., 2015). LSM simulations without GCM coupling but forced by an atmospheric dataset (also called "offline" or "stand-alone" mode) do not allow feedback between the atmosphere and land surface. Therefore, errors in the simulated values solely arise from deficiency model structure/parameterization, uncertainty in the forcing data (Yin et al., 2018), and mismatch in land cover between model and forcing data (Zabel et al., 2012). The foremost influential forcing factor on water cycle is precipitation (Qian et al., 2006; Decharme and Douville 2006), although radiation and land cover (i.e., vegetation) can also affect hydrological variables (Dirmeyer, 2001; Guo et al., 2006) such as surface soil moisture (SSM) or evapotranspiration (ET), depending on the temporal scale (Guo et al., 2006) and the hydrometeorological condition of the region (i.e., energy-limited or water-limited, Nasonova et al., 2011; Zabel et al., 2012). Anthropogenic factors (e.g., irrigation) may also cause errors in the simulated variables when not accounted for by the LSM (Yin et al., 2018). On the other hand, coupled LSM simulation are also affected by errors in atmospheric simulation, which can be enhanced through land-atmosphere interaction (Mahfouf et al., 1995; Liu et al., 2003; Wang, T. et al., 2015). Such errors occur at short time scale (i.e., several-days) up to seasonal time scale (Dirmeyer, 2001), via the interlinkage of hydrological variables (e.g., rainfall, SSM, ET, and infiltration) in the LSM scheme and thermal variables (Cheruy et al., 2017, AitMesbah et al., 2015).

Among various LSMs, we focused on the Organizing Carbon and Hydrology in Dynamic Ecosystems (ORCHIDEE) LSM (e.g., Krinner et al., 2005; d'Orgeval et al., 2008; Guimberteau et al., 2017), which enables the explicit representation of processes governing the water, carbon, and energy budgets with highly flexible spatial resolution (Raoult et al., 2019). We used the ORCHIDEE (revision 4783, tag 2.0) version, which is implemented in the IPSL's (Institute Pierre Simon Laplace) climate model configurations used for CMIP6 (Eyling et al., 2016a), including the Land Surface, Snow and Soil Moisture Model Intercomparison Project (LS3MIP) with offline simulations (van den Hurk et al., 2016). Through an in-depth assessment of five simulated variables (i.e., SSM, ET, leaf area index (LAI), surface albedo, and precipitation) that should be closely interlinked and a special focus on the differences between forced and coupled simulations, the aim of this study is to better understand which land surface processes deserve further improvements in the studied LSM and to investigate the land-atmosphere coupling role in diagnosed model uncertainties.

3

The Global Climate Observing System (GCOS, 2010) designates the five selected variables as being essential climate variables (ECVs), thereby allowing us to take advantage of recent progress in their global scale observation. Using satellite data, researchers have developed various retrieval algorithms to acquire SSM (Jackson et al., 1999; Wigneron et al., 2007, 2017), ET (Zhang et al., 2010; Miralles et al., 2011; Zeng et al., 2014), LAI (Zhu et al., 2013), albedo (Schaaf et al., 2002; Qu et al., 2014), and precipitation (Adler et al., 2003), which can be used as reference data for LSM evaluation. Empirical upscaling products from global *in situ* observations (Jung et al., 2011, 2019) can also be used. The selected variables are particularly interesting for land surface processes: SSM is a recognized driver of surface–atmosphere interactions (Seneviratne et al., 2010), constraining the partitioning of sensible/latent heat and plant activity and determining ET and vegetation dynamics (e.g., Gu et al., 2006). ET affects atmospheric humidity (usually described by the vapor pressure deficit) and cloud formation, creating feedback systems among SSM, ET, and precipitation (Yang et al., 2018). Accounting for long-term vegetation dynamics, which can be measured by LAI, interlinked with such hydrological processes, is important in monitoring carbon cycle and ecosystem services that are related to climate change (IPCC, 2014) and natural disasters (Adikari and Noro 2010). Another important parameter in the surface energy exchange is the surface albedo, which controls the reflection of incident solar radiation and is interlinked with hydrological processes (especially through surface snow cover) and vegetation dynamics (Bonan, 2008).

To investigate the potential sources of model uncertainty, we considered various landscape factors ("factor analysis") in addition to the traditional statistical evaluation. This work aims at increasing knowledge about the features and limitations of ORCHIDEE and is a practical example of in-depth model evaluation focusing on the differences between forced and coupled modes. The remainder of this paper is organized as follows. Section 2 describes the simulation setting, the reference datasets, and the factor analysis. Section 3 presents results for the spatiotemporal patterns of the model uncertainties and factor analysis. Finally, Sections 4 and 5 provide a discussion and conclusions, respectively.

2 Materials and methods

2.1 Model and simulations

2.1.1 Description of the land surface model

115 ORCHIDEE (ORganizing Carbon and Hydrology in Dynamic EcosystEms) is the LSM used in the IPSL Earth System model (ESM). This global process-based model of the land surface describes the complex links between the terrestrial biosphere and the water and the energy and carbon exchanges between the land surface and the atmosphere (Krinner et al., 2005). The used version in the IPSL-CM6 ESM for the CMIP6 simulations (Boucher et al., 2020), which is known as tag 2.0, was previously described in many

120 papers (Raoult et al., 2019; Boucher et al., 2020; Cheruy et al., 2020; Tafasca et al., 2020), and we only summarize its main features in this paper, with some details on the related parametrizations to the five studied ECVs.

The land cover is described with 15 plant functional types (PFTs), including one for bare soil, as seen in the full list in Table 2, and they can all coexist in each grid-cell, where the fractions taken

125 here are from the CMIP6 datasets (Boucher et al., 2020). For each PFT, the transpiration serves as a coupling flux between the water, energy budget, and photosynthesis process, which drive the evolution of the biomass and LAI owing to generic equations with PFT-specific parameters (Krinner et al., 2005). Evapotranspiration (ET) is controlled by the energy and water budget via a bulk aerodynamic approach, where four parallel fluxes are distinguished: sublimation, interception loss, soil evaporation, and

130 transpiration. In each grid-cell, the first two fluxes proceed at a potential rate from the grid-cell fractions with snow and canopy water, respectively. The soil evaporation and transpiration originate from the complementary snow-free fractions covered by bare soil and vegetation, which depend on LAI, as the effectively foliage-covered fraction exponentially increases with LAI. These two fluxes depend on soil moisture: transpiration is limited by the stomatal resistance, which

135 These two fluxes depend on soil moisture: transpiration is limited by the stomatal resistance, which increases when soil moisture drops from field capacity to wilting point; soil evaporation is not limited by a resistance but only by upward capillary fluxes, which control the soil propensity to meet the evaporation demand.

5

The soil moisture (SM) dynamics are described over a soil depth of 2 m and discretized into 11
140 soil layers to solve the Richards equation. SM in the top 10-centimeters is regarded as SSM. The hydraulic
conductivity and retention properties depend on the SM owing to the Van Genuchten-Mulaem equations,
with parameters depending on the soil texture (Tafasca et al., 2020),which is read from
the map of Zobler (1986). The infiltration is limited by the surface hydraulic conductivity, and it is
calculated with a time splitting procedure inspired by the Green-Ampt equation, where a sharp piston-
145 like wetting front is assumed (d'Orgeval et al., 2008; Vereecken et al., 2019). The surface runoff is made
of non-infiltrated water (infiltration-excess runoff); however, ponding is allowed in flat areas, and it can
reinfiltrate at later time steps. This so-called reinfiltration fraction linearly decreases from
1 in totally flat grid-cells to 0 where the mean grid-cell slope exceeds 0.5%
. For CMIP6, the ORCHIDEE model does not include the irrigation impact on the
150 SM, ET, and vegetation growth, although the model can simulate this anthropogenic pressure
(Xi et al., 2018).

The snow processes are described by a 3-layer scheme of intermediate complexity (Wang et al.,
2013), in which the snow albedo and insulating properties depend on the snow density and age. The
ORCHIDEE 2.0 also includes a revised parameterization of the interplay between the vegetation and the
155 snow albedo, and the optimized parameters match the remote sensing albedo data from the Moderate-
resolution Imaging Spectroradiometer (MODIS) sensor, distinguishing the visible and near-infrared
(NIR) bands (Boucher et al., 2020). For the calculation of the heat diffusion, which
includes the soil freezing effects (permafrost), the soil is extended to 90 m, and the moisture content of
the deepest hydrological layer is extrapolated to the entire profile between 2 and 90 m. The thermal soil
160 properties depend on the soil texture, moisture, and carbon content (Guimberteau et al., 2018).

2.1.2 Forced and coupled simulations

To separate the errors caused by the ORCHIDEE model structure/parameterization from the ones
resulting from the simulated climate through land-atmosphere coupling, we compared a forced and a
165 coupled simulation. In the coupled simulation, the ORCHIDEE LSM is coupled to the LMDZ6A
atmospheric GCM (Hourdin et al., 2020), as embedded in the IPSL-CM6 ESM for the CMIP6 simulations

6

(Boucher et al., 2020; Cheruy et al., 2020). The only difference between the atmospheric physics used in this paper and the one used for CMIP6 concern the parameterization of deep and shallow convection and their interaction, modified to improve the description of the intertropical convergence zone and the El Nino Southern Oscillation. The description of these differences and their impact on precipitation and other variables controlling the near surface climate can be found in Mignot et al., (2021).

The coupled simulation was run over 1985–2014 (following a 5-yr spin up) using a 'nudging' approach to constrain the large-scale atmosphere dynamics toward the synoptic atmospheric conditions (Cheruy et al., 2013). To this end, the simulated wind fields (zonal and meridional wind components) are relaxed toward the ERA-Interim winds (Dee et al., 2011) by adding a correction term in the evolution equation for the wind. By reducing the internal variability, this method allows the direct comparison of the observations and simulations, and it was successfully used for evaluating the coupled land-atmosphere parameterizations (Cheruy et al., 2013; Wang, et al., 2016), including with the IPSL-CM6 ESM (Cheruy et al., 2020).

In the forced simulation, which covers 1979–2009, the required near surface meteorological data by the ORCHIDEE LSM (liquid and solid precipitation, incoming longwave and shortwave radiation, 2-m air temperature and specific humidity, 10-m wind speed, surface pressure) are prescribed from the downscaled and bias-corrected reanalysis data [WATCH-Forcing-DATA-ERA-Interim (WFDEI)], provided at the 0.5° resolution with a 3-hourly time step (Weedon et al., 2011; Weedon et al., 2014). Precipitation is bias-corrected using monthly data from the Global Precipitation Climatology Centre (GPCC Version V6, Schneider et al., 2014), with a specific correction of undercatch errors following Adam and Lettenmaier (2003).

The spatial resolution differs between the two simulations, reflecting the grid of the atmospheric data: the coupled simulation has a coarser resolution (144 × 142, corresponding roughly to 2.5° in longitude and 1.25° in latitude) than that of the forced simulation (0.5° grid). To make the evaluation consistent and simple, we used the same spatial resolution for our analyses, and we oversampled the LMDz grid mesh to the finer resolution (0.5°) such as to keep as much spatial information as possible from the high-resolution offline grid mesh. To investigate variability patterns on seasonal to interannual scales, all the data were aggregated into monthly time steps. Five interlinked variables (SSM, ET, LAI,

7

195 albedo, and precipitation) were considered in this evaluation, and the study region was 60°S–90°N, 180°W–180°E (i.e., Antarctica and Greenland were excluded).

2.2 Reference data

2.2.1 Surface soil moisture

The SSM product provided by European Space Agency Climate Change Initiative (ESA CCI) (Liu et al.,
200 2012) was used as a reference. It is a merged product comprising multiple SSM data derived from various passive and active microwave satellites (i.e., SMMR, SSM/I, TMI, AMSR-E, Windsat, SMOS, AMSR2, AMI-WS, ASCAT-A, and ASCAT-B) providing a long-term (1978–2018) SSM dataset with 0.25° resolution. The CCI-SSM product has been evaluated extensively against *in situ* observations (e.g., Al-Yaari et al., 2019b), and their accuracy has been reported as being relatively high compared to that of
205 other existing products such as SMOS-L3, LPRM, and AMSR2 (Ma et al., 2019).

Because it includes low-quality data flags for snow, dense vegetation, and radio-frequency interference (RFI) (Oliva et al., 2012), we applied data screening following Al-Yaari et al. (2016). We screened out all the pixels where the provided uncertainty was larger than 0.06 m^3/m^3 (volumetric water content). Next, any data records in which the SSM was not in a valid range (either >0.6 or <0.0)
210 (Fernandez-Moran et al., 2017; Dorigo et al., 2013) were excluded. Finally, to exclude any areas covered by snow or dense vegetation and other unreliable regions, we kept only those areas in which the quality flag was zero (fine-quality pixels). The screened dataset was then aggregated into $0.5° \times 0.5°$ and monthly time steps. This screening process removed 3.6% of all the original pixels.

We performed an initial check on the time series of the global average of CCI-SSM and found an
215 artificial trend therein that depended on the availability of the observation data (Supplementary Fig. S6). As reported by other researchers (e.g., Loew et al., 2013), this artificial trend could lead to misinterpretation of long-term signals. To mitigate such artificial trends and initialization errors of each data, we selected a stable period (i.e., without discontinuous jump in time series) during 1993–1999 for comparison with both the forced and coupled simulations. Because of the differing natures of LSM-
220 simulated and observed SSM (e.g., dependence on meteorological forcing data/atmospheric model, model parameterization), their absolute SSM [m^3/m^3] values (i.e., magnitudes) are not comparable (Reichle et al.,

8

2004). In addition, since the CCI-SSM product is scaled by the comparison with a different LSM (GLDAS-Noah), a direct comparison between CCI-SSM and ORCHIDEE may lead to misleading results, as they have different soil representation (Raoult et al., 2019). Given this issue, the LSM- and satellite-based SSM were compared with statistically normalized values rather than absolute values of SSM (e.g., Polcher et al., 2016). Therefore, a spatiotemporal normalization (Equation 1) was applied to each co-masked dataset to eliminate systematic biases among the datasets and make the comparison reliable (Supplementary Fig. S7B):

$$SSM_{norm} = \frac{SSM - \overline{SSM}}{\sigma_{SSM}}, \qquad (1)$$

where SSM_{norm} is the normalized SSM, and $\overline{SSM}$ and σ_{SSM} are the mean and standard deviation, respectively, of all the available SSM sampled along spatial and temporal dimension during the period.

2.2.2 ET

In a preliminary study, we compared a ground-based machine-learning ET product (Jung et al., 2011; 2019), three remote-sensing-based physical model products (Miralles et al., 2011; Zhang et al., 2010; Zeng et al., 2014), and their ensemble (see Supplementary Figs. S2, S7). We found that they showed similar spatiotemporal structures although they differed in absolute values in some regions, and the ground-based product was the most consistent with the ensemble. Therefore, we decided to use the ground-based ET (mm/d) as a representative, from 1987 to 2009. It is also advantageous in that it is derived from upscaling of FLUXNET data (Jung et al., 2011; 2019) and is independent from specific ET-retrieval algorithms. The original spatial resolution (1°) of the data was resampled into 0.5° resolution to match that of forced simulation, and original temporal resolution was monthly time steps. A preliminary check of the time series and spatial patterns of the reference data revealed no artifact patterns (e.g., no abrupt jump in time series as found in CCI-SSM), so we used them with no pixel screening or normalization.

2.2.3 LAI

We used the global LAI dataset of Zhu et al. (2013), referred to hereinafter as LAI3g, which is based on a neural-network algorithm in conjunction with the third-generation Global Inventory Modeling and Mapping Studies (GIMMS 3g) and MODIS LAI

9

product, with an original spatial resolution was 0.5° and half-monthly temporal resolution. Considering
250 the common period among LAI3g and the coupled/forced simulations, we selected 1987–2009 as the
comparison period, and we resampled all data at 0.5° spatial resolution and aggregated them into monthly
time steps.

2.2.4 Albedo

We used the MODIS albedo product (Qu et al., 2014) as reference data, which provided the bi-
255 hemispherical reflectance (white-sky albedo) for the visible and NIR bands. The original 500-m spatial
resolution and 16-d temporal resolution were resampled (i.e., upscaled) into 0.5° resolution and monthly
time steps. The common period between simulations and observation, 2003–2009, was used for
evaluation. The pixels with retrieval failure of albedo were excluded from the analysis.

2.2.5 Precipitation

260 We also evaluated the simulated precipitation because it is the primary
factor that influences the hydrological variables (Qian et al., 2006; Decharme & Douville, 2006). To this
end, we used the GPCC dataset Version V6 (Schneider et al., 2014), which was also used to
bias-correct the WFDEI meteorological forcing of the offline ORCHIDEE simulation
(section 2.1.2). This gridded product at 0.5° provides monthly precipitation derived from quality-
265 controlled observed precipitation from over 65,000 stations world-wide, and accounts for a
climatological correction of undercatch based on Legates and Willmott (1990).

2.2.6 Data processing

270 For consistency between the observed and simulated data, we subjected the former to aggregation or
resampling toward the 0.5° spatial resolution and monthly time steps for each variable, as described
above. Due to the presence of data gaps in the reference data sets, which are either because of the
acquisition issues or the quality control and data screening, we masked the simulated datasets to match
the spatial-temporal data availability of the corresponding reference data. For the SSM, the dense snow

10

275 regions (with a snow water equivalent exceeding 48 mm) in the simulated data were further excluded so as to avoid unreliable comparisons with uncertain references. Also, co-masking was performed after the spatiotemporal resampling, followed by the statistical normalization (only for the SSM). The resulting coverage of the selected comparison period is summarized in Table 1 for each variable.

After the above-mentioned pre-processing, to compare the spatial patterns of the observed and
280 simulated data, we focus on three accuracy criteria calculated at the 0.5° scale along monthly time steps: the bias, Pearson's correlation coefficient (CC), and root-mean-square error (RMSE). The criteria were calculated along temporal axis for each pixel (i.e., the result was shown as one global map for a criterion). The statistical significance of the bias (compared to zero) and CC was assessed at each pixel with Student's t-test and Pearson's test, respectively, with a p-value of 5% in both cases. Note that the
285 evaluation periods were different among SSM (1993–1999), ET, LAI and precipitation (1987–2009), and albedo (2003–2009). However, the impact of the chosen period on the evaluation is likely to be limited (see Supplementary Table S1).

2.3 Factor analysis

To reveal features of the simulations in detail, the accuracy criteria were evaluated against various
290 landscape/meteorological factors (Figure 1), namely PFT, LAI, irrigation, precipitation, slope, snow, and ET. For each factor, time series were averaged temporally to make only one global map (i.e., the classification criteria were applied on long-term basis). The value of each factor was classified into a specific number of levels (classes), which were used as ordinal scales. Each factor was classified as given in Table 2, and each factor is described in detail below:
295 1) For PFT, we used the input dataset that is used in ORCHIDEE. This includes fractional coverages in each pixel of 15 PFTs. We created a dominant PFT map by picking up the PFT class that have maximum fractional coverage for each pixel.

2) For LAI, we used the LAI3g data (above subsection), classifying them into three levels (see Table 1 for the specific class definitions).

11

300 3) For irrigation, we used a global map of irrigation areas (Siebert et al., 2010), which indicates the fractional coverage (%) of an irrigated area with 5-arc-min spatial resolution. It was classified into six levels.

4) For precipitation, we used the pluri-annual mean of GPCC during the same period as the investigated ECVs. It was classified into five levels.

305 5) For slope, this classification was done by referring to the ETOPO5 DEM (5 arc-minute global relief model of Earth's surface; NOAA, 1988), which is also used in ORCHIDEE to control reinfiltration of the water.

6) We used the pluri-annual mean of the forced SWE for the factor analysis of the forced simulation, and that of the coupled SWE for the coupled simulation. 310 The SWE was classified into five levels.

7) For ET, we used the pluri-annual mean of Jung et al. (2011; 2019) during the same period as the investigated ECVs.

Here, the dominant PFTs, irrigation, slope, and SWE were only used for the factor analysis, while LAI, precipitation, and ET were also used for validation and come from independent sources. The PFT 315 fractions and SWE, however, were not independent from our simulations, but we assumed it was not problematic for factor analysis, which mostly aims at suggesting process-based explanations to the main model errors.

3 Results

3.1 Spatial and temporal patterns of model errors

320 Overall, the spatial structures of the ECVs simulated in both modes were consistent with those of the reference products, as shown by comparing the corresponding pluri-annual mean maps (Supplementary Figures S1-S5). To refine this comparison, Figure 2 shows the spatial bias patterns of the five variables (normalized SSM, ET, LAI, albedo, and precipitation), in both forced and coupled modes. Difference between forced and coupled modes are also 325 shown in Supplementary Figure S6. Spatiotemporal averages of bias, RMSE, and correlation coefficients

are summarized in Table 3. The spatial patterns of the temporal CC are also shown for SSM and albedo (Figure 3) for further discussion.

The precipitation bias in forced mode is very small in most regions since the simulation relies on bias-corrected precipitation (WFDEI), which relies on the GPCC dataset used here as reference data. Yet, it is not everywhere negligible (Figure 2A), and the forced ORCHIDEE precipitation (WFDEI) is higher than GPCC in small tropical pockets, the US Great Plains, and boreal zones, which are prone to precipitation undercatch because of strong winds and/or a large fraction of snowfall (Becker et al., 2013). The largest precipitation biases in coupled mode, in absolute value, are found in the wettest areas (humid tropics) and mountain ranges (Figure 2B), which is consistent with the analysis of Cheruy et al. (2020), in terms of bias sign and spatial pattern.

Figure 2C–D showed that the spatial pattern of normalized SSM bias in forced and coupled modes were consistent and delineated the biased regions clearly. The strong negative biases in normalized SSM was observed over the boreal region (except Eastern Siberia) with high SWE values (Figure 1J), suggesting the relation to snow or permafrost. Note that satellite observation uncertainties in such snowy regions could also be a reason for the discrepancy. The farm belt of India and China (with a lot of irrigation in Figure 1H) exhibit a systematic lower bias in SSM. Apart from those, arid (North Africa, middle of Australia, North China) and tropical (Congo and Amazon Basin) regions also showed lower correlation (Figure 3A–C), part of which can be attributed to the inherent feature that CC tends to be low when the range in which the sample varies is narrow. To better identify the error sources in SSM, we plotted the mean seasonal cycles (i.e., monthly climatology) separately for each latitude zone (Figure 4). Substantial parts of the time series were consistent between simulation and observation (except grayed-out period due to insufficient sample size and low reliability of the reference data). The underestimated simulated SSM values compared to the CCI-SSM values in the summer season in 30–60°N (Figure 4B) may be attributed to anthropogenic water input due to irrigation because this region includes large-scale agricultural fields (Figure 1H). In the low-latitude regions, the simulated values tend to underestimate SSM in the dry season, and to show larger seasonal change (Figure 4C, D).

Positive bias of LAI was observed in large areas globally (Figure 2G, H). Given the strong similarity between the forced and the coupled bias maps, it is suggested that the bias comes mostly from

the surface component, such as PFT maps, or reference data itself. In fact, LAI retrievals by spaceborne
355   sensors like MODIS may be saturated for large values of LAI (Zhao et al., 2016), resulting in
underestimation of LAI in reference. Despite such a positive bias tendency, the boreal region in Eastern
Siberia, the shores of the Great Lakes in North America, and the basin of the Mekong River all exhibited
negative bias of LAI. In addition, there were hotspots of negatively biased LAI in such regions as the
Zambezi River system lying across Angola and Zambia. Contrasting biases between the simulations were
360   observed around the Himalayas.

In most regions, the simulated bias for total albedo was small, and the bias patterns
were generally similar between the forced and coupled modes
(Figure 2I, J). The largest biases were the overestimation in the mountainous regions (especially the
(Figure 2I, J). The largest biases were the overestimation in the mountainous regions
365   (especially the Himalayas in the coupled mode) and the underestimation in the boreal and polar regions,
where snow affects the albedo. In addition, simulated and observed albedo were uncorrelated (or
negatively correlated) in many regions apart from the boreal one (Figure 3C–F). Low correlation
coefficients in the arid and tropical region can be attributed to the temporal invariance of the land surface.
However, even in some temperate and semi-arid zones where temporal variance is likely to be high, low
370   correlation was observed. In such regions, seasonal changes of the land surface (caused mainly by
vegetation phenology and the snowfall/snowmelt cycle) may not be described well in ORCHIDEE. In
fact, the global monthly climatology (Figure 5A, F) showed that global mean NIR albedo
was overestimated all year long except in the spring, when visible
albedo was underestimated. The main source of the NIR albedo overestimation seemed to come
375   from the temperate zone (30–60°N; Figure 5C), suggesting overestimated vegetation cover
(having high reflectance in NIR spectral region) there from summer to autumn. There was
a systematic overestimation of albedo in the tropical band (Figure 5D, E, I, J), and a small underestimation
in the snow-related season (winter to spring) of the boreal band (Figure 5B, G).

14

3.2 Factor analysis

380   The bivariate linear regressions between simulated ECV bias and factors (Table 4), and the boxplots against each factor class (Figures 6-8) firstly reveal a large bias variability within each class, resulting in a large part from the spatial variability of the simulated variables across the various climates and biomes of the globe. However, some controls could be identified despite this variability. It is particularly the case for irrigation, which has an obvious impact on the simulated hydrological variables (SSM, ET, LAI, and

385   precipitation; Figure 6A, C, E, G): both the coupled and forced models show negatively biased values in the largely irrigated areas (classes 5 and 6), except for the forced mode SSM. This is understandable because the simulations overlook irrigation, which creates artificial water input to the soil, resulting in additional ET and plant growth in reality. Interestingly, the coupled simulation underestimates the observed values more than does the forced one (Figure 6A, C, E, G), which probably relates to a positive

390   feedback driven by surface–atmosphere coupling (Mahfouf et al., 1995; Liu et al., 2003; Wang, T. et al., 2015). Since the forcing WFDEI precipitation is based on *in situ* rain gauges which integrate the impact of the real-world irrigation, this factor has a relatively weak effect in the forced mode.

        The contrasting ET bias pattern between forced and coupled modes in the Congo and the Amazon

395   (Figure 2E, F) was also confirmed in the factor analysis of precipitation (classes 4 and 5, which probably correspond to tropical regions; Figure 6D), PFT (class 2: broadleaf evergreen in Figure 7A), LAI (class 3 in Figure 7B), and ET (class 3 in Figure 8A). This also explains the contrasting correlation sign of ET bias with P, SSM, ET and LAI in Table 4.

        The factor analysis confirms the positive bias of LAI in the tropical regions, which are

400   characterized by high precipitation (classes 4 and 5 in Figure 6F), broadleaf evergreen forest (PFT 2 in Figure 7C), high LAI (class 3 in Figure 7D), and high ET (class 4 in Figure 8B). However, some of the positive bias in such tropical regions might be compensated by the negative bias of the simulated precipitation (especially in the Amazon; Figure 2B, also confirmed by class 3 in Figure 7J), resulting in a smaller bias of LAI in the coupled simulation than that in the forced simulation. Negative LAI bias in

405   the boreal region is also confirmed by the PFT factor analysis (classes 8, 9, and 15 in Figure 7C). Eastern Siberia is the main place with negative LAI biases (Figure 2G, H). Possible explanations include persistent

15

snowpack reducing the vegetation growing season; underestimated maximum LAI in the model; and errors in the reference LAI product especially at high latitudes, due to the less reliable assessment of solar reflectance from space (Guimberteau et al., 2018).

410   For albedo, the effect appeared in the factor analysis against slope (class 3 in Figure 8C, D) and SWE (classes 4 and 5 in Figure 8F, G) as a discrepancy between the coupled and forced simulations. In the steep regions, the coupled simulation tended to be positively biased because of the precipitation bias. In the high-SWE region, the negative bias of albedo was enhanced in the forced simulation. This is due to the already mentioned compensation of the positive bias in the mountainous region with the negative

415   bias in the boreal and polar regions (Figure 2I, J). The NIR albedo in the tropical region (classes 2 and 3 in Figure 7E and class 3 in Figure 7F) tended to be slightly large for both simulations. This is consistent with the positive bias of LAI in such regions (Figure 2G, H), although the range of bias was small.

4 Discussion

In general, the ORCHIDEE simulations show good spatial/temporal consistency with the reference data,

420   except for issues related to external water addition/subtraction and surface–atmosphere coupling. An example of the external source of water input is irrigation. Largely irrigated areas obviously lead to underestimated hydrology-related model parameters (i.e., SM, ET, and LAI). Although the impact of irrigation on ORCHIDEE SSM simulation has been suggested by Yin et al. (2019) over a specific region (China), our experiment demonstrated explicitly that the effect on SSM in the forced mode is relatively

425   small on the global scale, and rather larger on ET and LAI (Figure 6A, C, E). Integrating the irrigation process in ORCHIDEE with an ancillary agricultural map and data assimilation (Raoult et al., 2019) may improve the accuracy (de Rosnay et al., 2003). Through the land-atmosphere coupling (Al-Yaari et al., 2019a), the impact of the irrigation is emphasized in the coupled simulation (Figure 6A, C, E, G), where strong negative bias was observed in not only ET, LAI, and precipitation but also SSM over largely

430   irrigated areas. Specifically, a lack of description of the additional water input and man-made vegetation over irrigated agricultural land led to lower SSM and LAI, which in turn led to lower ET. In the coupled simulation, the lower SSM also led to lower humidity and lower precipitation, resulting in enhanced underestimation of SSM in the next time step (i.e., positive feedback). The enhanced SSM

16

underestimation caused enhanced ET underestimation, as well as enhanced LAI underestimation through the parametrizations of carbon assimilation and vegetation phenology. The underestimation of precipitation in the coupled simulation over irrigated areas (e.g., India in Figure 2B; classes 5 and 6 in Figure 6G) supports the validity of this emphasizing effect in the coupled model, which is consistent with other reports (Mahfouf et al., 1995; Liu et al., 2003; Wang, T. et al., 2015). The spatial similarity between the bias maps of SSM, ET, and precipitation (Figure 2A–F) over central-south Africa, Australia, and a large part of south and east Asia also suggests a strong interlink between them in coupled mode. A potential interpretation is that precipitation is the first-order control on SSM and ET in the region (i.e., water-limited ET). It also can be interpreted as a result from a positive feedback between precipitation, SSM, and ET in such regions, as reported by Yang et al. (2018).

However, other factors need to be considered regarding land-atmosphere feedback and their influence on coupled precipitation and ET. In particular, ET is not controlled solely by precipitation but also by radiation (Cheruy et al., 2020), and temperature determines the potential ET (Dirmeyer, 2001; Nasonova et al., 2011). The complex response of ET to precipitation in the present study suggests the importance of those factors.

For example, there may be a negative feedback in the Amazon and the maritime continent between precipitation and ET because these areas are strongly energy-limited (Seneviratne et al., 2010; McVicar et al., 2012). In the maritime continent, positive precipitation bias meant more cloud coverage than reality, which decreased the available energy and ET. In contrast, the negative precipitation bias in the Amazon meant less cloud coverage, larger available energy, and larger ET than in reality.

Although such feedback explains the overestimated ET in the Congo and the Amazon in the coupled simulation, it does not explain the underestimated ET there in the forced simulation.

17

Potential explanations are excessive water stress on ET, insufficient soil water holding capacity, underestimated precipitation or radiation in WFDEI, and/or an overestimation of ET by the Jung product, in the regions of high precipitation. Conversely, too weak water stress in dry areas (either for transpiration or soil evaporation) can also explain the negative correlation between forced-mode ET bias and precipitation (Table 4). A solution could be to activate a resistance to soil evaporation, increasing with the top soil dryness (Cheruy et al., 2020). Such contrasting results between the forced and coupled modes imply the importance of model evaluation under both modes to isolate the potential error sources.

Compared with the forced mode, the positively biased precipitation simulated by the coupled mode may positively bias the albedo, particularly in the mountainous areas (high slope class in Figure 8C–E) via considerable snow cover. This probably arose from incomplete atmospheric simulation of the local climate (Cheruy et al., 2020) such as an updrift along a mountain slope. Coarse spatial resolution of the atmospheric simulation in the coupled mode can also make it difficult to represent the impact of mountainous topography on local climate (Decharme and Douville, 2006). Theoretically, the overestimation of albedo should decrease the available energy at the surface, thereby decreasing ET and surface temperature. The slight negative bias of ET in the Himalayas (Figure 2F) despite the positive bias of precipitation (Figure 2B) can be explained by the decrease in available energy due to the increased albedo (Figure 2J). Such an ice–albedo interaction in the ORCHIDEE-LMDZ coupled mode has also been reported over the boreal region (Wang, T. et al., 2015; particularly pronounced in spring temperature over Eastern Siberia). Taking the ice–albedo feedback into consideration with the secondary factors (i.e., radiation and temperature) that affect ET, the link between precipitation and ET in the coupled mode is rather complex in the mountainous and boreal regions. Moreover, the deficit of available energy may reduce photosynthesis thus vegetation growth, causing a peaky underestimation of LAI in the Himalayas in the coupled mode (Figure 2H), which is not observed in the forced mode (Figure 2G).

Part of the positive biases in normalized SSM in the Eastern Siberia and polar region (Figure 2C, D) may be attributed to freezing/snowmelt and related vegetation phenology, as excessively large or fast snowmelt occurs in the spring in ORCHIDEE (Figure 5B, G). However, other control factors are likely involved in SSM overestimation, such as wetlands, permafrost, and albedo. The negative albedo bias found in the

18

490　boreal zone (Figure 2I, J) in spite of the positive snowfall bias (Figure 2A, B) can also be explained by the excessive snowmelt. This underestimation of albedo in many boreal areas, also noted in Cheruy et al. (2020), was expected to lead to overestimated ET, but it did not lead to an obvious ET bias because of the underestimated LAI (Figure 2G, H). Given the spectral features of land cover (Petty, 2006), the NIR albedo is related largely to an abundance of vegetation, i.e., LAI. Therefore, uncertainty

495　in snow and LAI leads to uncertainty in the surface albedo, which further propagates uncertainty in the energy balance and water cycles. Such a complicated relationship should be treated in the special tuning of ORCHIDEE for high latitudes (Druel et al., 2017; Guimberteau et al., 2018). In addition to high latitude regions, vegetation seasonality in the temperate zones seemed uncertain. In the temperate forests, the model is likely to simulate spring green-up that is considerable and/or very fast (Figure 5C), which

500　causes an overestimated NIR albedo and discrepancies in LAI and albedo seasonality.

　　　　Regardless of the origin (i.e., satellite, reanalysis, or *in situ*), observations inevitably contain inherent uncertainty, which leads to uncertainty in the model assessment. SSM retrieval over substantially high/low vegetation, tropical/arid regions, and highly heterogeneous and high-roughness regions remains challenging (Ma et al., 2019). Therefore, some part of the low SSM correlation in arid/tropical regions

505　(Figure 3A, B) can be attributed to uncertainties in the satellite products in addition to an inherent feature of CC. Snow cover and RFI (Oliva et al., 2012) may also cause uncertainties in satellite-based SSM estimation, although we attempted to remove such uncertain pixels by means of a preliminary quality check. Using multiple data sources (e.g., the Soil Moisture Active Passive (SMAP) product (Entekhabi et al., 2010)) as reference for model evaluation (Eyring et al.,

510　2016b) is a promising way to address such uncertainties. A brief attempt with SMOS-IC product (Fernandez-Moran et al., 2017) was shown in Supplementary Fig. S8. Inconsistency between the model-simulated SSM depth (up to 10 cm) and the penetration depth of satellite sensors (several centimeters) may also cause uncertainties in the assessment, although using normalized SSM instead of absolute SSM is likely to mitigate the effect to some extent.

515　The satellite-based LAI product (Zhu et al., 2013) may be affected by the saturation issue of optical satellite data (i.e., MODIS) in regions with high LAI. The snow albedo of the MODIS product (MCD43) has a slightly larger uncertainty (RMSE ≈ 0.07) (Stroeve et al., 2005; Stroeve et al., 2013) than

19

that of the snow-free daily mean albedo (RMSE = 0.034) (Wang, D. et al., 2015). However, this does not alter our conclusion about the ORCHIDEE albedo uncertainty in the snow region, but some of the uncertainty might be attributed to the error in satellite observation.

We depended largely on satellite-derived data for the SSM, LAI, and albedo evaluations. By contrast, we used a FLUXNET-based product (Jung et al., 2019) for the ET evaluation, which has potential uncertainties arising from (i) the statistical upscaling process (model tree ensemble: Jung et al., 2009), (ii) the input data required in machine-learning prediction, and (iii) the heterogeneous distribution of ground stations. Because the latter potential issue is particularly important for hardly accessible regions such as tropical and mountainous areas, progress in the data coverage of the FLUXNET network is desirable. Although ET products derived from satellite data (Miralles et al., 2011; Zhang et al., 2010; Zeng et al., 2014) can also be used, unlike the other variables (SSM, LAI, and albedo), the retrieval of ET is not done directly from the satellite observations but depends largely on the process-oriented models. Therefore, in addition to the uncertainties in the satellite observations themselves, such products have uncertainties that arise from ancillary data (e.g., atmospheric conditions, land cover) required in the model, as well as from imperfections in the model structure/parameterization (preliminary comparison among the different data sources can be found in Supplementary Fig. S9).

The difference between the forced ORCHIDEE precipitation (WFDEI) and GPCC (Figure 2A) probably comes from undercatch correction: based on Legates and Willmott (1990) for GPCC, and on Adam and Lettenmeier (2003) for WFDEI. Schneider et al. (2014) acknowledge that for the GPCC product, "the biggest uncertainty issue is the correction of the systematic gauge-measuring error (general undercatch of the true precipitation)", but this is very likely true for all precipitation products.

Note that the present study is based on a specific LSM (i.e., ORCHIDEE 2.0), atmospheric model (i.e., LMDZ6A), and forcing data (WFDEI). Future work should include addressing the uncertainties that arise from the LMDZ model structure/parameterization, as well as the resolution in the numerical simulation (Hourdin et al., 2013). Uncertainties that arise from the atmospheric model have been analyzed for ET and SSM by Cheruy et al. (2020). For China, forced ORCHIDEE simulations with WFDEI have been shown to perform better than with other forcing datasets (Princeton Global meteorological Forcing and Climatic Research Unit-National Center for Environmental Prediction

: Yin et al., 2018). More generally, the uncertainty arising from atmospheric forcing selection should be kept in mind, as it is comparable to the one of varying the LSM in forced simulations (Guo et al., 2006). Other perspectives include factor analysis against other hydrometeorological parameters, such as radiation, temperature, and precipitation frequency (Qian et al., 2006; Yin et al., 2018).

5 Conclusions

This paper has presented an in-depth evaluation of five interlinked essential climate variables (namely surface soil moisture, evapotranspiration, leaf area index, albedo, and precipitation) simulated by ORCHIDEE land surface model under different simulation modes (either forcing by WFDEI or coupled with LMDZ). Statistical evaluation was conducted using various reference-data sources (ESA CCI, upscaled FLUXNET, GIMSS 3g, MODIS products, and GPCC), and factor analysis was conducted against various landscape factors (namely plant functional type, leaf area index, irrigation, precipitation, slope, snow water equivalent, and evapotranspiration). Although ORCHIDEE consistently represented the spatiotemporal patterns of each essential climate variable in general, some issues were found relating to water cycles and their different consequences between the forced and coupled simulations. Errors relating to freezing/snowmelt, artificial water input such as irrigation, and precipitation bias propagated through surface–atmosphere coupling in the coupled mode. The factor analysis revealed a strong link between irrigation and precipitation (that further affected surface soil moisture, evapotranspiration, and leaf area index, particularly in the coupled mode) and a relatively complex link between precipitation and evapotranspiration that reflected the hydrometeorological regime of the region (energy-limited or water-limited) and the snow-albedo feedback in mountainous and boreal regions. In addition, the description of vegetation and snow seasonality seemed to be an issue in ORCHIDEE. Excessive and/or too fast green-up in temperate forest may lead an overestimation of leaf area index and near infrared albedo. Excessive and/or too fast snowmelt in spring in the boreal region may result in the underestimation of albedo in such regions, which can affect energy balance and water cycles. The different results between the forced and coupled modes stress the importance of model evaluation under both modes to determine each potential error source in model simulation.

21

Data availability

The ORCHIDEE simulation data used for this study (revision 4783) are referred by doi:10.14768/d2569664-3578-4c8f-8a45-25a927c8ed64, and freely accessible from https://data.ipsl.fr/catalog/srv/eng/catalog.search#/metadata/d2569664-3578-4c8f-8a45-25a927c8ed64. The reference data are freely available from the following links. All the links were confirmed to be

- The latest version of ESA CCI SM. https://www.esa-soilmoisture-cci.org/node/145. ESA also provides the previous version (version 4.4 was used in this study) upon request.
- The evapotranspiration product by Jung et al. (2019). https://www.bgc-jena.mpg.de/geodb/projects/FileDetails.php (FLUXCOM data portal, on CC4.0 BY license).
- The latest version of LAI3g. https://daac.ornl.gov/VEGETATION/guides/Mean_Seasonal_LAI.html. The original version used in this study (Zhu et al., 2013) can be obtained on request at http://cliveg.bu.edu/modismisr/lai3g-fpar3g.html.
- The MODIS albedo product (MCD43C3). https://lpdaac.usgs.gov/products/mcd43c3v006/.

22

- The latest version of GPCC (version 7). https://psl.noaa.gov/data/gridded/data.gpcc.html. The version used in this study (version 6, 0.5-deg) is available at https://opendata.dwd.de/climate_environment/GPCC/html/fulldata_v6_doi_download.html.
- Global Map of Irrigation Areas (GMIA; Sibert, 2013). http://www.fao.org/aquastat/en/geospatial-information/global-maps-irrigated-areas/latest-version/.
- ETOPO5. https://www.ngdc.noaa.gov/mgg/global/etopo5.HTML.

Code availability

The version of the ORCHIDEE model used for this study (revision 4783) is very close of tag 2.0, freely available from http://forge.ipsl.jussieu.fr/orchidee/browser/tags/ORCHIDEE_2_0/ORCHIDEE/ (Peylin et al., 2020). Tag 2.0 is based on revision 4783, with updates regarding the number and format of output variables to comply the CMIP6 requirements, and a few very minor bug corrections regarding the carbon cycle. The code of revision 4783 can be obtained upon request to the corresponding author.

Competing interests

The author declare that they have no conflict of interest.

References

Adam, J. C., and Lettenmaier, D. P.: Adjustment of global gridded precipitation for systematic bias. J. Geophys. Res., 108, 4257, doi:10.1029/2002JD002499, 2003.

Adikari, Y., and Noro, T.: A global outlook of sediment-related disasters in the context of water-related disasters. International Journal of Erosion Control Engineering. 3, 110–116, https://doi.org/10.13101/ijece.3.110, 2010.

Adler, R.F., Huffman, G.J., Chang, A., Ferraro, R., Xie, P., Janowiak, J., Rudolf, B., Schneider, U., Curtis, S., Bolvin, D., Gruber, A., Susskind, J., and Arkin, P.: The Version 2 Global Precipitation Climatology Project (GPCP) Monthly Precipitation Analysis (1979-Present). J. Hydrometeor., 4, 1147-1167, https://doi.org/10.1175/1525-7541(2003)004<1147:TVGPCP>2.0.CO;2, 2003.

23

Ait-Mesbah, S., Dufresne, J.L., Cheruy, F., and Hourdin, F.: The role of thermal inertia in the representation of mean and diurnal range of surface temperature in semiarid and arid regions. Geophysical Research Letters, 42, 7572-7580, https://doi.org/10.1002/2015GL065553, 2015.

Al-Yaari, A., Wigneron, J.-P., Kerr, Y., de Jeu, R., Rodriguez-Fernandez, N., van der Schalie, R., Al Bitar, A., Mialon, A., Richaume, P., Dolman, A., and Ducharne, A.: Testing regression equations to derive long-term global soil moisture datasets from passive microwave observations. Remote Sens. Environ., 180, 453–464, https://doi.org/10.1016/j.rse.2015.11.022, 2016.

Al-Yaari, A., Ducharne, A., Cheruy, F., Crow, W. T, and Wigneron, J. -P.: Satellite-based soil moisture provides missing link between summertime precipitation and surface temperature biases in CMIP5 simulations over conterminous United States. Sci. Rep., 9, 1657, https://doi.org/10.1038/s41598-018-38309-5, 2019a.

Al-Yaari, A., Wigneron, J.-P., Dorigo, W., Colliander, A., Pellarin, T., Hahn, S., Mialon, A., Richaume, P., Fernandez-Moran, R., Fan, L., Kerr, Y.H., and de Lannoy, G.: Assessment and inter-comparison of recently developed/reprocessed microwave satellite soil moisture products using ISMN ground-based measurements. *Remote Sens. Environ.* 224, 289-303, https://doi.org/10.1016/j.rse.2019.02.008, 2019b.

Becker, A., Finger, P., Meyer-Christoffer, A., Rudolf, B., Schamm, K., Schneider, U., and Ziese, M.: A Becker, A., Finger, P., Meyer-Christoffer, A., Rudolf, B., Schamm, K., Schneider, U., and Ziese, M.: Becker, A., Finger, P., Meyer-Christoffer, A., Rudolf, B., Schamm, K., Schneider, U., and Ziese, M.:

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., van den Hurk, B. J. J., Nearing, G. S., Pak, B., Peters-Lidard, C., Santanello, Jr. J. A., Stevens, L., and Vuichard, N.: The plumbing of land surface models: benchmark model performance. J. Hydrometeor. 16(3), 1425-1442, https://doi.org/10.1175/JHM-D-14-0158.1, 2015.

Boisier, J. P., de Noblet-Ducoudré, N., and Ciais, P.: Historical land-use-induced evapotranspiration changes estimated from present-day observations and reconstructed land-cover maps, Hydrol. Earth Syst. Sci., 18, 3571–3590, https://doi.org/10.5194/hess-18-3571-2014, 2014.

24

Bonan, G. B.: Forests and climate change: forcing, feedbacks, and the climate benefits of forests. Science, 320, 1444-1449, https://doi.org/10.1126/science.1155121, 2008.

Boucher, O. Servonnat, J. et 77 co-authors. Presentation and evaluation of the IPSL-CM6A-LR climate model, JAMES, 12, e2019MS002010, https://doi.org/10.1029/2019MS002010, 2020.

655   Cheruy, F., Campoy, A., Dupont, J.-C., Ducharne, A., Hourdin, F., Haeffelin, M., Chiriaco, M., and Idelkadi, A.: Combined influence of atmospheric physics and soil hydrology on the simulated meteorology at the SIRTA atmospheric observatory, Clim. Dyn., 40, 2251-2269, https://doi.org/10.1007/s00382-012-1469-y, 2013.

Cheruy, F., Ducharne A., Hourdin, F., Musat, I., Vignon, E., Gastineau, G., Bastrikov, V., Vuichard, N.,
660   Diallo, B., Dufresne, J.L., Ghattas, J., Grandpeix, J.Y., Idelkadi, A., Mellul, L., Maigna, F., Nenegoz, M., Ottlé, C., Peylin, P., Wang, F., and Zhao, Y.: Improved near surface continental climate in IPSL-CM6A-LR by combined evolutions of atmospheric and land surface physics. JAMES, 12, e2019MS002005,   https://doi.org/10.1029/2019MS002005, 2020.

665   Cheruy, F., Dufresne, J.L., Ait-Mesbah, S., Grandpeix, J.Y., and Wang, F.: Role of soil thermal inertia in surface temperature and soil moisture-temperature feedback, JAMES, 9(8), 2906-2919, https://doi.org/10.1002/2017MS001036, 2017.

Decharme, B., and Douville, H.: Uncertainties in the GSWP-2 precipitation forcing and their impacts on regional and global hydrological simulations. Clim. Dyn., 27, 695-713,
670   https://doi.org/10.1007/s00382-006-0160-6, 2006.

Dee, D., Uppala, S., Simmons, A., Berrisford, P., Poli, P., Kobayashi, S., et 29 co-authors.: The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. Q. J. R. Meteorol. Soc., 137(656), 553-597, https://doi.org/10.1002/QJ.828, 2011.

de Rosnay, P., Polcher, J., Laval, K., and Sabre, M.: Integrated parameterization of irrigation in the land
675   surface model ORCHIDEE. Validation over Indian Peninsula. Geophysical Research Letters. 30(19), 1–4, https://doi.org/10.1029/2003GL018024, 2003.

Dirmeyer, P.A.: Climate drift in a coupled land-atmosphere model. *J. Hydrometeor*. 2, 89-100, https://doi.org/10.1175/1525-7541(2001)002<0089:CDIACL>2.0.CO;2, 2001.

25

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, https://doi.org/10.5194/hess-12-1387-2008, 2008.

Dorigo, W. A., Xaver, A., Vreugdenhil, M., Gruber, A., Hegyiová, A., Sanchis-Dufau, A. D., Zamojski, D., Cordes, C., Wagner,W., and Drusch, M.: Global Automated Quality Control of In Situ Soil Moisture Data from the International Soil Moisture Network. Vadose Zone Journal, 12(3), https://doi.org/10.2136/vzj2012.0097, 2013.

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). Geoscientific Model Development. 1-51, https://doi.org/10.5194/gmd-2017-65, 2017.

Entekhabi, D., Njoku, E. G., O'Neill, P. E., Kellogg, K. H., Crow, W. T., Edelstein, W. N., Entin, J. K., Goodman, S. D., Jackson, T. J., Johnson, J., Kimball, J., Piepmeier, J. R., Koster, R. D., Martin, N., McDonald, K. C., Moghaddam, M., Moran, S., Reichle, R., Shi, J. C., Spencer, M. W., Thurman, S. W., Tsang, L., Zyl, J. V.: The soil moisture active passive (SMAP) mission, in Proceedings of the IEEE, 98(5), 704-716, https://doi.org/10.1109/JPROC.2010.2043918, 2010.

Eyling, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., and Taylor, K. E.: Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization, Geosci. Model Dev. 9, 1937-1958. https://doi.org/10.5194/gmd-9-1937-2016, 2016a.

Eyring, V., Gleckler, P. J., Heinze, C., Stouffer, R. J., Taylor, K. E., Balaji, V., Guilyardi, E., Joussaume, S., Kindermann, S., Lawrence, B. N., Meehl, G. A., Righi, M., and Williams, D. N.: Towards improved and more routine Earth system model evaluation in CMIP. Earth Syst. Dynam., 7, 813-830, https://doi.org/10.5194/esd-7-813-2016, 2016b.

26

Eyring, V., Righi, M., Lauer, A., Evaldsson, M., Wenzel, S., Jones, C., Anav, A., Andrews, O., Cionni, I., Davin, E. L., Deser, C., Ehbrecht, C., Friedlingstein, P., Gleckler, P., Gottschaldt, K. -D. Hagemann, S., Juckes, M., Kindermann, S., Krasting, J., Kunert, D., Levine, R., Loew, A., Mäkelä, J., Martin, G., Mason, E., Phillips, A. S., Read, S., Rio, C., Roehrig, R., Senftleben, D., Sterl, A., van Ulft, L. H., Walton, J., Wang, S., and Williams, K. D.: ESMValTool (v1.0) – a community diagnostic and performance metrics tool for routine evaluation of Earth system models in CMIP, GeoSci. Model Dev., 9, 1747-1802, https://doi.org/10.5194/gmd-9-1747-2016, 2016c.

Fernandez-Moran, R., Al-Yaari, A., Mialon, A., Mahmoodi, A., Al Bitar, A., De Lannoy, G., Rodriguez-Fernandez, N., Lopez-Baeza, E., and Wigneron, J. P.: SMOS-IC: An alternative SMOS soil moisture and vegetation optical depth product. Remote Sensing, 9(5), 1–21, https://doi.org/10.3390/rs9050457, 2017.

Flato, G., Marotzke, J., Abiodun, B., Braconnot, P., Chou, S. C., Collins, W., Cox, P., Driouech, F., Emori, S., Eyring, V., Forest, C., Glecker, P., Guilyardi, E., Jakob, C., Kattsov, V., Reason, C., and Rummukainen, M.: Evaluation of Climate Models. In: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, edited by: Stocker, T. F., Qin, D., Plattner, G.-K., Tignor, M., Allen, S. K., Boschung, J., Nauels, A., Xia, Y., Bex, V. Midgley, P. M., Cambridge University Press, Cambridge, UK and New York, NY, USA, 741-866, 2013.

GCOS: Implementation plan for the global observing system for climate in support of the UNFCCC, August 2010, 1-29, 2010.

Gleckler, P. J., Doutriaux, C., Durack, P. J., Taylor, K. E., Zhang, Y., Williams, D. N., Mason, E., and Servonnat, J.: A more powerful reality test for climate models, Eos Trans. AGU, 97, https://doi.org/10.1029/2016EO051663, 2016.

Gu, L., Meyers, T., Pallardy, S. G., Hanson, P. J., Yang, B., Heuer, M., Hosman, K. P., Riggs, J. S., Sluss, D., and Wullschleger, S. D.: Direct and indirect effects of atmospheric conditions and soil moisture on surface energy partitioning revealed by a prolonged drought at a temperate forest site. Journal of Geophysical Research. 111, D16102, https://doi.org/10.1029/2006JD007161, 2006.

27

Guimberteau, M., Ciais, P., Ducharne, A., Boisier, J. P., Dutra Aguiar, A. P., Biemans, H., De Deurwaerder, H., Galbraith, D., Kruijt, B., Langerwisch, F., Poveda, G., Rammig, A., Rodriguez, D. A., Tejada, G., Thonicke, K., Von Randow, C., Von Randow, R. C. S., Zhang, K., and Verbeeck, H.: Impacts of future deforestation and climate change on the hydrology of the Amazon Basin: a multi-model analysis with a new set of land-cover change scenarios, Hydrol. Earth Syst. Sci., 21, 1455–1475, https://doi.org/10.5194/hess-21-1455-2017, 2017.

Guimberteau, M., Zhu, D., Maignan, F., Huang, Y., Yue, C., Dantec-Nédélec, S., Ottlé, C., Jornet-Puig, A., Bastos, A., Laurent, P., Goll, D., Bowring, S., Chang, J., Guenet, B., Tifafi, M., Peng, S., Krinner, G., Ducharne, A., Wang, F., Wang, T., Wang, X., Wang, Y., Yin, Z., Lauerwald, R., Joetzjer, E., Qiu, C., Kim, H., and Ciais, P.: ORCHIDEE-MICT (v8.4.1), a land surface model for the high latitudes: model description and validation. Geosci. Model Dev., 11, 121-163, https://doi.org/10.5194/gmd-11-121-2018, 2018.

Guo, Z., Dirmeyer, P.A., Hu, Z.Z., Gao, X., and Zhao, M.: Evaluation of the Second Global Soil Wetness Project soil moisture simulations: 2. Sensitivity to external meteorological forcing. J. Geophys. Res. Atmospheres. 111(D22), https://doi.org/10.1029/2006JD007845, 2006.

Hourdin, F., Foujols, M.A., Codron, F., Guemas, V., Dufresne, J.L., Bony, S., Denvil, S., Guez, L., Lott, F., Ghattas, J., Braconnot, P., Marti, O., Meurdesoif, Y., and Bopp, L.: Impact of the LMDZ atmospheric grid configuration on the climate and sensitivity of the IPSL-CM5A coupled model. *Clim. Dyn*. 40, 2167-2192, https://doi.org/10.1007/s00382-012-1411-3, 2013.

Hourdin, F., Rio, C., Grandpeix, J.Y., Madeleine, J.B., Cheruy, F., Rochetin, N., Jam, A., Musat, I., Idelkadi, A., Fairhead, L. et 15 co-authors. LMDZ6A: the atmospheric component of the IPSL climate model with improved and better tuned physics, JAMES, 12, e2019MS001892, https://doi.org/10.1029/2019MS001892, 2020.

IPCC. Climate Change 2014: Mitigation of Climate Change. Contribution of Working Group III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2014.

Jackson, T.J, Le Vine, D.M., Hsu, A.Y., Oldak, A., Starks, P.J., Swift, C.T., Isham, J.D., and Haken, M.: Soil moisture mapping at regional scales using microwave radiometry: the Southern Great Plains

28

Hydrology Experiment. IEEE Transactions on Geoscience and Remote Sensing. 37, 2136-2150, https://doi.org/10.1109/36.789610, 1999.

Jung, M., Reichstein, M., and Bondeau, A.: Towards global empirical upscaling of FLUXNET eddy covariance observations: validation of a model tree ensemble approach using a biosphere model. Biogeoscience, 6, 2001-2013, https://doi.org/10.5194/bg-6-2001-2009, 2009.

Jung, M., Reichstein, M., Margolis, H. A., Cescatti, A., Richardson, A. D., Arain, M. A., Arneth, A., Bernhofer, C., Bonal, D., Chen, J., Gianelle, D., Gobron, N., Kiely, G., Kutsch, W., Lasslop, G., Law, B. E., Lindroth, A., Merbold, L., Montagnani, L., Moors, E. J., Papale, D., Sottocornola, M., Vaccari, F., and Williams, C.: Global patterns of land-atmosphere fluxes of carbon dioxide, latent heat, and sensible heat derived from eddy covariance, satellite, and meteorological observations. Journal of Geophysical Research: Biogeosciences, 116(3), 1–16, https://doi.org/10.1029/2010JG001566, 2011.

Jung, M., Koirala, S., Weber, U., Ichii, K., Gans, F., Camps-Valls, G., Papale, D., Schwalm, C., Tramontana, G., Reichstein, M. The FLUXCOM ensemble of global land-atmosphere energy fluxes, Scientific Data, 6, 74, https://doi.org/10.1038/s41597-019-0076-8, 2019.

Krinner, G., Viovy, N., de Noblet‐Ducoudré, N., 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), GB1015, https://doi.org/10.1029/2003GB002199, 2005.

Legates, D.R., and Willmott, C.J.: Mean seasonal and spatial variability in gauge-corrected, global precipitation. Int. J. Climatol. 10, 111–127, http://dx.doi.org/10.1002/joc.3370100202, 1990.

Liu, Y., Bastidas, L.A., Gupta, H.V., and Sorooshian, S.: Impacts of a parameterization deficiency on offline and coupled land surface model simulations. J. Hydrometeor. 4, 901-914, https://doi.org/10.1175/1525-7541(2003)004<0901:IOAPDO>2.0.CO;2, 2003.

Liu, Y. Y., Dorigo, W. A., Parinussa, R. M., De Jeu, R. A. M., Wagner, W., McCabe, M. F., Evans, J. P., and van Dijk, A. I. J. M.: Trend-preserving blending of passive and active microwave soil moisture retrievals. Remote Sens. Environ., 123, 280–297, https://doi.org/10.1016/j.rse.2012.03.014, 2012.

29

Loew, A., Stacke, T., Dorigo, W., De Jeu, R., and Hagemann, S.: Potential and limitations of multidecadal satellite soil moisture observations for selected climate model evaluation studies. Hydrology and Earth System Sciences, 17(9), 3523–3542, https://doi.org/10.5194/hess-17-3523-2013, 2013.

790   Ma., H., Zeng, J., Chen, N., Zhang, X., Cosh, M.H., and Wang, W.: Satellite surface soil moisture from SMAP, SMOS, AMSR2, and ESA CCI: A comprehensive assessment using global ground-based observations. Remote Sens. Environ. 231(15), 111215, https://doi.org/10.1016/j.rse.2019.111215, 2019.

Mahfouf, J.-F., Manzi A.O., Noilhan J., Giordani H., and Deque M.: The land surface scheme ISBA

795   within the Meteo-France climate model ARPEGE. Part I: Implementation and preliminary result. J. Clim, 8, 2039–2057, https://doi.org/10.1175/1520-0442(1995)008<2039:TLSSIW>2.0.CO;2, 1995.

McVicar, T. R., Roderick, M. L., Donohue, R. J., Li, L. T., Van Niel, T. G. Thomas, A., Grieser, J., Jhajharia, D., Himri, Y., Mahowald, N. M, Mescherskaya, A. V., Kruger, A. C., Rehman, S., and Dinpashoh, Y.: Global review and synthesis of trends in observed terrestrial near-surface wind speeds:

800   Implications for evaporation. Journal of Hydrology. 416-417, 182-205, https://doi.org/10.1016/j.jhydrol.2011.10.024, 2012.

Mignot, J., Hourdin, F., Deshayes, J., Boucher, O., Gastineau, G., Musat, I., Vancoppenolle, M., Servonat, J., Caubel, A., Cheruy, F., Denvil, S., Dufresne, J.-L., Ethe, C., Fairhead, L., Foujols, M.-A., Grandpeix, J.-Y., Levavasseur, G., Marti, O., Menary, M., Rio, C., Rousset, C.: The tuning strategy of IPSL-

805   CM6A-LR, JAMES, revision submitted.

Miralles, D. G., Holmes, T. R. H., De Jeu, R. A. M., Gash, J. H., Meesters, A. G. C. A., and Dolman, A. J.: Global land-surface evaporation estimated from satellite-based observations. Hydrology and Earth System Sciences, 15(2), 453–469, https://doi.org/10.5194/hess-15-453-2011, 2011.

Nasonova, O.N., Gusev, Y.M., and Kovalev, Y.E.: Impact of uncertainties in meteorological forcing data

810   and land surface parameters on global estimates of terrestrial water balance components. Hydrol. Process. 25, 1074-1090. https://doi.org/10.1002/hyp.7651, 2011.

NOAA: Data Announcement 88-MGG-02, digital relief of the surface of the Earth. National Geophysical Data Center, Boulder, Colorado, 1988.

30

Oliva, R., Daganzo-Eusebio, E., Kerr, Y. H., Mecklenburg, S., Nieto, S., Richaume, P., and Gruhier, C.

815   SMOS radio frequency interference scenario: Status and actions taken to improve the RFI environment in the 1400-1427-MHZ passive band. IEEE Transactions on Geoscience and Remote Sensing, 50(5 PART 1), 1427–1439. https://doi.org/10.1109/TGRS.2012.2182775, 2012.

Petty, G. W.: A first course in atmospheric radiation. 2nd ed, Sundog Publishing, Madison, Wisconsin, USA, 100, 2006.

820   Peylin, P., Ghattas, J., Cadule, P., Cheruy, F. , Ducharne, A., Guenet, B., Lathière, J., Luyssaert S., Maignan, F., Maugis, P., Ottlé, C., Polcher, J., Viovy, N., Vuichard, N., Bastrikov, V., Guimberteau, M., Lanso, A.-S., MacBean, N., Mcgrath, M., Tafasca, S., and Wang, F.: Description and evaluation of the global land surface model ORCHIDEE - Tag2.0, available at: http://forge.ipsl.jussieu.fr/orchidee/browser/tags/ORCHIDEE_2_0/ORCHIDEE/, 2020, last accessed:

825   27 January 2021

Polcher, J., Piles, M., Gelati, E., Barella-Ortiz, A., and Tello, M.: Comparing surface-soil moisture from the SMOS mission and the ORCHIDEE land-surface model over the Iberian Peninsula. Remote Sens. Environ., 174(1), 69-81, https://doi.org/10.1016/j.rse.2015.12.004, 2016.

Qian, T. Dai, A. Trenberth, K.E., and Oleson, K.W.: Simulation of global land surface conditions from

830   1948 to 2004. Part I : Forcing data and evaluations. J. Hydrometeor. 7, 953-975, https://doi.org/10.1175/JHM540.1, 2006.

Qu, Y., Liu, Q., Liang, S., Wang, L., Liu, N., and Liu, S. Direct-estimation algorithm for mapping daily land-surface broadband albedo from MODIS data. IEEE Transactions on Geoscience and Remote Sensing. 52(2), 907- 919, https://doi.org/10.1109/TGRS.2013.2245670, 2014.

835   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 Sens., 10, 1786, https://doi.org/10.3390/rs10111786, 2019.

Reichle, R.H., Koster, R.D., Dong, J., and Berg, A.A.: Global soil moisture from satellite observations, land surface models, and ground data: implications for data assimilation. J. Hydrol. 5, 430-442,

840   https://doi.org/10.1175/1525-7541(2004)005<0430:GSMFSO>2.0.CO;2, 2004.

Schaaf, C. B., Gao, F., Strahler, A. H., Lucht, W., Li, X., Tsang, T., Strugnell, N. C., Zhang, X., Jin, Y., Muller, J.-P., Lewis, P., Barnsley, M., Hobson, P., Disney, M., Roberts, G., Dunderdale, M., Doll, C., d'Entremont, R. P., Hu, B., Liang, S., Privette, J. L., and Roy, D.: First operational BRDF, albedo nadir reflectance products from MODIS. Remote Sens. Environ. 83(1-2), 135-148,

845   https://doi.org/10.1016/S0034-4257(02)00091-3, 2002.

Schneider, U., Becker, A., Finger, P., Meyer-Christoffer, A., Ziese, M., and Rudolf, B.: GPCC's new land surface precipitation climatology based on quality-controlled in situ data and its role in quantifying the global water cycle. Theoretical and Applied Climatology, 115(1–2), 15–40, https://doi.org/10.1007/s00704-013-0860-x, 2014.

850   Seneviratne, S. I., Corti, T., Davin, E. L., Hirschi, M., Jaeger, E. B., Lehner, I., Orlowsky, B., and Teuling, A. J: Investigating soil moisture-climate interactions in a changing climate: A review. Earth-Sci. Rev., 99, 125–161, https://doi.org/10.1016/j.earscirev.2010.02.004, 2010.

Siebert, S., Burke, J., Faures, J. M., Frenken, K., Hoogeveen, J., Döll, P., and Portmann, F. T.: Groundwater use for irrigation - A global inventory. Hydrology and Earth System Sciences, 14(10),

855   1863–1880, https://doi.org/10.5194/hess-14-1863-2010, 2010.

Stroeve, J., Box, J. E., Gao, F., Liang, S., Nolin, A., and Schaaf, C.: Accuracy assessment of the MODIS 16-day albedo product for snow: Comparisons with Greenland in situ measurements. Remote Sensing of Environment, 94(1), 46–60, https://doi.org/10.1016/j.rse.2004.09.001, 2005.

Stroeve, J., Box, J. E., Wang, Z., Schaaf, C., and Barrett, A.: Re-evaluation of MODIS MCD43 greenland

860   albedo accuracy and trends. Remote Sens. Environ., 138, 199–214. https://doi.org/10.1016/j.rse.2013.07.023, 2013.

Tafasca, S., Ducharne, A., and Valentin, C.: Weak sensitivity of the terrestrial water budget to global soil texture maps in the ORCHIDEE land surface model, Hydrol. Earth Syst. Sci., 24, 3753–3774, https://doi.org/10.5194/hess-24-3753-2020, 2020.

865   Taylor, K. E., Stouffer, R. J., and Meehl, G. A.: An overview of CMIP5 and the experiment design, *B. Am. Meteor. Soc.*, 93, 485-498. https://doi.org/10.1175/BAMS-D-11-00094.1, 2012.

van den Hurk, B., Kim, H., Krinner, G., Seneviratne, S. I., Derksen, C., Oki, T., Douville, H., Colin, J., Ducharne, A., Cheruy, F., Viovy, N., Puma, M. J., Wada, Y., Li, W., Jia, B., Alessandri, A., Lawrence,

D. M., Weedon, G. P., Ellis, R., Hagemann, S., Mao, J., Flanner, M. G., Zampieri, M., Materia, S., Law, R. M., and Sheffield, J.: LS3MIP (v1.0) contribution to CMIP6: the Land Surface, Snow and Soil moisture Model Intercomparison Project - aims, setup and expected outcome. Geoscientific Model Development. 9(8), 2809-2832, https://doi.org/10.5194/gmd-9-2809-2016, 2016.

Vereecken, H., Weihermüller, L., Assouline, S., Šimůnek, J., Verhoef, A., Herbst, M., Nicole, A.,

Vereecken, H., Weihermüller, L., Assouline, S., Šimůnek, J., Verhoef, A., Herbst, M., Nicole, A.,

Vereecken, H., Weihermüller, L., Assouline, S., Šimůnek, J., Verhoef, A., Herbst, M., Nicole, A.,

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, https://doi.org/10.1002/jgrd.50395, 2013.

Wang, D., Liang, S., He, T., Yu, Y., Schaaf, C., Wang, Z.: Estimating daily mean land surface albedo from MODIS data. Journal of Geophysical Research: Atmospheres, 120, 4825–4841. https://doi.org/10.1002/2015JD023178, 2015.

Wang, T., Peng, S., Krinner, G., Ryder, J., Li, Y., Dantec-Nedelec, S. and Ottle, C.: Impacts of satellite-based snow albedo assimilation on offline and coupled land surface model simulations. PLOS ONE. 10(9), e0137275. https://doi.org/10.1371/journal.pone.0137275, 2015.

Wang, F. Cheruy, F., Dufresne, J.L.: The improvement of the soil thermodynamics and its effects on the land surface meteorology in the IPSL-CM. Geoscientific model development discussions. *Geosci. Model Dev.,* 9(1), 363-381, https://doi.org/10.5194/gmd-9-363-2016, 2016.

Weedon, G. P., Gomes, S., Viterbo, P., Shuttleworth, W. J., Blyth, E., Österle, H., Adam, J. C., Bellouin, N., Boucher, O., and Best, M.: Creation of the WATCH forcing data and its use to assess global and regional reference crop evaporation over land during the twentieth century. Journal of Hydrometeorology. 12, 823-848, https://doi.org/10.1175/2011JHM1369.1, 2011.

Weedon, G. P., Balsamo, G., Bellouin, N., Gomes, S., Best, M. J., and Viterbo, P.: Data methodology applied to ERA-Interim reanalysis data. Water Resources Research, 50, 7505–7514. https://doi.org/10.1002/2014WR015638, 2014.

33

Wigneron, J.-P., Jackson, T.J., O'Neill, P., De Lannoy, G., de Rosnay, P., Walker, J.P., Ferrazzoli, P., Mironov, V., Bircher, S., Grant, J.P., Kurum, M., Schwank, M., Munoz-Sabater, J., Das, N., Royer, A., Al-Yaari, A., Al Bitar, A., Fernandez-Moran, R., Lawrence, H., Mialon, A., Parrens, M., Richaume, P., Delwart, S., and Kerr, Y.: Modelling the passive microwave signature from land surfaces: A review of recent results and application to the L-band SMOS & SMAP soil moisture retrieval algorithms. Remote Sens. Environ. 192, 238-262, https://doi.org/10.1016/j.rse.2017.01.024, 2017.

Wigneron, J.-P., Kerr, Y.H., Waldteufel, P., Saleh, K., Escorihuela, M-J., Richaume, P., Ferrazzoli, P., de Rosnay, P., Gurney, R., Calvet, J.C., Grant, J.P., Guglielmetti, M., Hornbuckle, B., Matzler, C., Pellarin, T., and Schwank, M.: L-band Microwave Emission of the Biosphere (L-MEB) Model: description and calibration against experimental data sets over crop fields. Remote Sens. Environ. 107(4), 639-655. https://doi.org/10.1016/j.rse.2006.10.014, 2007.

Xi, Y., Peng, S., Ciais, P., Guimberteau, M., Li, Y., Piao, S., Wang, X., Polcher, J., Yu, J., Zhang, X., Zhou, F., Bo, Y., Ottle, C., and Yin, Z.: Contributions of Climate Change, CO2, Land-Use Change, and Human Activities to Changes in River Flow across 10 Chinese Basins, J. Hydrometeorol., 19, 1899-1914, https://doi.org/10.1175/JHM-D-18-0005.1, 2018.

Yang, L., Sun, G., Zhi, L., and Zhao, J. Negative soil moisture-precipitation feedback in dry and wet regions. Scientific Reports, 8(1), 1–9. https://doi.org/10.1038/s41598-018-22394-7, 2018.

Yin, Z., Ottle, C., Ciais, P., Guimberteau, M., Wang, X., Zhu, D., Maignan, F., Peng, S., Piao, S., Polcher, J., Zhou, F., and Kim, H.: Evaluation of ORCHIDEE-MICT-simulated soil moisture over China and impacts of different atmospheric forcing data. Hydrol. Earth Syst. Sci. 22(10), 5463-5484. https://doi.org/10.5194/hess-22-5463-2018, 2018.

Zabel, F., Mauser, W., Marke, T., Pfeiffer, A., Zangl, G., and Wastl, C.: Inter-comparison of two land-surface models applied at different scales and their feedbacks while coupled with a regional climate model. Hydrol. Earth Syst. Sci. 16, 1017-1031, https://doi.org/10.5194/hess-16-1017-2012, 2012.

Zeng, Z., Wang, T., Zhou, F., Ciais, P., Mao, J., Shi, X., and Piao, S.: A worldwide analysis of spatiotemporal changes in water balance-based evapotranspiration from 1982 to 2009. J. Geophys. Res. Atmos, 119, 1186–1202, https://doi.org/10.1002/2013JD020941, 2014.

Zhang, K., Kimball, J. S., Nemani, R. R., and Running, S. W.: A continuous satellite-derived global record of land surface evapotranspiration from 1983 to 2006. Water Resources Research, 46(9), 1–21. https://doi.org/10.1029/2009WR008800, 2010.

925

Zhao, J., Wang, Y., Zhang, H., Zhang, Z., Guo, X., Yu, S., and Du, W.: Spatially and temporally continuous LAI datasets based on the mixed pixel decomposition method. Springerplus. 5, 516, https://doi.org/10.1186/s40064-016-2166-9, 2016.

Zhu, Z., Bi, J., Pan, Y., Ganguly, S., Anav, A., Xu, L., Samanta, A., Piao, S., Nemani, R. R., and Myneni,

930

R. B.: Global data sets of vegetation leaf area index (LAI)3g and fraction of photosynthetically active radiation (FPAR)3g derived from global inventory modeling and mapping studies (GIMMS) normalized difference vegetation index (NDVI3G) for the period 1981 to 2011. Remote Sensing, 5(2), 927–948, https://doi.org/10.3390/rs5020927, 2013.

Zobler, L.: A world soil file for global climate modeling, National Aeronautics and Space Administration,

935

Technical Memorandum 87802, 1986.

Table 1. Overview of the selected reference datasets and period of analysis. The last column gives the percent of land pixels in the maps of Figs. 2 and 3 with observed values. Smaller amount of SSM data is available than the others as a result of relatively strict quality control. For ET and LAI, no data were available in extremely arid regions.

Variable	Reference product	Evaluation period	Observed fraction of land area (%)
SSM	ESA CCI v4.4 (Liu et al., 2012)	1993–1999	44.1
ET	Jung et al. (2019)	1987–2009	89.3
LAI	LAI3g (Zhu et al., 2013)	1987–2009	87.8
Precipitation	GPCC (Schneider et al., 2014)	1987–2009	98.5
Albedo	MODIS (Qu et al., 2014)	2003–2009	87.4

Table 2. Correspondence between classification levels and values for each factor.

Factor	Reference data	How classified	Fraction of land area (%)
PFT	ORCHIDEE-defined plant functional types	class 1: bare soil is dominant	15.4
		class 2: tropical broadleaf evergreen forest is dominant	6.9
		class 3: tropical broadleaf raingreen forest is dominant	3.2
		class 4: temperate needleleaf evergreen forest is dominant	2.0
		class 5: temperate broadleaf evergreen forest is dominant	3.4
		class 6: temperate broadleaf summergreen forest is dominant	3.6
		class 7: boreal needleleaf evergreen forest is dominant	8.6
		class 8: boreal broadleaf summergreen forest is dominant	6.5
		class 9: boreal needleleaf summergreen forest is dominant	8.6
		class 10: temperate C3 grasses are dominant	6.6
		class 11: C4 grasses are dominant	8.8
		class 12: C3 crops are dominant	
		class 13: C4 crops are dominant	

36

		class 14: tropical C3 grasses are dominant	9.2
		class 15: boreal C3 grasses are dominant	2.1
			2.7
			12.4
LAI	Zhu et al. (2013)	class 1 (low LAI): —0 – 1.0 —m^2/m^2	41.9
		class 2 (middle LAI): 1.0 – 3.0 -m^2/m^2	37.1
		class 3 (high LAI): —3.0 – ——m^2/m^2	8.7
ET	Jung et al. (2019)	class 1: less than 1 mm/d	45.8
		class 2: 1–2 mm/d	23.6
		class 3: 2–3 mm/d	13.0
		class 4: more than 3 mm/d	6.9
Precipitation	GPCC, Schneider et al. (2014)	class 1 (extremely dry): less than 1 mm/d	41.0
		class 2 (dry): 1 to 2 mm/d	24.1
		class 3 (moderate): 2 to 4 mm/d	17.2
		class 4 (wet): 4 to 7 mm/d	11.0
		class 5 (extremely wet): more than 7 mm/d	5.3
SWE	ORCHIDEE-simulated SWE	class 1: 0 mm	33.3
		class 2: 0–16 mm	33.6
		class 3: 16–32 mm	8.7
		class 4: 32–48 mm	9.5
		class 5: more than 48 mm	14.8
Irrigated area	Siebert et al. (2010)	class 1: 0%	56.6
		class 2: 0–5%	34.7
		class 3: 5–10%	3.6
		class 4: 10–20%	2.6
		class 5: 20–50%	1.9
		class 6: 50–100%	0.5

Slope	ETOPO5+ (Amante & Eakins, 2009NOAA, 1988)	class 1 (flat): 0–0.5 degree	3.0
		class 2 (middle): 0.5–2.0 degree	28.1
		class 3 (steep): 2.0– degree	67.4

Table 3. Land averages of the evaluation criteria (bias, RMSE, and correlation coefficient CC) for the selected variables and reference data sets. The same bias sign on average over land for all variables wasere observed between forced and coupled simulations. Positive systematic bias were observed for LAI, and large uncertainty (i.e., RMSE) was observed for SSM and ET. ET shows the best correlation coefficient. Overall, coupled simulation tends to behave more realistically, despite overestimation of precipitation.

		Forced	Coupled
Bias	Precipitation (mm/d)	0.112	0.186
	SSM (normalized)	-0.072	-0.062
	ET (mm/d)	-0.231	-0.133
	LAI (-)	0.325	0.220
	Albedo (-)	-0.000	0.009
RMSE	Precipitation (mm/d)	1.057	1.680
	SSM (normalized)	0.546	0.560
	ET (mm/d)	0.513	0.540
	LAI (-)	0.586	0.554
	Albedo (-)	0.048	0.047
CC	Precipitation (mm/d)	0.790	0.605
	SSM (normalized)	0.581	0.551
	ET (mm/d)	0.744	0.692
	LAI (-)	0.328	0.340
	Albedo (-)	0.395	0.426

950

955

Table 4. Spatial correlation coefficients (SCC) between biases (in forced and coupled modes) and potential explanatory factors. PFT was excluded in the Table because it is nominal scale. Statistically insignificant SCCs appear in italic. SSM tended to be underestimated in high P, SSM, ET, LAI regions for both forced and coupled modes. Between forced and coupled ET, opposite association with P, ET, and LAI was observed. LAI in both modes were positively biased in high P, ET, SSM region (probably corresponding to the tropical region). Albedo and coupled P were strongly associated with slope. Irrigation is likely to bias SSM and ET negatively, and the effect was more enhanced in the coupled mode.

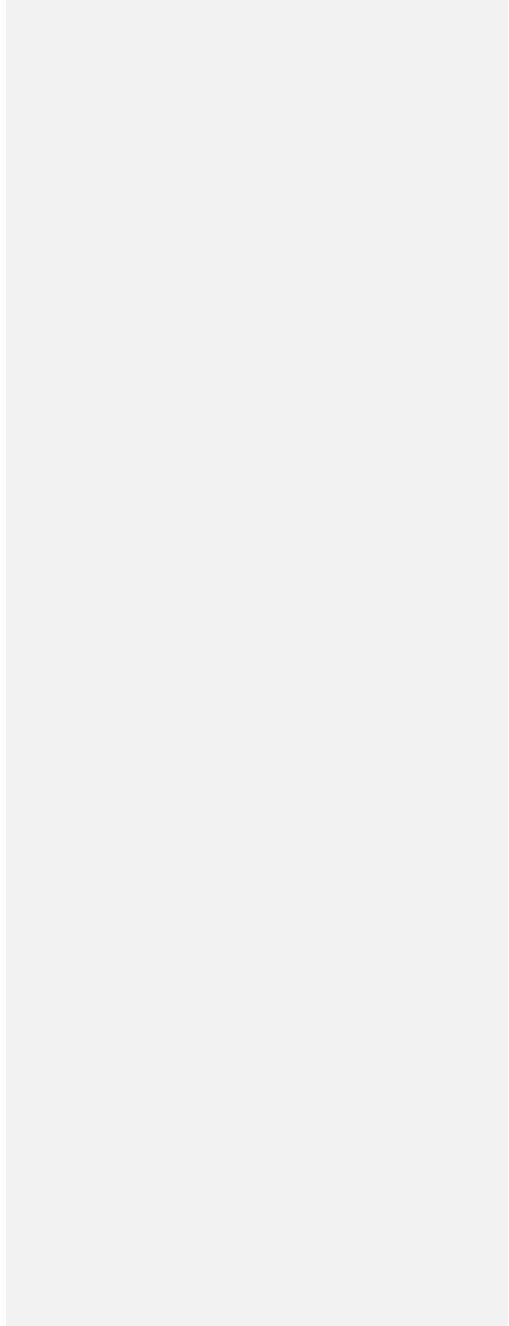|	Biases of forced simulations				
Factors	P	SSM-CCI	ET	LAI	Albedo
P	0.054	-0.203	-0.168	0.375	0.283
ET	0.064	-0.163	-0.277	0.357	0.344
LAI	0.068	-0.083	-0.127	0.263	0.275
SWE	0.096	0.024	-0.103	-0.090	0.181
Irrigated fraction	-0.060	-0.066	-0.170	0.012	0.059
Slope	-0.027	-0.068	-0.010	0.027	0.023
	Biases of coupled simulations				
Factors	P	SSM-CCI	ET	LAI	Albedo
P	-0.108	-0.234	0.200	0.258	0.130
ET	*0.006*	-0.164	0.163	0.245	0.139
LAI	*-0.004*	-0.121	0.264	0.092	0.103
SWE	0.034	*-0.009*	-0.134	-0.060	0.073
Irrigated fraction	-0.071	-0.118	-0.213	-0.012	0.030
Slope	0.267	0.085	-0.022	-0.031	0.249
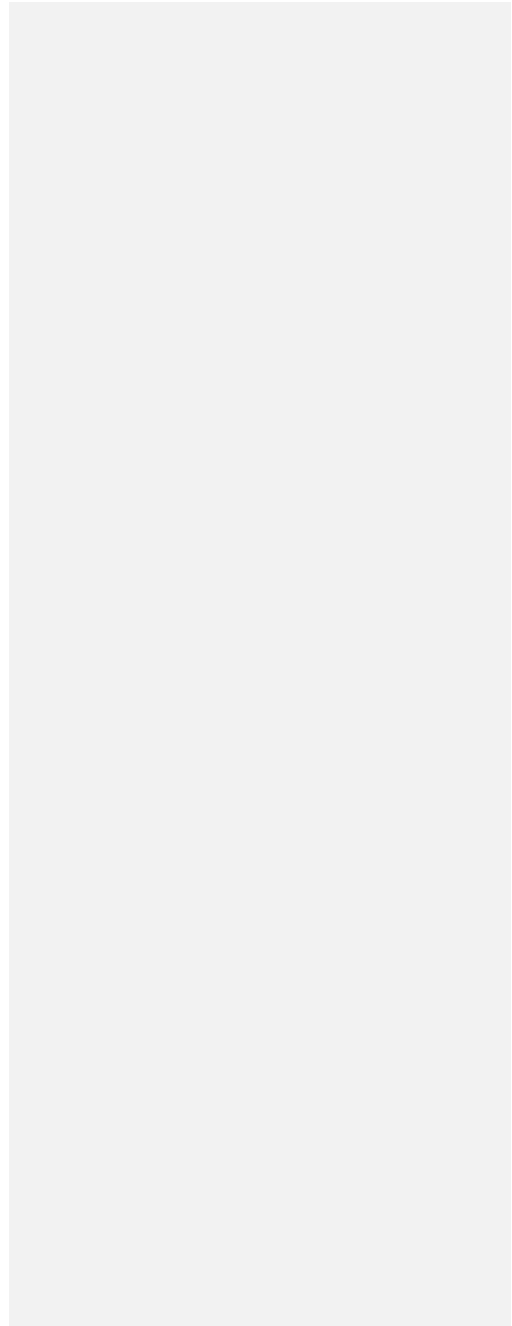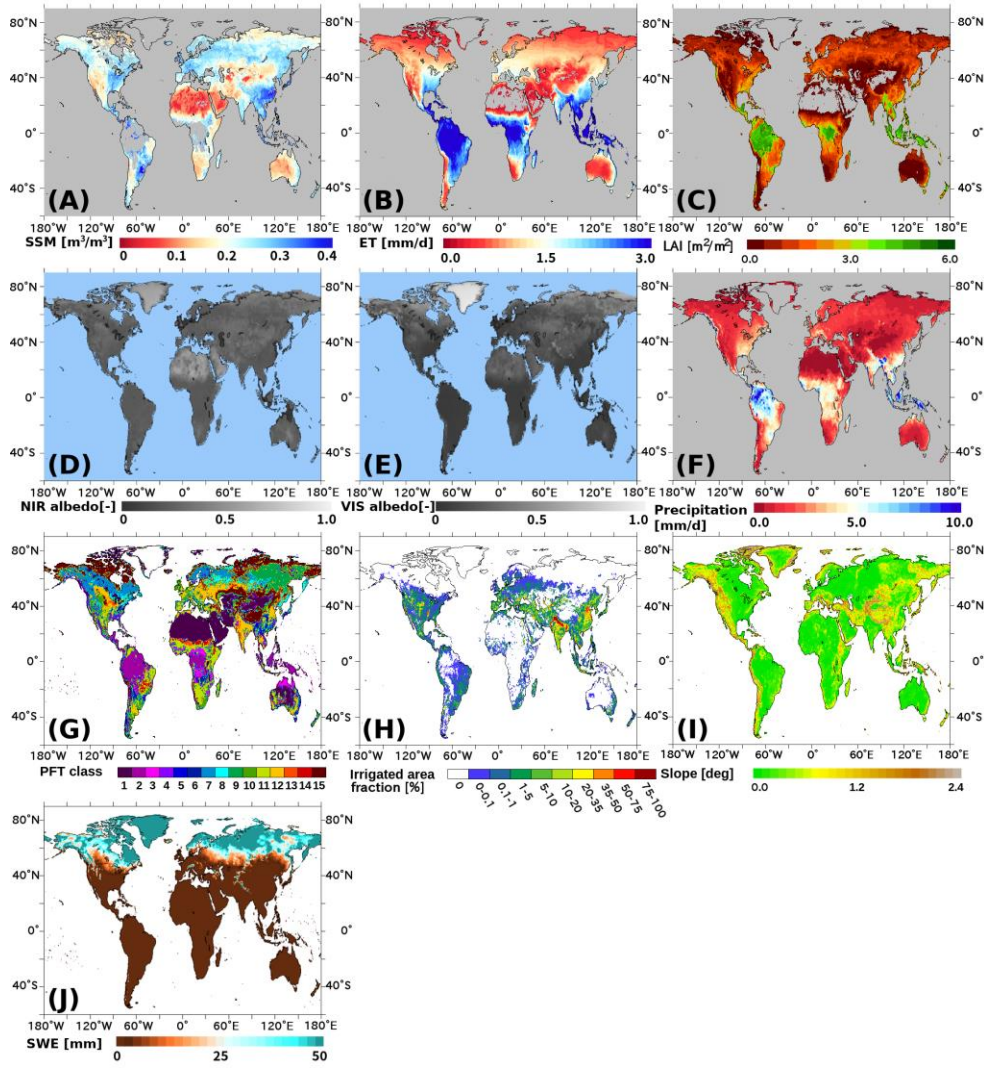
Figure 1: Spatial patterns of <u>pluri-annual mean reference data used for validation (marked with a +) and/or factor analysis (marked</u> <u>with a x): (A+) SSM from ESA CCI; (B+x) ET product of Jung et</u>

al. (2019); (C+x) LAI3g; (D+) MODIS NIR albedo; (E+) MODIS visible albedo; (F+x) GPCC precipitation data; (Gx) dominant plant functional type used in ORCHIDEE (see Table 2 for the class definition); (Hx) fractional area equipped with irrigation; (Ix) slope derived from ETOPO5; (Jx) snow water equivalent derived from the forced-mode ORCHIDEE.

980

42
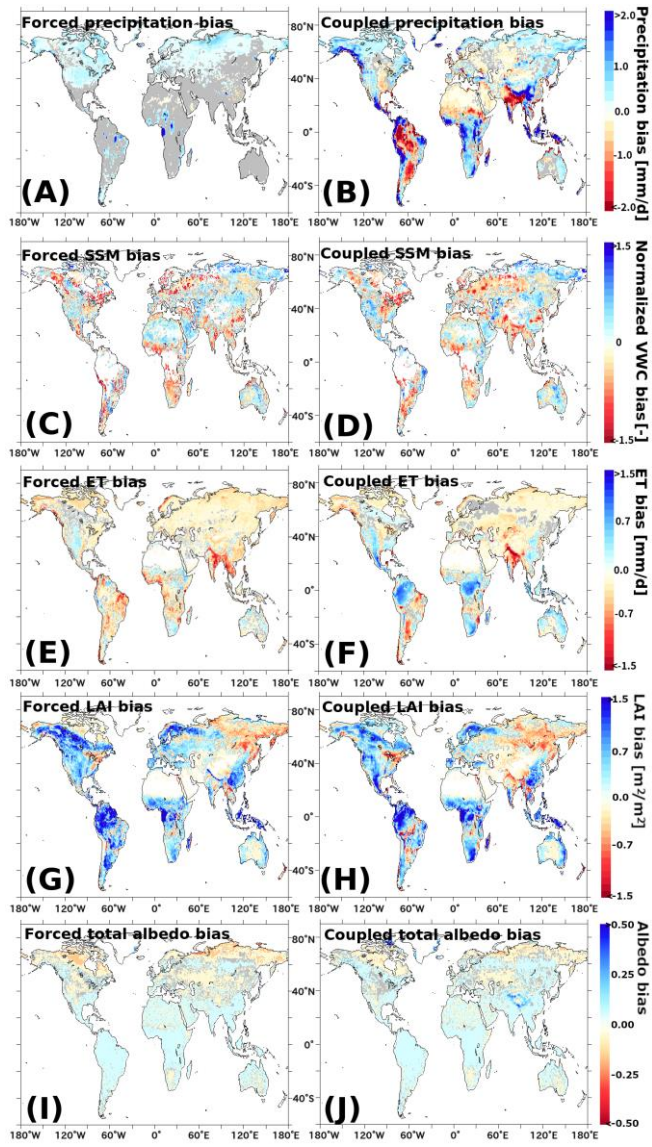
985

Figure 2: Pluri-annual average of bias (i.e., simulated values minus observed values) for the five evaluated variables, simulated in forced mode (left) and coupled mode (right): (A, B): Precipitation bias against GPCC; (C, D): SSM bias against CCI-SSM, in normalized VWC during period 1 (1993–1999); (E, F): ET bias against upscaled FLUXNET data; (H, H): LAI bias against LAI3g data; (I, J): total-albedo bias against MODIS albedo product. Grey areas are statistically insignificant pixels.

990

995

44

(A) Forced SSM correlation
(B) Coupled SSM correlation
(C) Forced NIR albedo correlation
(D) Coupled NIR albedo correlation
(E) Forced visible albedo correlation
(F) Coupled visible albedo correlation

Correlation coefficient
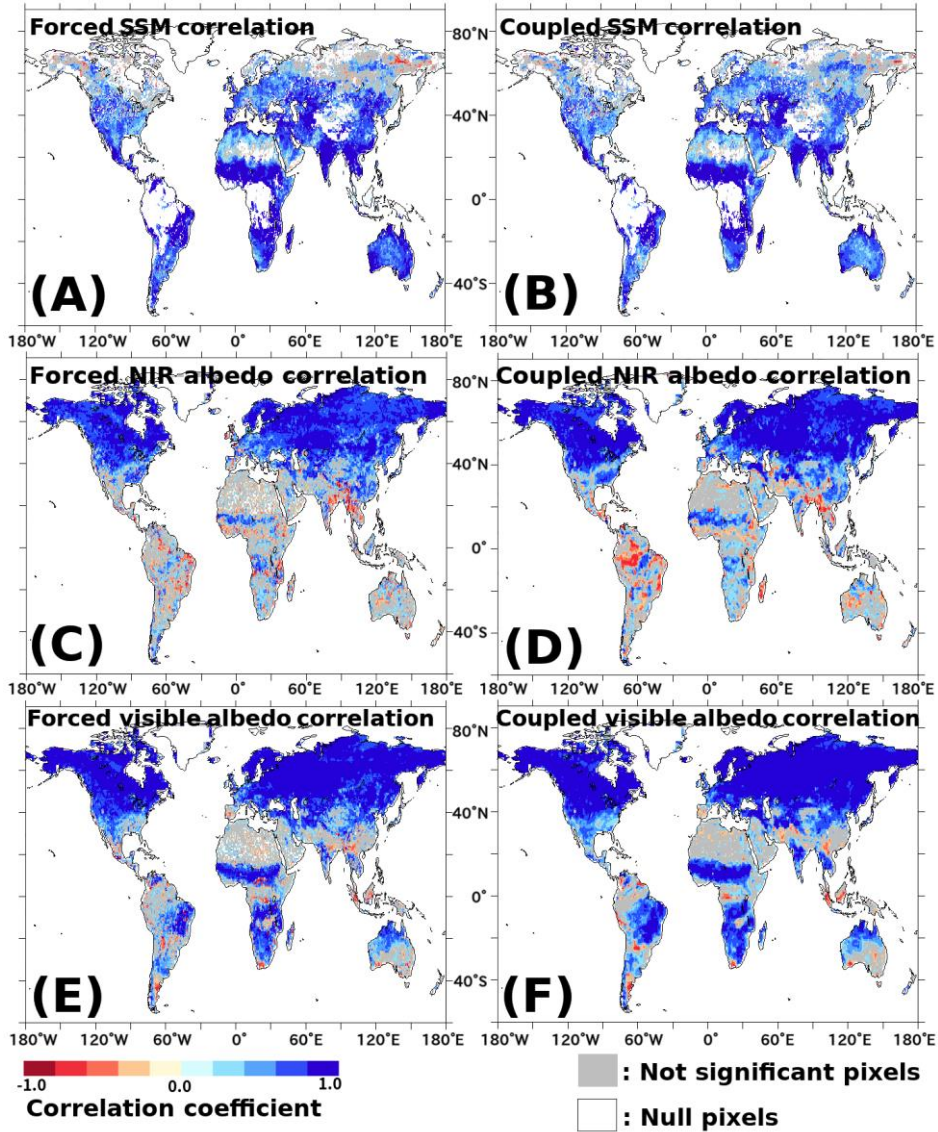-1.0   0.0   1.0

: Not significant pixels
: Null pixels

Figure 3: Spatial patterns of correlation coefficient along time series (with monthly time steps) per pixel, for SSM and albedo. (A), (B): correlation coefficient between simulated SSM and CCI-SSM in normalized VWC for forced and coupled mode, respectively. (C), (D): correlation coefficient between simulated and observed (MODIS) albedo in NIR band for forced and coupled mode, respectively. (E), (F): those in visible band. White areas are null pixels that were excluded by the quality control, and grey areas are statistically not significant pixels.

1000
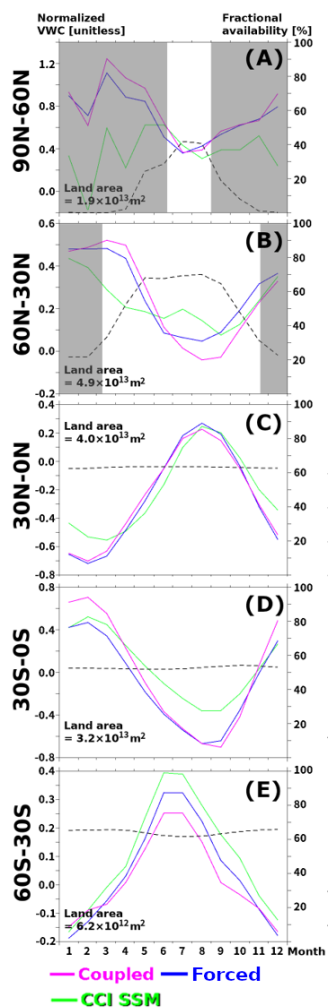
Figure 4: Comparison of SSM seasonal variations among reference (CCI-SSM) and simulations (forced and coupled) for each latitude zone. Dashed black line is the fraction of available pixels to all land pixels over each zone. Depending on the snow mask, the number of available pixels varied along the season in high latitude regions. To avoid misleading interpretation by the small number of samples with unreliable SSM reference, periods of less pixel availability (<30%) are greyed out.
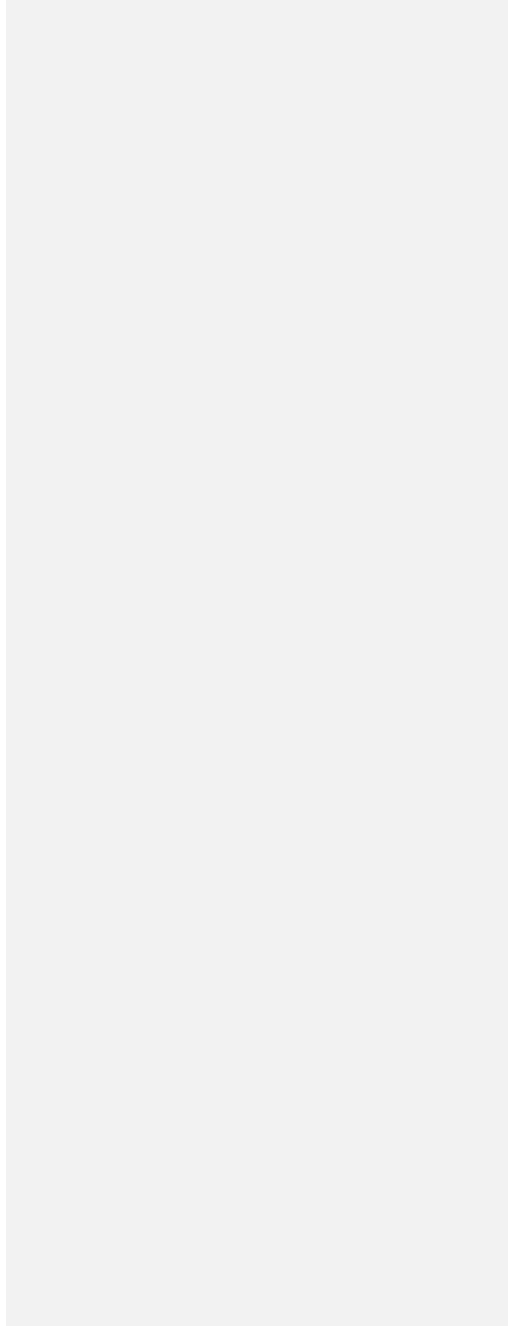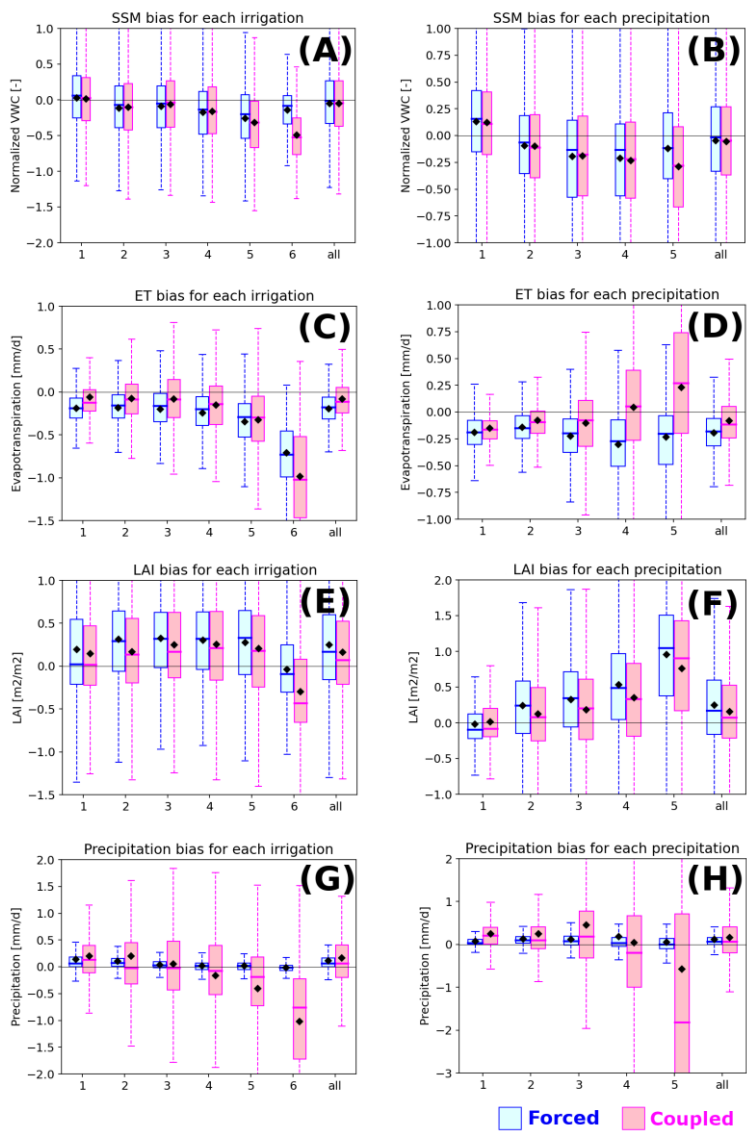
1005
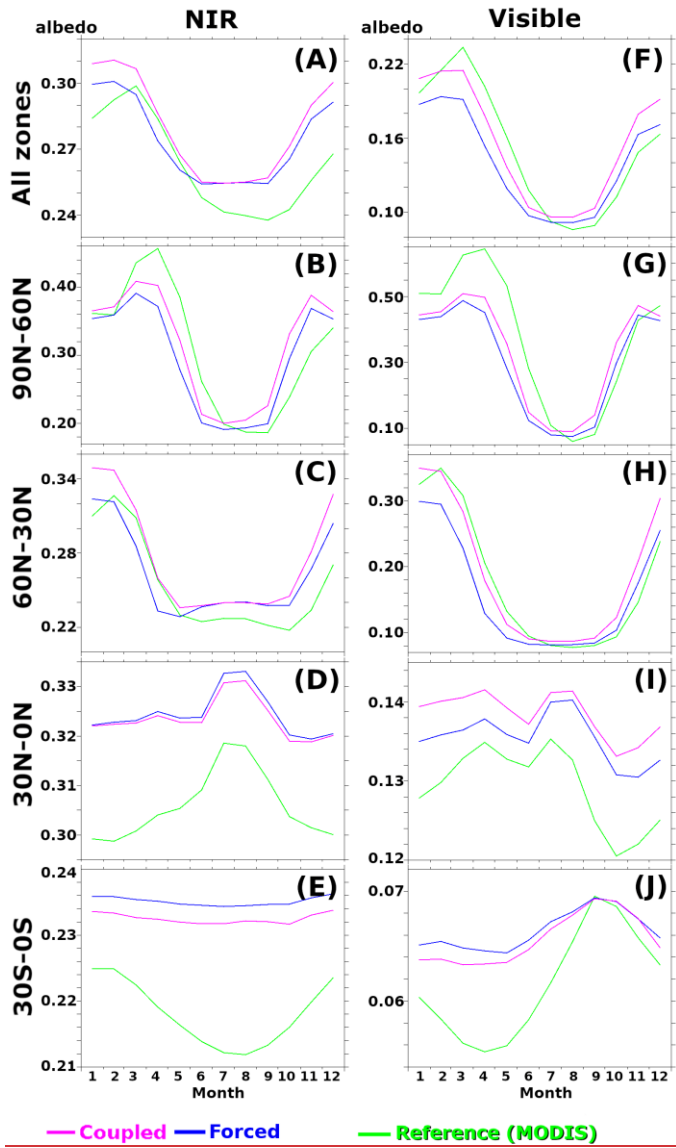
1010 **Figure 5:** Monthly climatological time series of global and zonal mean albedo. The left column shows the NIR band (A: global average; B–E: zonal average for each 30° in latitude), while the right column shows the visible band (the vertical arrangement is the same as that for NIR).

SSM bias for each irrigation (A)
SSM bias for each precipitation (B)
ET bias for each irrigation (C)
ET bias for each precipitation (D)
LAI bias for each irrigation (E)
LAI bias for each precipitation (F)
Precipitation bias for each irrigation (G)
Precipitation bias for each precipitation (H)

Forced   Coupled

1015   **Figure 6: Boxplots of mean biases (simulated minus observed values) of SSM, ET, LAI (forced/coupled), and precipitation (coupled only) against each class of irrigation and precipitation. The upper limit, middle line, and lower limit of the boxes correspond 25-, 50- and 75- percentile values, respectively. The upper and lower limits of whiskers are maximum and minimum values, respectively. The diamond indicates mean value of the class. (A), (B): SSM bias; (C), (D): ET bias; (E), (F): LAI bias; (G), (H): precipitation bias vs. irrigation and precipitation classes, respectively. Blue and pink boxes correspond to forced and coupled mode,**
1020   **respectively. The horizontal black line shows zero. Each class of landscape factors (i.e., x-axis) is defined in Table 1.**

51
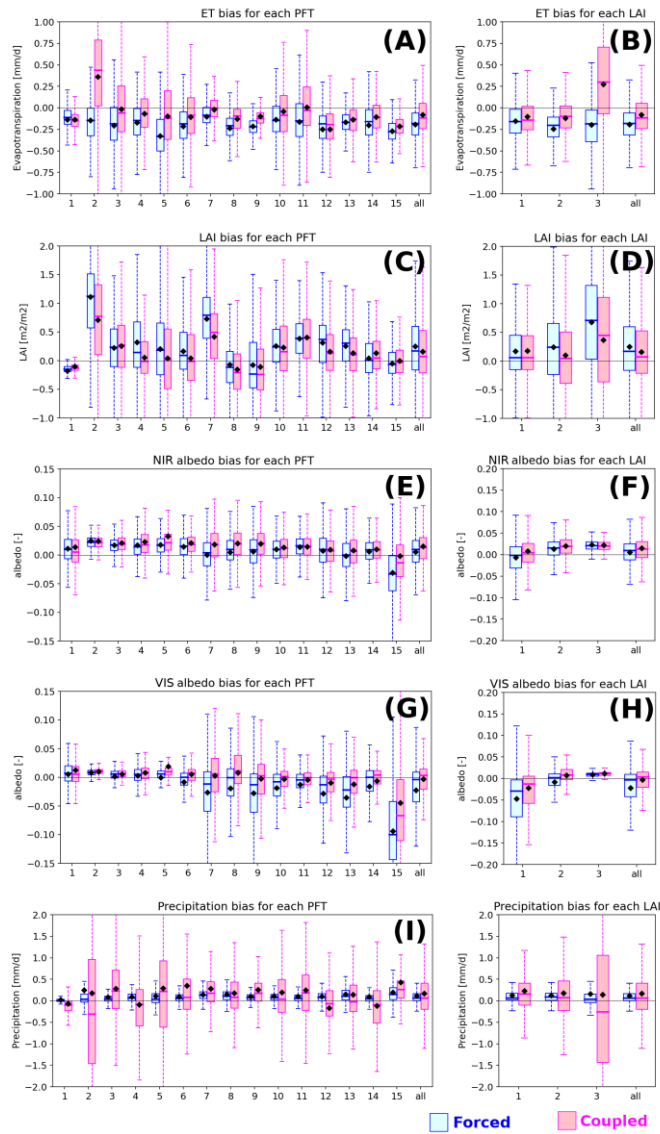
Figure 7: Boxplots of mean biases (simulated minus observed values) of ET, LAI, NIR/visible albedo (forced/coupled), and precipitation (coupled only) against each class of PFT and LAI. (A), (B): ET bias; (C), (D): LAI bias; (E), (F): NIR_albedo bias; (G), (H): visible_albedo bias; (I), (J): precipitation bias vs. PFT and LAI classes, respectively. Legends and axes are the same as in Figure 6, and each class of landscape factors (i.e., x-axis) is defined in Table 1.
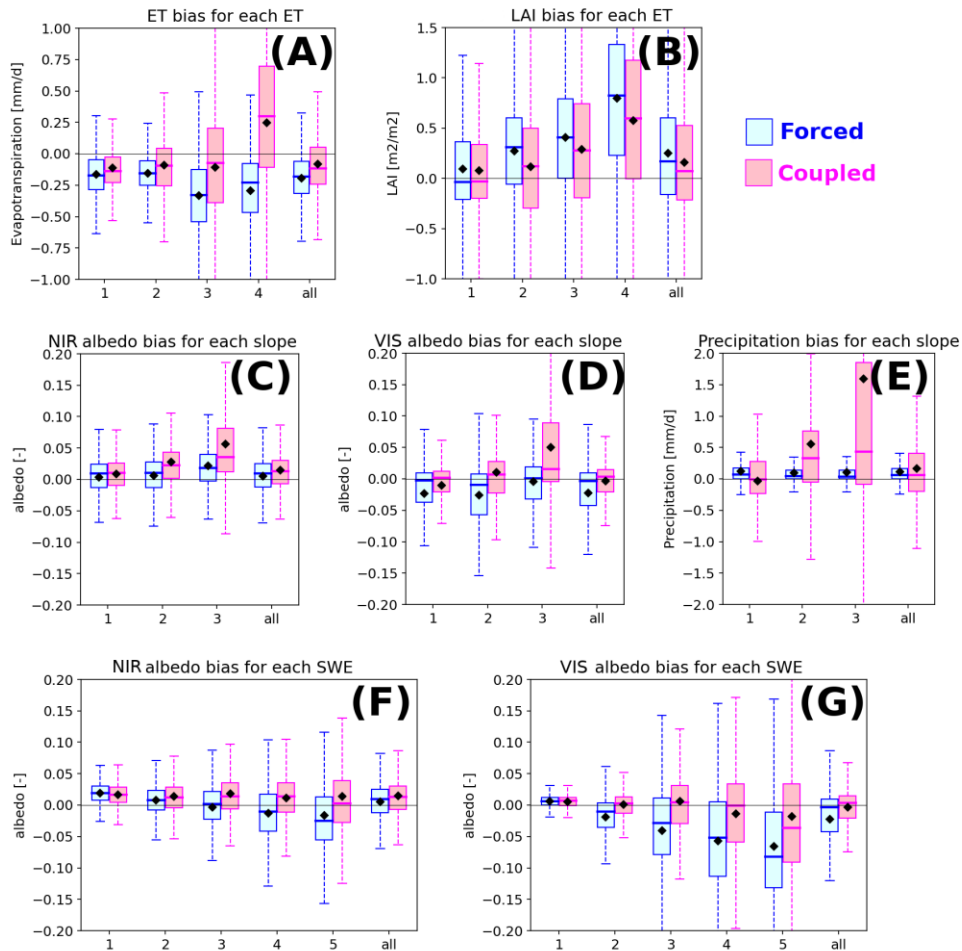
Figure 8: Boxplots of mean biases (simulated minus observed values) against each class of ET, slope, and SWE. (A), (B): ET and LAI bias vs. ET class; (C), (D), (E): NIR albedo, visible albedo, and precipitation bias vs. slope class; (F), (G): NIR and visible albedo bias against SWE class, respectively. Legends and axes are the same as in Figure 6, and each class of landscape factors (i.e., x-axis) is defined in Table 1.