

**Interannual land cover and vegetation variability based on remote sensing data in the HTESSEL land surface model implementation and effects on simulated water dynamics**

van Oorschot, Fransje; van der Ent, Ruud J.; Hrachowitz, M.; Di Carlo, Emanuele; Catalano, Franco; Boussetta, Souhail; Balsamo, Gianpaolo; Alessandri, Andrea

**DOI**

[10.5194/esd-14-1239-2023](https://doi.org/10.5194/esd-14-1239-2023)

**Publication date**

2023

**Document Version**

Final published version

**Published in**

Earth System Dynamics

**Citation (APA)**

van Oorschot, F., van der Ent, R. J., Hrachowitz, M., Di Carlo, E., Catalano, F., Boussetta, S., Balsamo, G., & Alessandri, A. (2023). Interannual land cover and vegetation variability based on remote sensing data in the HTESSEL land surface model: implementation and effects on simulated water dynamics. *Earth System Dynamics*, 14(6), 1239-1259. <https://doi.org/10.5194/esd-14-1239-2023>

**Important note**

To cite this publication, please use the final published version (if applicable). Please check the document version above.

**Copyright**

Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

**Takedown policy**

Please contact us and provide details if you believe this document breaches copyrights. We will remove access to the work immediately and investigate your claim.



# Interannual land cover and vegetation variability based on remote sensing data in the HTESSEL land surface model: implementation and effects on simulated water dynamics

Fransje van Oorschot<sup>1,2</sup>, Ruud J. van der Ent<sup>2</sup>, Markus Hrachowitz<sup>2</sup>, Emanuele Di Carlo<sup>1</sup>, Franco Catalano<sup>3</sup>, Souhail Boussetta<sup>4</sup>, Gianpaolo Balsamo<sup>4</sup>, and Andrea Alessandri<sup>1</sup>

<sup>1</sup>Institute of Atmospheric Sciences and Climate, National Research Council of Italy (CNR-ISAC), Bologna, Italy

<sup>2</sup>Department of Water Management, Faculty of Civil Engineering and Geosciences, Delft University of Technology, Delft, the Netherlands

<sup>3</sup>Italian National Agency for New Technologies, Energy and Sustainable Economic Development (ENEA), Rome, Italy

<sup>4</sup>European Centre for Medium-Range Weather Forecasts (ECMWF), Reading, UK

**Correspondence:** Fransje van Oorschot (f.vanoorschot@tudelft.nl)

Received: 21 April 2023 – Discussion started: 22 May 2023

Revised: 4 October 2023 – Accepted: 9 October 2023 – Published: 29 November 2023

**Abstract.** Vegetation largely controls land surface–atmosphere interactions. Although vegetation is highly dynamic across spatial and temporal scales, most land surface models currently used for reanalyses and near-term climate predictions do not adequately represent these dynamics. This causes deficiencies in the variability of modeled water and energy states and fluxes from the land surface. In this study we evaluated the effects of integrating spatially and temporally varying land cover and vegetation characteristics derived from satellite observations on modeled evaporation and soil moisture in the Hydrology Tiled ECMWF Scheme for Surface Exchanges over Land (HTESSEL) land surface model. Specifically, we integrated interannually varying land cover from the European Space Agency Climate Change Initiative and interannually varying leaf area index (LAI) from the Copernicus Global Land Services (CGLS). Additionally, satellite data on the fraction of green vegetation cover (FCover) from CGLS were used to formulate and integrate a spatially and temporally varying effective vegetation cover parameterization. The effects of these three implementations on model evaporation fluxes and soil moisture were analyzed using historical offline (land-only) model experiments at the global scale, and model performances were quantified with global observational products of evaporation ( $E$ ) and near-surface soil moisture ( $SM_s$ ). The interannually varying land cover consistently altered the evaporation and soil moisture in regions with major land cover changes. The interannually varying LAI considerably improved the correlation of  $SM_s$  and  $E$  with respect to the reference data, with the largest improvements in semiarid regions with predominantly low vegetation during the dry season. These improvements are related to the activation of soil moisture–evaporation feedbacks during vegetation–water-stressed periods with interannually varying LAI in combination with interannually varying effective vegetation cover, defined as an exponential function of LAI. The further improved effective vegetation cover parameterization consistently reduced the errors of model effective vegetation cover, and it regionally improved  $SM_s$  and  $E$ . Overall, our study demonstrated that the enhanced vegetation variability consistently improved the near-surface soil moisture and evaporation variability, but the availability of reliable global observational data remains a limitation for complete understanding of the model response. To further explain the improvements found, we developed an interpretation framework for how the model development activates feedbacks between soil moisture, vegetation, and evaporation during vegetation water stress periods.

## 1 Introduction

Land surface–atmosphere interactions are largely controlled by vegetation, which is dynamic across spatial (local, regional, and global) and temporal (seasonal, interannual, and decadal) scales (Seneviratne et al., 2010). Land surface models (LSMs) aim to describe these interactions and are therefore a crucial aspect of models used for climate reanalysis and climate predictions. However, most state-of-the-art LSMs do not adequately represent the temporal and spatial variability of vegetation, resulting in weaknesses in the associated variability of modeled surface water and energy states and fluxes (e.g., Alessandri et al., 2007; Pitman et al., 2009; Ukkola et al., 2016; Fisher and Koven, 2020; Hersbach et al., 2020; van Oorschot et al., 2021).

To improve the representation of land surface–atmosphere dynamics, products based on satellite remote sensing data have been widely used in LSMs. Global satellite-derived maps of land cover and albedo have been directly used as boundary conditions (Faroux et al., 2013; Alessandri et al., 2017; Boussetta et al., 2021). In addition, leaf area index (LAI) derived from satellite remote sensing has been assimilated in several LSMs for different spatial scales, generally leading to improved water, energy, and carbon fluxes (Kumar et al., 2019; Ling et al., 2019; Rahman et al., 2020, 2022). Albergel et al. (2017, 2018) also combined LAI assimilation with the assimilation of remote-sensing-based surface soil moisture in the LSM called ISBA (interactions between soil–biosphere–atmosphere). This resulted in reduced errors of modeled soil moisture, evaporation, river discharges, and gross primary production with respect to observations. Furthermore, satellite products have been used to improve model parameterizations of, for example, leaf phenology, surface roughness, soil characteristics, and subsurface water storage (Lo et al., 2010; Trigo et al., 2015; MacBean et al., 2015; Yang et al., 2016; Orth et al., 2017). Moreover, LSMs have been evaluated using global satellite products of, e.g., land surface temperatures, snow depth, and soil moisture (Balsamo et al., 2018; Johannsen et al., 2019; Dong et al., 2020; Nogueira et al., 2020, 2021; Boussetta et al., 2021).

Recent studies have exploited the latest satellite campaigns to update land cover (LC) and leaf area index (LAI) representation into the land surface model Carbon-Hydrology ECMWF Tiled Scheme for Surface Exchanges over Land (CHTESSEL) (Johannsen et al., 2019; Nogueira et al., 2020, 2021; Boussetta et al., 2021) as part of the Integrated Forecasting System (IFS) of the European Centre for Medium-Range Weather Forecasts (ECMWF). These studies replaced the original fixed map of land cover from the Global Land Cover Characteristics (GLCC) dataset representing the early 1990s (Loveland et al., 2000) with an updated map obtained from the latest-generation estimates of land cover from the European Space Agency Climate Change Initia-

tive (ESA-CCI) (Poulter et al., 2015). Similarly, the LAI climatology from the Moderate Resolution Imaging Spectroradiometer (MODIS) (Boussetta et al., 2013) was replaced with updated climatology from the recent Copernicus Global Land Service (CGLS) LAI dataset (Verger et al., 2014). The integration of these satellite-derived variables considerably reduced the bias of model land surface temperatures (Johannsen et al., 2019; Nogueira et al., 2020, 2021). In addition, Boussetta et al. (2021) showed an overall reduction of model annual mean evaporation bias when using the updated LC and LAI in CHTESSEL.

LAI in LSMs can be coupled to the effective vegetation cover ( $C_{\text{eff}}$ ), which characterizes the density of the vegetated surface from a top view that effectively contributes to the water and energy balances. The organization structure of leaves inside the canopy is reported as vegetation clumping. In previous modeling studies, the seasonal variations in  $C_{\text{eff}}$  have been described as an exponential function of LAI considering vegetation clumping in (C)HTESSEL (Alessandri et al., 2017; Nogueira et al., 2020; Boussetta et al., 2021) and in other land modeling efforts (Anderson et al., 2005; Krinner et al., 2005; Le Moigne, 2012). The shape of the exponential relation between  $C_{\text{eff}}$  and LAI in state-of-the-art land surface models has, to our knowledge, been assumed to be constant in time and space so far (Krinner et al., 2005; Alessandri et al., 2017; Nogueira et al., 2020; Boussetta et al., 2021). However, studies have shown that the degree of vegetation clumping, and therefore the shape of this relation, actually varies for different vegetation types (Chen et al., 2005; Ryu et al., 2010; Zhang et al., 2014).

The research gap that we identified is that most previous LSM studies using HTESSEL aimed at improving the temporally fixed boundary condition of land cover and the monthly seasonal cycle of LAI, while not exploring the effects of interannual variations of LC and LAI. Moreover, these studies have generally used one spatially fixed relationship between effective vegetation cover and LAI, while there is considerable evidence that this relationship is vegetation-type-dependent (Chen et al., 2005; Ryu et al., 2010; Zhang et al., 2014).

The objective of this research is to evaluate the effects of integrating temporal and spatial variations of land cover and vegetation characteristics derived from satellite observations on modeled evaporation and soil moisture in the land surface model HTESSEL. Specifically, we will integrate annually varying LC from ESA-CCI as well as seasonally and interannually varying LAI from CGLS. Additionally, the CGLS fraction of green vegetation cover (FCover; Verger et al., 2014) is used to formulate and implement a spatially (i.e., vegetation-dependent) and temporally (i.e., interannually) variable effective vegetation cover parameterization in HTESSEL.

## 2 Methods

This section describes how we integrated temporal and spatial variations of land cover and vegetation characteristics in HTESSEL. In Sect. 2.1 we describe the land cover and vegetation data used, in Sect. 2.2 we describe the model characteristics with relevance to water dynamics, and in Sect. 2.3 the model developments performed in this study are reported. Finally, the model experiments and model evaluation are described in Sect. 2.4 and Sect. 2.5, respectively.

### 2.1 Land cover and vegetation data

Here we used yearly land cover maps at a 300 m spatial resolution from ESA-CCI for the time period 1993–2019 (Defourny et al., 2017; Copernicus Climate Change Service, 2019). In this dataset the land cover is classified into 22 classes based on the United Nations Land Cover Classification System (LCCS) (Di Gregorio and Jansen, 2005).

LAI and FCover data were obtained from CGLS for 1999–2019 (Copernicus Global Land Service, 2022). We used the 1 km version 2 collection in which both products were derived at a 10-daily resolution from the top-of-canopy reflectance measurements by the SPOT/VEGETATION (1999–2013) and PROBA-V (2014–2019) sensors (Verger et al., 2019). These two time series were homogenized using a cumulative distribution function (CDF) approach following Boussetta and Balsamo (2021). For model spin-up purposes, the CGLS LAI (1999–2019) was further extended backwards with former-generation LAI data from the Advanced Very-High-Resolution Radiometer (AVHRR) for 1993–1999 at a 0.05° resolution (Pacholczyk and Verger, 2020). The AVHRR LAI (1993–1999) was interpolated using conservative interpolation (Schulzweida, 2022) to the CGLS 1 km resolution and harmonized with CGLS (1999–2019) using CDF matching (Boussetta and Balsamo, 2021).

### 2.2 Relevant model components for water cycle representation

Here we used the HTESSEL land surface model (Balsamo et al., 2009) as it was developed and implemented for climate predictions with the EC-Earth3 Earth system model (Döscher et al., 2022). This version already implements a temporally, but not spatially, varying effective vegetation cover, which is further developed in this work (Alessandri et al., 2017). This section describes the relevant model representations of land cover (Sect. 2.2.1), leaf area index (Sect. 2.2.2), and effective vegetation cover (Sect. 2.2.3) in the current HTESSEL version as part of the EC-Earth3 Earth system model (ESM) and the role of these representations in the modeled water cycle. Section 2.3 describes the adaptations of these model components made in this study.

#### 2.2.1 Land cover representation

In HTESSEL the vegetated area of a grid cell is divided into high- and low-vegetation tiles. In the case of snow there are separate model tiles for snow on bare ground with low vegetation and snow beneath high vegetation (Balsamo et al., 2009). Figure 1a represents an example of the vegetation types and cover fractions for a single grid cell that were originally based on the GLCC land cover dataset (Loveland et al., 2000). The low-vegetation (L) and high-vegetation (H) types with the largest cover fraction in each grid cell (see example in Fig. 1a) are used in HTESSEL as dominant vegetation types  $T_L$  and  $T_H$  (Fig. 1b). The corresponding HTESSEL vegetation cover fractions  $A_L$  and  $A_H$  are based on the total low- and high-vegetation grid cell cover fractions.

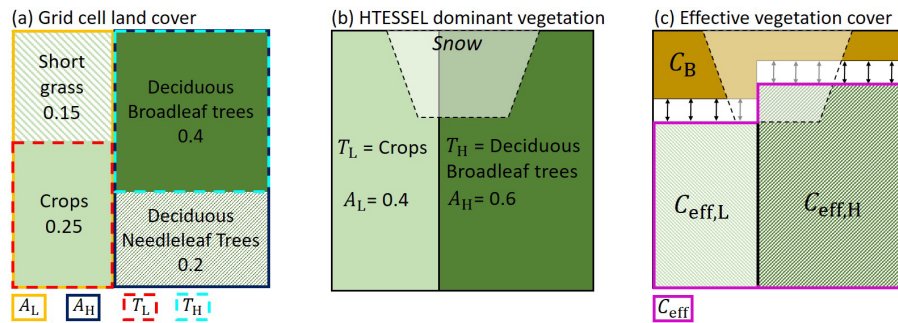
$T_L$  and  $T_H$  directly control surface water and energy fluxes because model parameters such as vegetation root distribution, minimum canopy resistance, and roughness lengths for momentum and heat are obtained from lookup tables based on the vegetation type (ECMWF, 2015). Surface fluxes are calculated separately for low- and high-vegetation tiles and combined based on the fractions  $A_L$  and  $A_H$ . Here we only focus on the surface evaporation flux that we define as the sum of transpiration, soil evaporation, interception evaporation, and, in the case of lakes, also open-water evaporation (Savenije, 2004; Miralles et al., 2020). The subsurface in HTESSEL consists of four soil layers with thicknesses of 7, 21, 72, and 189 cm, totaling a depth of 289 cm. In this study we differentiate between near-surface soil moisture ( $SM_s$ ) in the top layer (0–7 cm) and the subsurface soil moisture ( $SM_{sb}$ ) in the three deeper layers (7–289 cm).

#### 2.2.2 Leaf area index representation

LAI is defined separately for the high- and low-vegetation tiles ( $LAI_L$  and  $LAI_H$ ). In the original HTESSEL model,  $LAI_L$  and  $LAI_H$  are prescribed as a seasonal cycle that is derived from a satellite-based climatology based on MODIS (Boussetta et al., 2013) and the vegetation cover fractions  $A_L$  and  $A_H$ . The LAI controls the canopy resistance  $r_c$  of the high- and low-vegetation tiles through the following linear relation:

$$r_c = \frac{r_{s,\min}}{LAI} f_1(R_s) f_2(D_a) f_3(\overline{SM}), \quad (1)$$

with  $r_{s,\min}$  being the prescribed vegetation-specific minimum canopy resistance that does not change in time and  $f_1(R_s)$ ,  $f_2(D_a)$ , and  $f_3(\overline{SM})$  being functions describing the dependencies on shortwave radiation ( $R_s$ ), atmospheric water vapor deficit ( $D_a$ ), and weighted average soil moisture based on the root distribution over the four soil layers ( $\overline{SM}$ ), respectively. The effects of  $CO_2$  changes on  $r_c$  are not explicitly taken into account in the present study. The root fractions are generally the largest in soil layers 2 and 3, and therefore transpiration mostly originates from the  $SM_{sb}$ . The transpiration



**Figure 1.** Vegetation representation in a grid cell with example vegetation types and cover fractions. (a) Grid cell vegetation type and cover fraction based on land cover dataset. (b) HTESSEL dominant low- and high-vegetation type ( $T_L$  and  $T_H$ ) and cover fraction ( $A_L$  and  $A_H$ ). (c) HTESSEL effective vegetation cover with  $C_{\text{eff,L}}$  and  $C_{\text{eff,H}}$  being the effective low- and high-vegetation cover fraction,  $C_B$  the bare soil fraction, and  $C_{\text{eff}} = C_{\text{eff,L}} + C_{\text{eff,H}}$ , with the arrows indicating the temporal variability of  $C_{\text{eff}}$  as discussed in Sect. 2.2.3.

is linearly related to  $r_c$  and atmospheric variables. Furthermore, the LAI controls the capacity of the model interception reservoir  $W_{1m}$  by

$$W_{1m} = W_{1\text{max}} \cdot (C_B + C_L \cdot \text{LAI}_L + C_H \cdot \text{LAI}_H), \quad (2)$$

with  $W_{1\text{max}} = 0.0002 \text{ m}$  and  $C_B$ ,  $C_L$ , and  $C_H$  being the fractions of bare soil, effective low vegetation, and effective high vegetation, respectively (Sect. 2.2.3). The interception evaporation per time step follows from the water content of the interception reservoir (calculated from precipitation),  $W_{1m}$ , and the potential evaporation.

### 2.2.3 Effective vegetation cover representation

The model effective low vegetation cover and high vegetation cover ( $C_{\text{eff,L}}$  and  $C_{\text{eff,H}}$ ) represent the part of the model vegetation cover fraction ( $A_L$  and  $A_H$ ) that is actively contributing to the water balance through transpiration and interception evaporation (Fig. 1c). The fraction of the grid cell not covered by effective vegetation is treated as bare soil ( $C_B$ ), where only soil evaporation takes place. Soil evaporation only occurs in the top soil layer (0–7 cm) and therefore originates only from  $\text{SM}_s$ . The model resistance to soil evaporation ( $r_{\text{soil}}$ ) is described by

$$r_{\text{soil}} = r_{\text{soil,min}} f_3(\text{SM}_s), \quad (3)$$

with  $r_{\text{soil,min}} = 50 \text{ s m}^{-1}$  and  $f_3(\text{SM}_s)$  representing the dependency on the first layer soil moisture content. The effective vegetation cover fractions  $C_{\text{eff,L}}$  and  $C_{\text{eff,H}}$  as well as bare soil fraction  $C_B$  are described by

$$C_{\text{eff,L}} = c_{v,L} \cdot A_L, \quad (4)$$

$$C_{\text{eff,H}} = c_{v,H} \cdot A_H, \quad (5)$$

$$C_{\text{eff}} = C_{\text{eff,L}} + C_{\text{eff,H}}, \quad (6)$$

$$C_B = 1 - C_{\text{eff}}, \quad (7)$$

with  $c_{v,L}$  and  $c_{v,H}$  representing the low and high vegetation density. Originally,  $c_{v,L}$  and  $c_{v,H}$  were described by a

lookup table with vegetation-specific values, allowing for spatial variation of the  $C_{\text{eff,L}}$ ,  $C_{\text{eff,H}}$ , and  $C_B$  fractions. However, this approach does not represent temporal variations in vegetation density. To allow for temporal variability in  $C_{\text{eff}}$  (represented by the arrows in Fig. 1c),  $c_{v,L}$  and  $c_{v,H}$  were linked to the seasonal variability of LAI by the following exponential relation (Alessandri et al., 2017):

$$c_{v,L} = 1 - e^{-k \text{LAI}_L}, \quad (8)$$

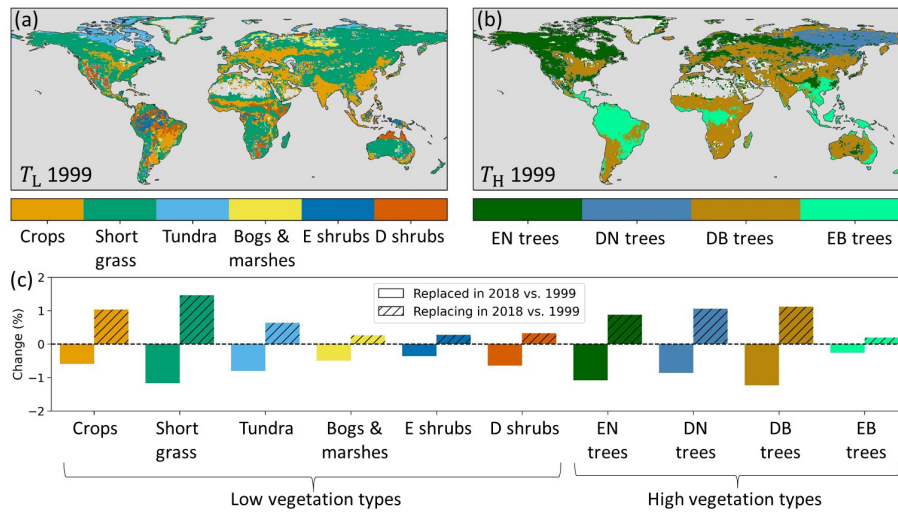
$$c_{v,H} = 1 - e^{-k \text{LAI}_H}, \quad (9)$$

with  $k$  being the canopy light extinction coefficient that represents the degree of vegetation clumping (Anderson et al., 2005). Previously  $k$  was generally set to constant values of 0.5 (Krinner et al., 2005; Alessandri et al., 2017) or 0.6 (Boussetta et al., 2021) for all vegetation types. As a consequence, the vegetation-dependent spatial variability in  $k$  was not accounted for.

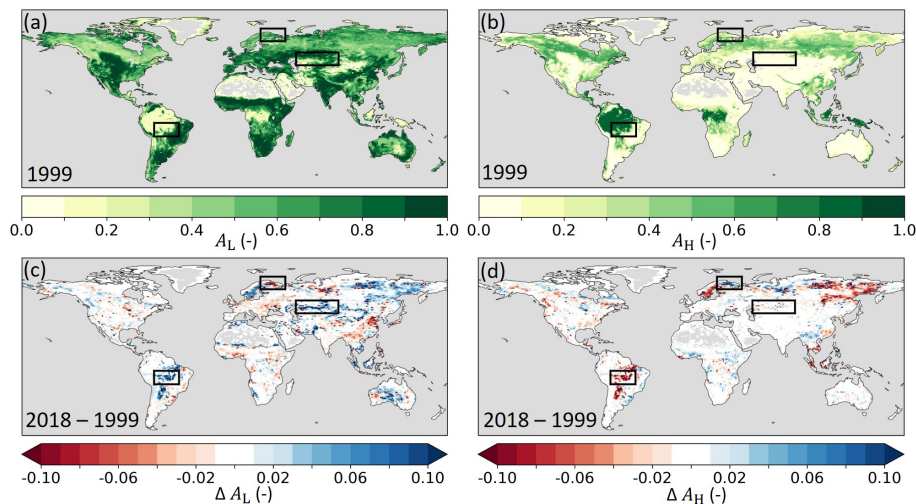
## 2.3 Model developments

### 2.3.1 The implemented land cover variability

Here we implemented the annually varying ESA-CCI land cover (LC) data for the 1993–2019 period (Sect. 2.1), as developed by Boussetta and Balsamo (2021), for the HTESSEL vegetation types and spatial resolution. For consistency with the other model adaptations and evaluations (Sect. 2.3.2, 2.3.3 and 2.5), our LC analyses were based on 1999–2018. The interannually varying LC from ESA-CCI introduced a change in  $T_L$  in 5 % and  $T_H$  in 4 % of the land grid cells between the first (1999) and the last (2018) year of the considered study period (Fig. 2). Figure 2c shows the fraction of land grid cells in which each vegetation type (dominant in 1999) is replaced by another type in 2018 (plain colors) and conversely how often each vegetation type replaces the 1999 dominant one in 2018 (hatched colors). The figure shows that crops and short grass relatively often replaced other low-vegetation types (relatively large hatched bars), while ever-



**Figure 2.** (a) Model low ( $T_L$ ) and (b) high ( $T_H$ ) dominant vegetation types in 1999 based on ESA-CCI land cover. (c) Changes in low- and high-vegetation type as a percent of the total land points, with plain colors indicating that the vegetation type was replaced in 2018 compared to 1999 and hatched colors that the vegetation replaced another type in 2018 compared to 1999. Note that low vegetation and high vegetation are treated separately and do not replace each other. E stands for evergreen, D for deciduous, N for needleleaf, and B for broadleaf.



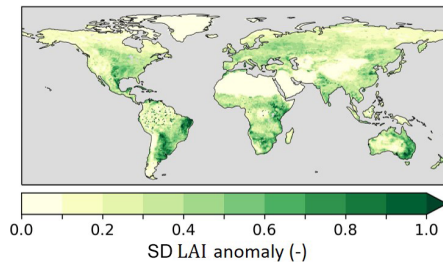
**Figure 3.** (a) Model low- ( $A_L$ ) and (b) high-vegetation ( $A_H$ ) cover fraction in 1999 as well as the absolute difference in (c)  $A_L$  and (d)  $A_H$  between 2018 and 1999 (2018–1999) based on ESA-CCI land cover. Blue (red) indicates an increased (reduced) cover in 2018. The black boxes highlight the three regions of the southern Amazon, Lapland, and central Asia that are further analyzed in Sect. 3.1.

green needleleaf (EN) and deciduous broadleaf (DB) trees were relatively often replaced by other high-vegetation types (relatively large plain bars). The low- and high-vegetation cover fractions changed in many regions according to the ESA-CCI LC dataset (Fig. 3). During the 1999–2018 period, low vegetation replaced high vegetation in the southern Amazon and northeastern Siberia. Conversely, high vegetation replaced low vegetation in the boreal regions of Lapland and northwestern Siberia. Moreover, arid regions such as central Asia and Australia experienced an expansion of low vegetation over the 1999–2018 period. In Fig. 3 we highlight the southern Amazon, Lapland, and central Asia where the veg-

etation cover fraction changed considerably. These regions are further analyzed in Sect. 3.1.

### 2.3.2 The implemented leaf area index variability

We used the monthly CGLS LAI data described in Sect. 2.1 to prescribe model LAI, representing both the seasonal cycle and interannual variability of LAI. The 1 km LAI data were interpolated using conservative interpolation to the HTESSEL grid (Schulzweida, 2022). Next, LAI was disaggregated into low and high LAI ( $LAI_L$  and  $LAI_H$ ) based on the low- and high-vegetation cover fractions ( $A_L$  and  $A_H$ ) for use



**Figure 4.** Standard deviation (SD) of monthly interannual anomaly CGLS LAI for 1999–2018 as implemented in experiment IAK5 (Table 1).

in the HTESSEL model setup with separate low- and high-vegetation tiles (Boussetta et al., 2021; Boussetta and Balsamo, 2021). Figure 4 shows the LAI interannual variability as integrated here in HTESSEL, quantified with the standard deviation. The effects of this added variability are presented in Sect. 3.2.

### 2.3.3 The implemented vegetation-specific effective vegetation cover parameterization

The CGLS FCover and LAI data were used (Sect. 2.1) to further develop the model effective vegetation cover parameterization as described by Eqs. (4)–(9). The constant  $k = 0.5$  parameter was replaced with a vegetation-specific  $k$  to improve spatial and temporal variability of the model  $C_{\text{eff}}$ . We assumed that the model  $C_{\text{eff}}$  is equivalent to the CGLS FCover data. Following the model  $C_{\text{eff}}$  parameterization, FCover is then described as follows:

$$\text{FCover} = 1 - e^{-k \text{LAI}}. \quad (10)$$

We estimated  $k$  for different HTESSEL vegetation types by solving the minimization problem in Eq. (11) using a nonlinear least-squares optimization at a 1 km spatial resolution.

$$\min \|1 - e^{-k \text{LAI}} - \text{FCover}\|_2 \quad (11)$$

To discriminate vegetation types, the grid cells where each vegetation type maximizes its cover fraction based on the ESA-CCI LC developed in Boussetta and Balsamo (2021) were selected for each year. For each set of grid cells corresponding to each vegetation type, the FCover and LAI 10-daily 1 km data for 1999–2019 were extracted. Here we used a  $1 \times 1$  km resolution for LAI, FCover, and LC in order to obtain the most representative discrimination of vegetation types and to minimize vegetation mixing within each resolved grid cell. For the optimization of  $k$ , a randomly selected subsample of 2000 grid points of the LAI and FCover time steps (10-daily) for each vegetation type was used to keep the analysis computationally feasible, while ensuring a representative sample with robust significance of the fit. In this way, we obtained a sample of 2000 grid cells times 36

time steps per year times 20 years, which equals 1 440 000 data points to be used for the optimization for each vegetation type. This optimization resulted in vegetation-specific  $k$  values that were implemented in the HTESSEL code as in Eqs. (4)–(9). The robustness of the optimization was verified by repeating the random selection procedure several times, which resulted in negligible changes in the  $k$  estimates.

## 2.4 Model experiments

We performed experiments with an offline, uncoupled version of HTESSEL to evaluate the effect of the implemented vegetation variability as described in Sect. 2.3. HTESSEL was forced with atmospheric hourly forcing from ECMWF Reanalysis v5 (ERA5) and simulations were performed from 1993–2019, with 1993–1999 as the spin-up period (details in Table S1 in the Supplement). The model spatial resolution is the  $n_{128}$  reduced Gaussian grid corresponding to grid cells of  $\sim 75 \times 75$  km. In total, four different model experiments were performed (Table 1). In the first experiment, as a benchmark and control experiment (CTR) the land cover of all years was set to the ESA-CCI land cover of 1993, the LAI of all years was set to the 1993–2019 climatology, and the  $C_{\text{eff}}$  parameterization with  $k = 0.5$  was used. This reflects standard settings of the EC-Earth3 version of HTESSEL. In the second experiment (IALC) the interannually varying ESA-CCI LC was included, while in the third experiment (IAK5) we further added interannually varying CGLS LAI. Finally, the model sensitivity to the vegetation-specific  $C_{\text{eff}}$  parameterization (see Sect. 2.2.3) was evaluated in the fourth experiment (IAKV). The model experiments were evaluated for 1999–2018, which is the longest period to consistently assess all three model implementations with the available evaluation data described in Sect. 2.5.

## 2.5 Model evaluation

### 2.5.1 Model variables

The effects of the vegetation-specific  $C_{\text{eff}}$  parameterization on the model  $C_{\text{eff}}$  were assessed in IAKV compared to IAK5. Furthermore, we analyzed the effects of the increasingly detailed model land cover and vegetation variability in the three experiments (IALC, IAK5, IAKV) on total evaporation ( $E$ ) and the evaporation components, i.e., transpiration ( $E_t$ ), soil evaporation ( $E_s$ ), and interception evaporation ( $E_i$ ). In addition, the effects on model near-surface soil moisture ( $\text{SM}_s$ ) and subsurface soil moisture ( $\text{SM}_{\text{sb}}$ ) were analyzed.

### 2.5.2 Reference data

The modeled  $C_{\text{eff}}$  was compared to the CGLS FCover data (Sect. 2.1) at the model spatial resolution. As a benchmark for total evaporation we used the Derived Optimal Linear Combination Evapotranspiration version 3 (DOLCEv3), which is a linear combination of estimates from ERA5-land,

**Table 1.** Details of model experiments.

Experiment	Land cover	Leaf area index	Effective vegetation cover
CTR	ESA-CCI fixed	CGLS climatology	$k = 0.5$
IALC	ESA-CCI interannual	CGLS climatology	$k = 0.5$
IAK5	ESA-CCI interannual	CGLS interannual	$k = 0.5$
IAKV	ESA-CCI interannual	CGLS interannual	$k$ vegetation-specific

GLEAM v3.5a and v3.5b, and FLUXCOM-RSMETEO that was regionally weighted based on the performance in reproducing FLUXNET tower evaporation observations (Hobeichi et al., 2021). The associated uncertainty estimate is spatially and temporally varying based on the gridded evaporation and flux tower observations (Hobeichi et al., 2018). This dataset was selected because it is intended to better capture evaporation temporal variations compared to previous DOLCE versions (v1 and v2) and was therefore found to be suitable for evaluating the effects of the modified temporal and spatial variability of vegetation on evaporation (Hobeichi et al., 2021). Daily evaporation and associated uncertainty at a  $0.25^\circ$  resolution were used for 1999–2018 and were interpolated here using conservative interpolation (Schulzweida, 2022) to the model spatial resolution.

Model near-surface soil moisture ( $SM_s$ ) (0–7 cm) was compared to the combined active–passive ESA-CCI soil moisture product (ESA-CCI SM v06.1), which is generated from satellite-based active and passive microwave products that are combined using the absolute values and dynamic range of the modeled soil moisture of the top 10 cm soil layer from the Global Land Data Assimilation System (GLDAS)-Noah LSM (Liu et al., 2012; Dorigo et al., 2017; Gruber et al., 2017). This dataset provides near-surface ( $\sim 0$ –5 cm) soil moisture at a daily resolution on a  $0.25^\circ$  grid. Here we used the combined active–passive product interpolated using conservative interpolation (Schulzweida, 2022) to the model spatial resolution ( $\sim 75 \times 75$  km) for 1999–2018 (European Space Agency, 2022). The uncertainty estimates for ESA-CCI SM were also considered as they were provided with the data product and based on error variance of the data used to generate the product (Dorigo et al., 2017). ESA-CCI SM contains spatial and temporal gaps due to densely vegetated areas (tropical forests) and snow coverage. Here only grid cells with a temporal coverage larger than 60 % were used, and, as a consequence, model performance metrics for  $SM_s$  were only calculated for these grid cells.

### 2.5.3 Evaluation metrics

The hourly model output was first averaged to monthly values, based on which annual means, monthly climatology, and interannual anomalies were then calculated. To differentiate the seasons (June, July, and August: JJA; September, October and November: SON; December, January, and February:

DJF; March, April, and May: MAM), the monthly values were averaged to seasonal means, and interannual seasonal anomalies were calculated. For the evaluation of  $E$  and  $SM_s$  with respect to reference data, we used the Pearson correlation coefficients  $r$  of the interannual monthly and seasonal anomalies. To calculate  $r$  of the interannual monthly and seasonal anomalies, the anomalies were detrended assuming a linear trend. Detrending was not applied for the effects of the modified LC, as the interannually varying LC mostly influenced the trend. In addition, we quantified the effects of the improved vegetation variability with the root mean squared error (RMSE). For  $C_{\text{eff}}$  and  $E$  RMSE we used monthly values, while for  $SM_s$  we used standardized interannual anomalies. Model  $SM_s$  and reference ESA-CCI SM cannot be compared directly in absolute terms due to the different representative soil layers and the imposed dynamic range from the GLDAS-Noah model (Liu et al., 2012), potentially resulting in different temporal variability (Sect. 2.5.2). To overcome this limitation, we standardized the interannual anomalies for model and reference  $SM_s$  by dividing the monthly  $SM_s$  by the climatological standard deviation.

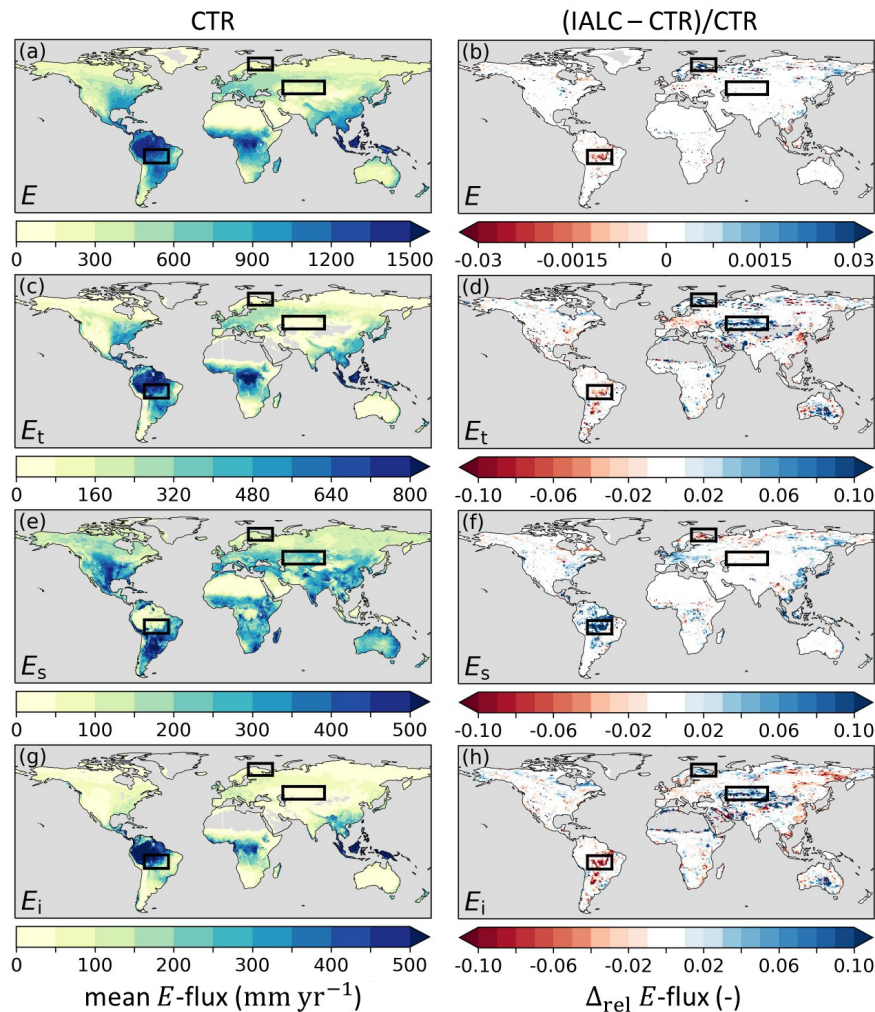
To test the significance of the  $r$  and RMSE differences between the experiments we used a bootstrap, in which 1000 data samples were randomly created by resampling the data of model 1 and model 2 with replacement for each time step. We tested the null hypothesis that the  $r$  and RMSE of model 1 and model 2 with respect to the reference data are not significantly different from each other at the 10 % significance level.

## 3 Results

### 3.1 Land cover interannual variability effects

The interannually varying land cover from ESA-CCI in experiment IALC resulted in a shift in mean evaporation components (i.e.,  $E_t$ ,  $E_s$ , and  $E_i$ ) compared to the CTR experiment (Fig. 5). The last 5 years of the simulations (2014–2019) are considered because the effects of the interannually varying land cover mostly emerge in this period. In the southern Amazon, where  $A_H$  was reduced on average from 0.64 to 0.57 in IALC compared to CTR (Fig. 3), the mean  $E_t$  was reduced by 3 % from 633 to 615  $\text{mm yr}^{-1}$  and  $E_i$  was reduced by 6 % from 384 to 363  $\text{mm yr}^{-1}$ , while  $E_s$  increased by 17 % from 156 to 183  $\text{mm yr}^{-1}$ . In this region, the total evapora-



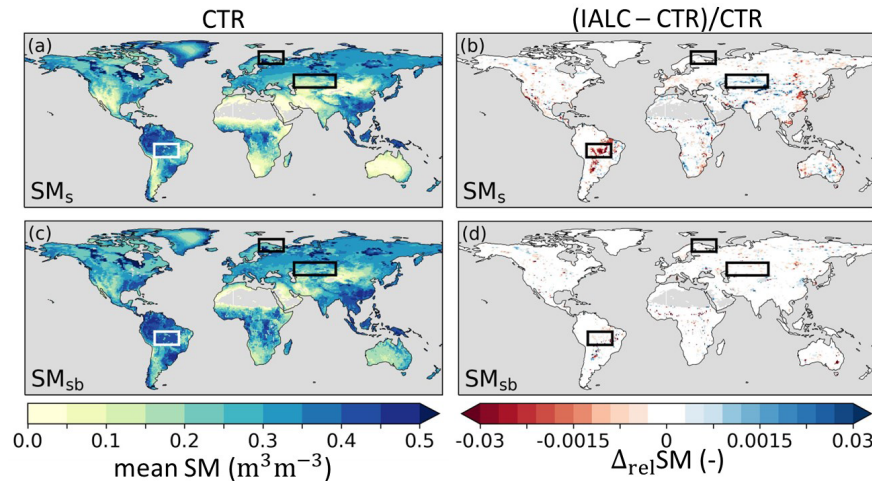


**Figure 5.** Annual mean evaporation fluxes (2014–2018) in experiment CTR with (a) total evaporation ( $E$ ), (c) transpiration ( $E_t$ ), (e) soil evaporation ( $E_s$ ), and (g) interception evaporation ( $E_i$ ), as well as the relative difference ( $\Delta_{\text{rel}}$ ) between annual mean evaporation fluxes in experiments IALC and CTR  $((\text{IALC}-\text{CTR})/\text{CTR})$  for (b)  $E$ , (d)  $E_t$ , (f)  $E_s$ , and (h)  $E_i$ . Blue (red) indicates an increased (reduced) flux. Grey land areas indicate regions with annual mean  $E$  fluxes  $< 0.1 \text{ mm yr}^{-1}$ . The boxes highlight the three regions of the southern Amazon, Lapland, and central Asia with major land cover changes (Fig. 3). Results with respect to DOLCEv3  $E$  are presented in Fig. S1. See Table 1 for details of the experiments.

tion was reduced only by 1% from 1174 to 1162  $\text{mm yr}^{-1}$  in IALC compared to CTR because the reductions in  $E_t$  and  $E_i$  were partially compensated for by increased  $E_s$ . The reduced  $E$  in IALC is closer to the DOLCEv3  $E$ , which in this region is 1160  $\text{mm yr}^{-1}$ . We also found an evaporation compensation effect in Lapland, where  $A_H$  increased from 0.24 to 0.30, and central Asia, where  $A_L$  increased from 0.66 to 0.71 (Fig. 3). In Lapland and in central Asia  $E$  increased by 2% and 0.1%, respectively, moving closer to the DOLCEv3  $E$  (Fig. 5b; Table S2). In contrast to the small changes in  $E$ , the individual  $E$  fluxes changed considerably in these two cases (Fig. 5d, f, h).

The changes in  $E_t$  and  $E_s$  also induced changes in soil moisture because  $E_s$  extracts water exclusively from the near-surface soil layer ( $\text{SM}_s$ ), while  $E_t$  originates mostly

from deeper soil layers ( $\text{SM}_{\text{sb}}$ ). However, we observed only marginal differences between mean  $\text{SM}_s$  and  $\text{SM}_{\text{sb}}$  in IALC compared to CTR (Fig. 6), except for the southern Amazon  $\text{SM}_s$ . The increased  $E_s$  in the southern Amazon reduced the  $\text{SM}_s$  by 2%, as more near-surface soil moisture was extracted (Fig. 6b). Individual evaporation fluxes influence the timing of total evaporation and soil moisture differently. However, the overall minor magnitude of changes in total  $E$  and  $\text{SM}_s$  in IALC compared to CTR led to marginal changes in RMSE and Pearson correlation coefficients with respect to the reference data in the three highlighted cases (Table S3, Figs. S1–S3).



**Figure 6.** Annual mean soil moisture (2014–2018) (SM) in experiment CTR with (a) near-surface soil moisture (SM<sub>s</sub>) and (c) subsurface soil moisture (SM<sub>sb</sub>), as well as the relative difference ( $\Delta_{\text{rel}}$ ) between annual mean SM in experiments IALC and CTR  $((\text{IALC}-\text{CTR})/\text{CTR})$  for (b) SM<sub>s</sub> and (d) SM<sub>sb</sub>. Blue (red) indicates increased (reduced) soil moisture. Grey land areas indicate regions with annual mean SM < 0.01 m<sup>3</sup> m<sup>-3</sup>. The boxes highlight the three regions of the southern Amazon, Lapland, and central Asia with major land cover changes (Fig. 3). See Table 1 for details of the experiments.

### 3.2 Leaf area index interannual variability effects

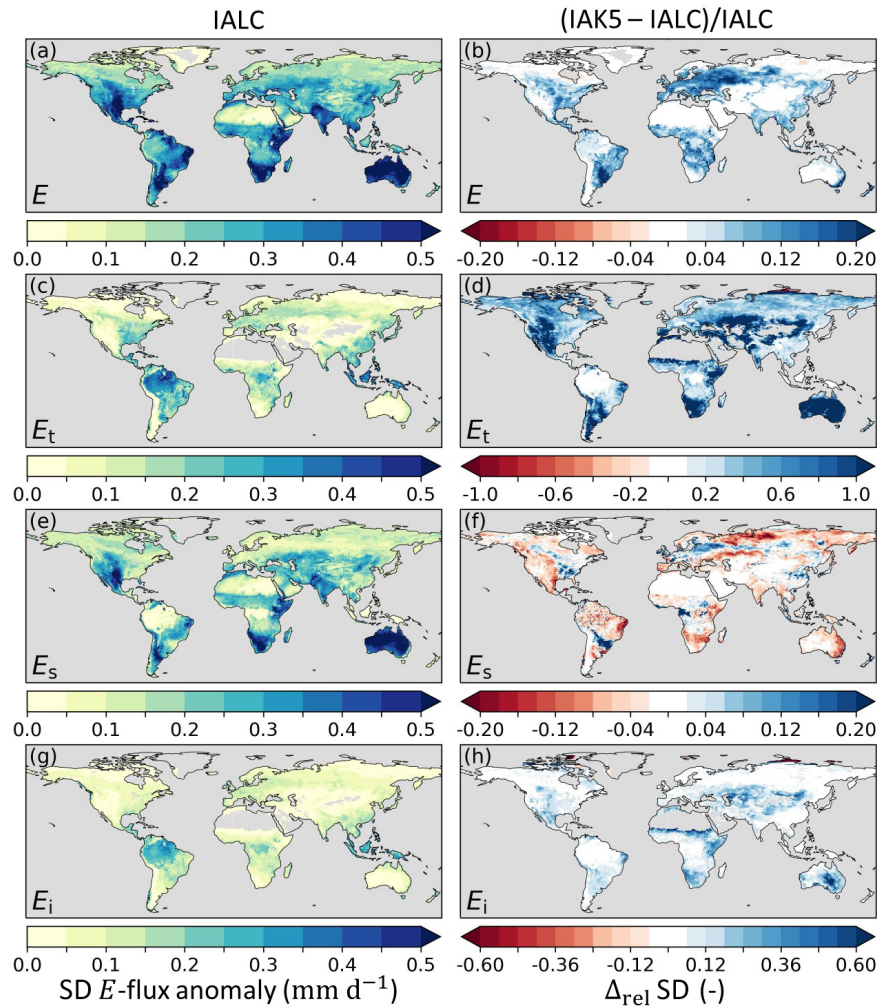
The inclusion of interannual LAI variability in IAK5 (Fig. 4) generally led to an increased anomaly standard deviation (i.e., variability) of  $E$  (Fig. 7a, b). This effect is mostly dominated by  $E_t$  (Fig. 7d), which, in the model, is linearly related to LAI (Eq. 1). Figure 7d and h show that the variability in  $E_t$  and  $E_i$  was mostly increased in semiarid regions such as the Great Plains region of the US, central Asia, and southern Africa, with a stronger effect for  $E_t$  than for  $E_i$ . In contrast, the  $E_s$  variability was reduced with the enhanced LAI variability in these semiarid regions but was increased in more temperate regions such as in Europe, the eastern US, and the La Plata Basin in South America (Fig. 7e, f). While the  $E_t$  anomaly variability considerably increased in IAK5 compared to IALC in semiarid regions, the anomaly variability in subsurface soil moisture (SM<sub>sb</sub>) that acts as the main source of  $E_t$  was reduced in these regions (Fig. 8c, d). On the other hand, the anomaly variability of SM<sub>s</sub> increased (Fig. 8a, b), while the  $E_s$  variability was reduced.

Figure 9 shows that the Pearson correlation coefficient ( $r$ ) of anomaly  $E$  with respect to DOLCEv3 increased in IAK5 compared to IALC in 85% of the land area in which the  $r$  was significantly different in IAK5 compared to IALC. Consistently, the  $r$  of anomaly SM<sub>s</sub> with respect to ESA-CCI SM also improved in 85% of the significantly changing land area. For both  $E$  and SM<sub>s</sub>  $r$  increased mostly in semiarid regions with predominantly low vegetation (Fig. 3).

### 3.3 Vegetation-specific effective vegetation cover parameterization effects

The observed relationship of LAI and FCover in Fig. 10 is broadly consistent with the shape of the exponential functions with the vegetation-specific  $k$ , with RMSEs between 0.018 and 0.053 for the individual vegetation types. All optimized LAI–FCover relations are characterized by  $k$  values that at 0.351–0.458 are consistently lower than the original  $k = 0.5$ , which has been used as the constant default value in most HTESSEL applications so far (Alessandri et al., 2017; Boussetta et al., 2021). We found that the  $k$  values for low-vegetation types (0.438–0.458) are higher than for high-vegetation types (0.351–0.396), except for tundra regions (0.375) (Fig. 10 and Table S3). These findings are in line with our expectations, as leaf organization of low vegetation is more regular (larger  $k$ ) than leaf organization of high vegetation, where leaves are found more on top of each other (smaller  $k$ ) (Chen et al., 2005, 2021).

The vegetation-specific  $C_{\text{eff}}$  parameterization (IAKV) generally reduced the  $k$  values compared to the  $k = 0.5$  setup (IAK5), and as a consequence the associated vegetation densities  $c_{v,L}$  and  $c_{v,H}$  also decreased (Eqs. 8 and 9). On average, the global mean  $C_{\text{eff}}$  was reduced from 0.21 in IAK5 to 0.19 in IAKV (Fig. 11). The reduced  $C_{\text{eff}}$  considerably reduced the RMSE with respect to the FCover data in IAKV compared to IAK5 (Fig. 12), as expected from the parameterization optimization presented in Fig. 10. The RMSE was reduced the most over the boreal and tropical forests, with an average RMSE reduction from 0.12 to 0.06 for evergreen needleleaf trees and from 0.05 to 0.03 for evergreen broadleaf trees. On the other hand, the differences in regions

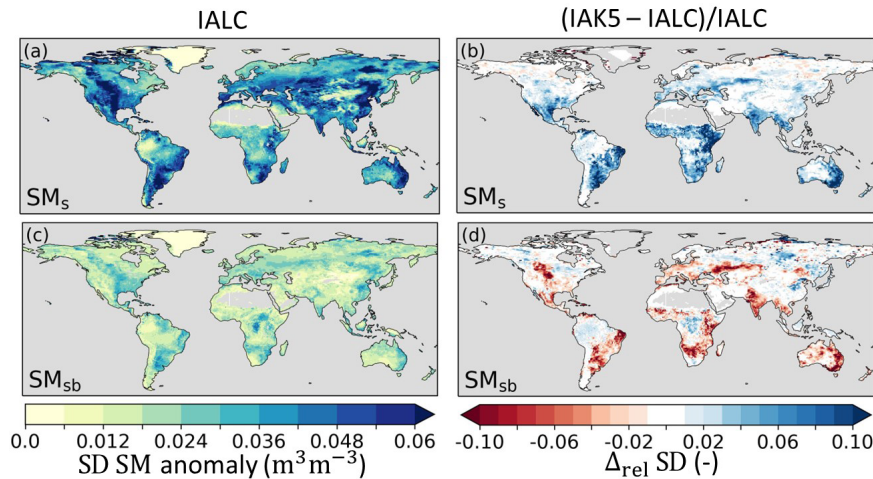


**Figure 7.** Standard deviation (SD) of anomaly evaporation fluxes in experiment IALC with (a) total evaporation ( $E$ ), (c) transpiration ( $E_t$ ), (e) soil evaporation ( $E_s$ ), and (g) interception evaporation ( $E_i$ ), as well as the relative difference ( $\Delta_{\text{rel}}$ ) between the anomaly  $E$  SD in experiments IAK5 and IALC ( $(\text{IAK5} - \text{IALC})/\text{IALC}$ ) for (b)  $E$ , (d)  $E_t$ , (f)  $E_s$ , and (h)  $E_i$ . Blue (red) indicates an increased (reduced) SD. See Table 1 for details of the experiments.

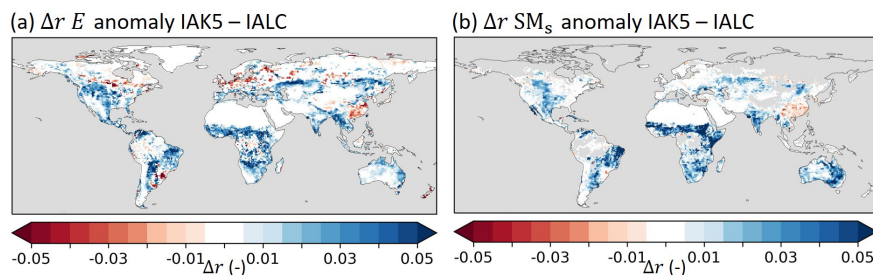
with predominantly low vegetation were smaller because the fitted  $k$  value was closer to the original  $k = 0.5$ , with an average RMSE reduction from 0.06 in IAK5 to 0.05 in IAKV for crops and from 0.04 to 0.03 for short grass. For low vegetation, the effects were not consistent throughout the seasons, with RMSE increasing at high latitudes in JJA (Fig. 12d). Here the  $C_{\text{eff}}$  in IAK5 was smaller than the CGLS FCover and is further reduced in IAKV, increasing the RMSE. This was likely caused by a poor fit for short grass at  $\text{LAI} > 2$  (Fig. 10b) and tundra at  $\text{LAI} > 1$  (Fig. 10g).

The reduced model  $C_{\text{eff}}$  in IAKV compared to IAK5 led to a shift in individual evaporation fluxes. On average,  $E_s$  increased and  $E_t$  and  $E_i$  were reduced, while the total  $E$  was not very affected (Fig. S6). These shifts led to changes in the temporal distribution of the evaporation. Figure 13 shows quite mixed results of the vegetation-specific  $C_{\text{eff}}$  parameterization for the model  $E$  RMSE with respect to DOLCEv3.

The RMSE was consistently reduced during summer months in temperate regions such as in Europe, the eastern US, and eastern China (JJA), as well as in southeastern Brazil and southern Africa (DJF). On the other hand, the results for tropical and boreal regions were less consistent throughout the seasons (Fig. 13). The effects of the vegetation-specific  $C_{\text{eff}}$  on  $\text{SM}_s$  RMSE with respect to ESA-CCI SM show consistent RMSE reductions in the JJA period for Canada and southeastern Brazil and in the DJF period for the Sahel (Fig. 14). Consistent with the  $C_{\text{eff}}$  RMSE increase in boreal regions in JJA (Fig. 12), the Pearson correlation coefficient for monthly anomaly  $E$  with respect to DOLCEv3  $E$  was significantly reduced in these regions in IAKV compared to IAK5, while other regions were not very affected (Fig. S14). On the other hand, the correlation of monthly anomaly  $\text{SM}_s$  with respect to ESA-CCI SM did not considerably change (Fig. S14).



**Figure 8.** Standard deviation (SD) of anomaly soil moisture (SM) in experiment IALC with (a) near-surface soil moisture ( $SM_s$ ) and (c) subsurface soil moisture ( $SM_{sb}$ ), as well as the relative difference ( $\Delta_{rel}$ ) between the anomaly SM SD in experiments IAK5 and IALC ( $(IAK5 - IALC)/IALC$ ) for (b)  $SM_s$  and (d)  $SM_{sb}$ . Blue (red) indicates an increased (reduced) variability. See Table 1 for details of the experiments.



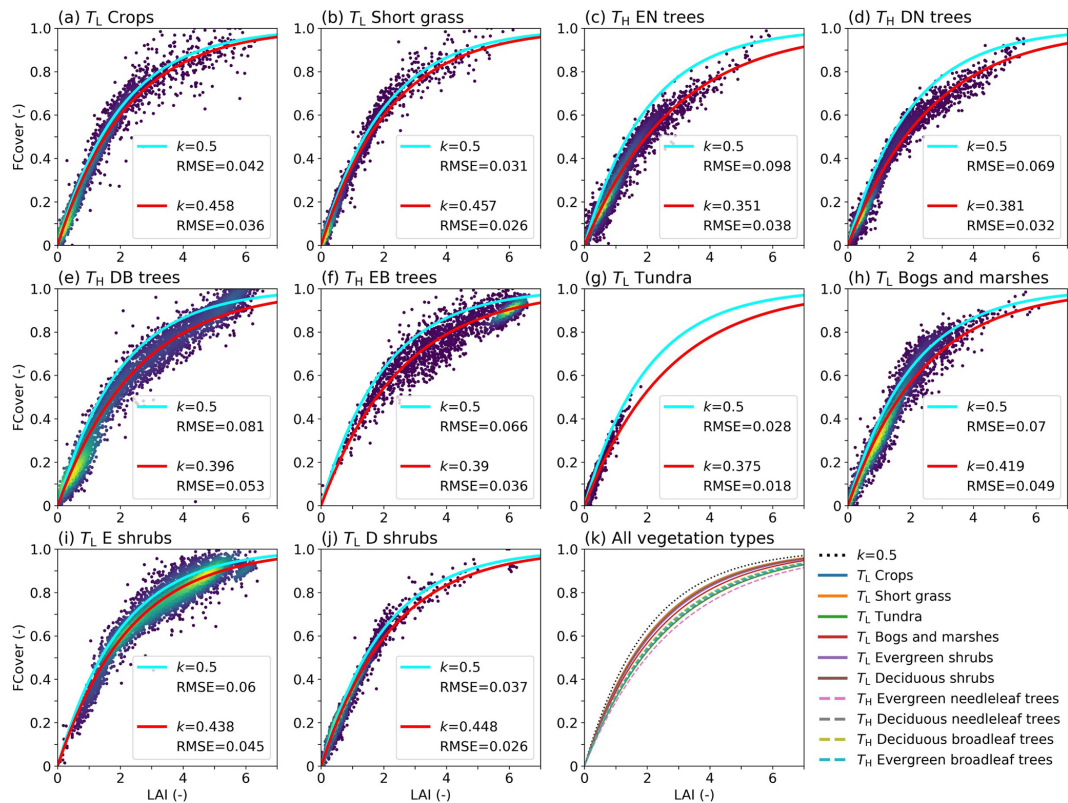
**Figure 9.** Pearson correlation difference ( $\Delta r$ ) between experiments IALC and IAK5 ( $IAK5 - IALC$ ) for (a) monthly anomaly total evaporation ( $E$ ) with respect to DOLCEv3 evaporation and (b) monthly anomaly near-surface soil moisture ( $SM_s$ ) with respect to ESA-CCI SM. Blue (red) indicates an increased (reduced) correlation in IAK5 compared to IALC, white indicates small and/or insignificant  $\Delta r$ , and grey indicates no data points. See Table 1 for details of the experiments. Similar figures for seasonal anomalies are presented in Figs. S4–S5.

### 3.4 Combined effects of land cover, leaf area index, and effective vegetation cover

The results presented in Sect. 3.2 demonstrate that the interannually varying LAI in experiment IAK5 considerably improved the correlation of  $E$  and  $SM_s$  with respect to reference data. On the other hand, the annually varying LC and vegetation-specific  $C_{eff}$  affected correlations merely to a minor degree (Sect. 3.1 and 3.3). Here, we further elaborate on the effects of combining the enhanced variability in LC, LAI, and  $C_{eff}$  on correlation of  $E$  and  $SM_s$ .

Figure 15 shows that the  $E$  correlation improved in 68 % (JJA) and 54 % (DJF) of the land area in which the  $r$  significantly changed in IAKV compared to CTR. Significant reduction of  $r$  is found over boreal regions, which is related to the effects of the effective vegetation cover parameterization, as discussed in Sect. 3.3 and shown in Fig. S14. Figures 15b and d show that the  $SM_s$  correlation consistently and significantly improved in 83 % (JJA) and 76 % (DJF) of the

land area in which the  $r$  significantly changed in IAKV compared to CTR. The  $E$  and  $SM_s$  correlations got consistently stronger during dry periods in regions with a semiarid climate and predominantly low vegetation (Figs. 3 and 15). For example, in northeastern Brazil during the dry JJA season, the correlation coefficient for  $E$  increased from  $r = 0.79$  in CTR to 0.84 in IAKV with respect to DOLCEv3 and for  $SM_s$  from  $r = 0.57$  to 0.67 with respect to ESA-CCI SM. Similarly, in western India during the dry DJF season, the correlation coefficient for  $E$  increased from  $r = 0.78$  to 0.85 and for  $SM_s$  from  $r = 0.45$  to 0.73. To further explore the effects in these semiarid regions, we zoom in to northeastern Brazil for the 2010–2013 period (Fig. 16). This period is characterized by positive LAI and  $C_{eff}$  anomalies in JJA 2011 and negative LAI and  $C_{eff}$  anomalies in JJA 2012 (Fig. 16a, b). The negative LAI and  $C_{eff}$  anomalies in 2012 characterize a dry period in which the negative  $E$  anomaly was magnified in IAKV compared to CTR (Fig. 16c). During this dry period,  $E_t$  was reduced, while  $E_s$  increased. This is consistent



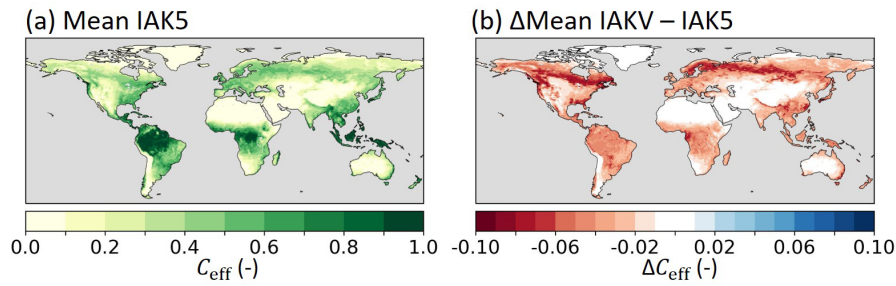
**Figure 10.** (a–j) LAI vs. FCover for a subsample (5000) of the selected points used for the least-squares optimization for all vegetation types with the optimized LAI–FCover relation in red (Eq. 10) and the  $k = 0.5$  relation in light blue, with RMSE values of the data points with respect to the curve. The colors indicate the point density, with purple indicating low density and yellow high density. (k) The optimized LAI–FCover relation for all vegetation types. E stands for evergreen, D for deciduous, N for needleleaf, and B for broadleaf. Values of  $k$  and RMSE are also presented in Table S3.

with the soil moisture response presented in Fig. 16d, as the  $SM_s$  was reduced (due to more  $E_s$ ) and the  $SM_{sb}$  increased (due to less  $E_t$ ) during the 2012 dry period. Opposite effects were found for the 2011 period with positive LAI and  $C_{eff}$  anomalies. So in this specific case, the variability in  $E_t$  and  $SM_s$  anomalies was enhanced in IAKV compared to CTR, while the variability in  $E_s$  and  $SM_{sb}$  anomalies was dampened. This is consistent with the results presented in Figs. 7 and 8, in which the effects of the interannually varying LAI on the variability of  $E$  and  $SM$  are presented.

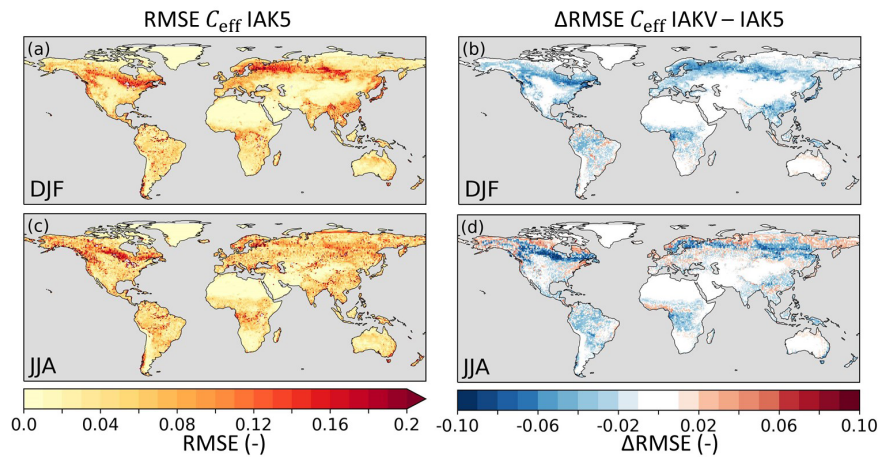
The opposing effects of the enhanced LAI variability on anomaly  $E_t$  and  $SM_{sb}$  can be explained by a negative feedback between vegetation and soil moisture schematized on the right side in Fig. 17a. During dry periods, the soil moisture is reduced; this lower soil water availability can result in vegetation water stress, consequently leading to lower vegetation activity in terms of transpiration and primary production, which is reflected, for example, in the typical dry season browning of grass species in low-vegetation regions and in the model represented by negative LAI and  $C_{eff}$  anomalies (Fig. 16a, b). As transpiration is reduced (Fig. 16c), the negative subsurface soil moisture anomaly is similarly reduced

because less water is extracted (Fig. 16d). On the other hand, the enhanced vegetation variability activated a positive feedback between anomaly vegetation activity and anomaly  $SM_s$ , as illustrated on the left side of Fig. 17a. Reduced vegetation activity is reflected in the model by a reduced  $C_{eff}$  and an increased bare soil fraction (Sect. 7), which leads to an increased  $E_s$  (Fig. 16b) and, as a consequence, less  $SM_s$  during a dry period as long as soil moisture is available (Fig. 16d).

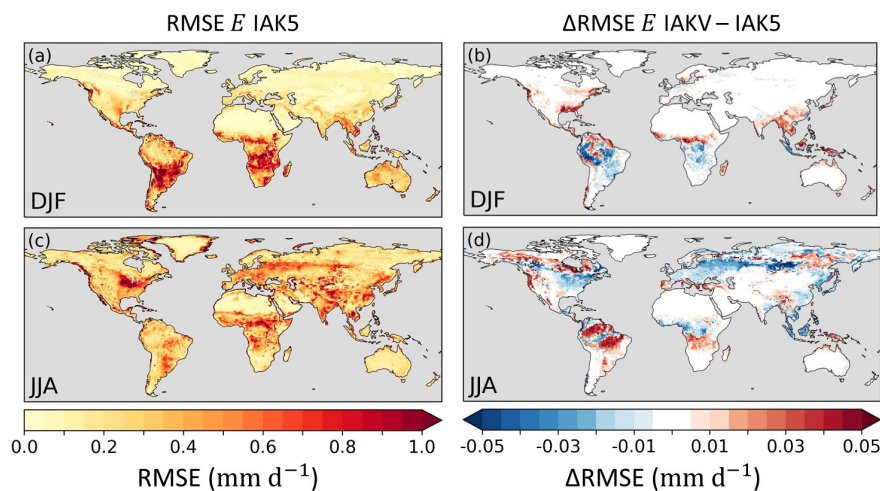
Figure 17b and c show that the positive feedback between  $E_s$  and  $SM_s$ , as introduced by the improved vegetation variability, is the strongest over semiarid regions with low vegetation, while the negative feedback between  $E_t$  and  $SM_{sb}$  is more pronounced for temperate regions with deciduous vegetation and crops, where the interannual LAI variability is larger (Fig. 4).



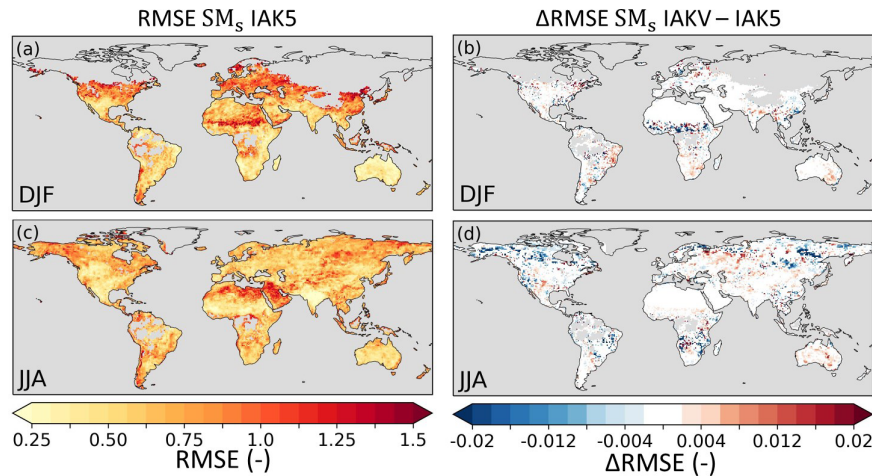
**Figure 11.** (a) Mean monthly model effective vegetation cover ( $C_{\text{eff}}$ ) in experiment IAK5 and (b) the absolute difference between IAKV and IAK5 (IAKV–IAK5) mean monthly  $C_{\text{eff}}$ . Red (blue) indicates a reduced (increased)  $C_{\text{eff}}$  in IAKV compared to IAK5. Details of model experiments are in Table 1.



**Figure 12.** Root mean squared error (RMSE) of model seasonal  $C_{\text{eff}}$  in experiment IAK5 with respect to CGLS FCover for DJF (a) and JJA (c), with red indicating a larger RMSE. The difference between RMSE in IAK5 and IAKV (IAKV–IAK5) for DJF (b) and JJA (d) with blue (red) indicating a reduced (increased) RMSE. White indicates small and/or insignificant  $\Delta\text{RMSE}$ . See Table 1 for details of the experiments. Similar figures for monthly values and all the seasons are presented in Figs. S8–S9.



**Figure 13.** Root mean squared error (RMSE) of model seasonal  $E$  in experiment IAK5 with respect to DOLCEv3  $E$  for DJF (a) and JJA (c), with red indicating a larger RMSE. The difference between RMSE in IAK5 and IAKV (IAKV–IAK5) for DJF (b) and JJA (d) with blue (red) indicating a reduced (increased) RMSE. White indicates small and/or insignificant  $\Delta\text{RMSE}$ . See Table 1 for details of the experiments. Similar figures for monthly values and all the seasons are presented in Figs. S10–S11.



**Figure 14.** Root mean squared error (RMSE) of model standardized interannual seasonal anomaly  $SM_s$  in experiment IAK5 with respect to ESA-CCI SM for DJF (a) and JJA (c). The difference between RMSE in IAK5 and IAKV (IAKV–IAK5) for DJF (b) and JJA (d) with blue (red) indicating a reduced (increased) RMSE. White indicates small and/or insignificant  $\Delta RMSE$ , and grey indicates no data points. See Table 1 for details of the experiments. Similar figures for monthly values and all the seasons are presented in Figs. S12–S13.

## 4 Discussion

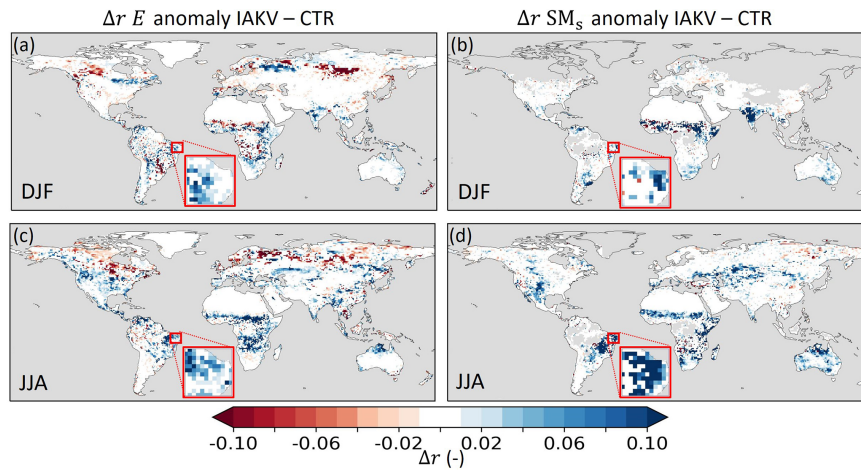
### 4.1 Synthesis and implications

The results presented in Sect. 3.4 indicate overall improvements of correlation coefficients of  $E$  and  $SM_s$  with all three aspects of vegetation variability implemented. We attribute these effects primarily to the implementation of interannually varying LAI, as the effects of the LC variability and vegetation-specific  $C_{\text{eff}}$  on  $E$  and  $SM_s$  were smaller (Sect. 3.1 and 3.3). The pronounced improvements in  $SM_s$  and  $E$  correlation in semiarid regions (Fig. 15) are directly related to the feedback mechanisms typical of water-limited regions that were activated by the vegetation variability. Regions where the positive feedback is strong (Fig. 17b) coincide with the regions that exhibit a strengthening of the correlations. In the model setup with seasonally varying LAI only (experiments CTR and IALC), the feedbacks in Fig. 17 are not represented because the interaction between SM and LAI is activated by the interannually varying LAI. In particular, the interactions between LAI,  $C_{\text{eff}}$ , and bare soil cover are only captured if model  $C_{\text{eff}}$  is exponentially related to LAI (Sect. 2.2.3). This finding complements the arguments from previous studies for using the exponential LAI– $C_{\text{eff}}$  relation instead of the lookup-table  $C_{\text{eff}}$  in HTESSEL (Alessandri et al., 2017; Johannsen et al., 2019; Nogueira et al., 2020, 2021).

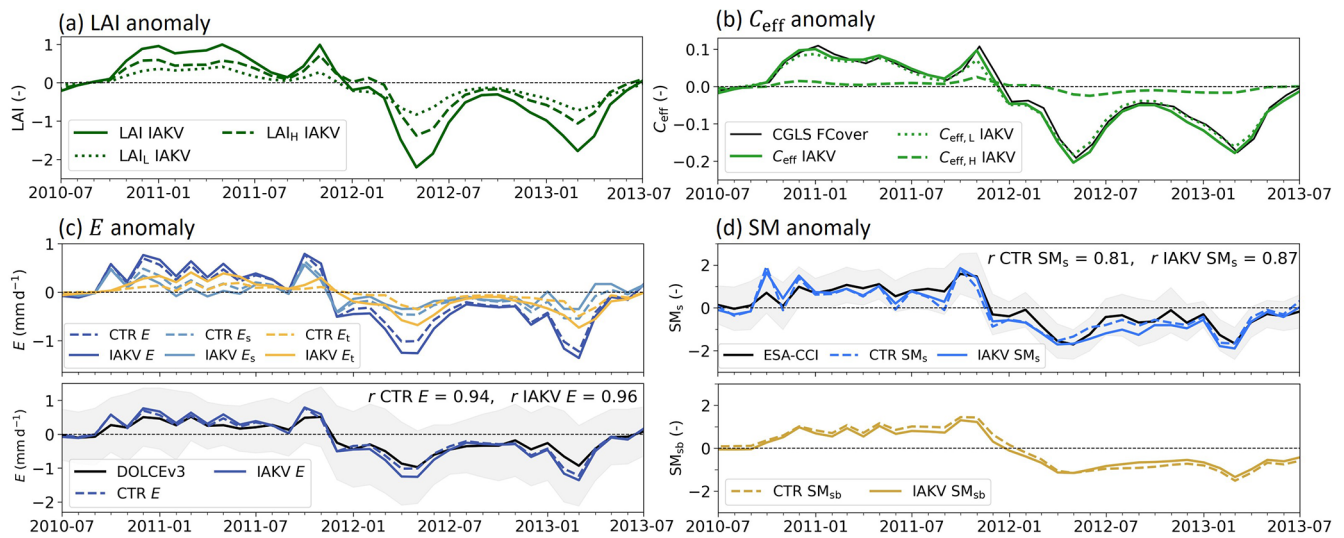
Recent studies also applied data assimilation methods to integrate satellite-based LAI in LSMs. For example, Rahman et al. (2022) found improved anomaly correlations of transpiration in many areas when integrating satellite-based LAI in the LSM called Noah-MP (Noah Multi-Parameterization), with the largest effects in the regions where  $E$  and  $SM_s$  anomaly correlations consistently improved in our results

(Fig. 9). However, this study also found limited sensitivity of model surface and root zone soil moisture when only LAI assimilation was applied (Rahman et al., 2022). Similarly, Albergel et al. (2017) concluded that LAI assimilation only affected deeper SM. In contrast, our results showed considerable changes in near-surface soil moisture when integrating CGLS LAI; this can be explained by the interplay between LAI, effective vegetation cover, soil evaporation, and near-surface soil moisture schematized in Fig. 17, which apparently differs from the interplay in Noah-MP (Rahman et al., 2022) and ISBA (Albergel et al., 2017).

The vegetation-specific effective vegetation cover parameterization presented in Sect. 3.3 generally resulted in an improved match of model  $C_{\text{eff}}$  and CGLS FCover (Fig. 12), which was expected because the FCover data were used for the estimation of the exponential coefficient  $k$  based on least-squares minimization (Sect. 2.3.3). CGLS FCover explicitly represents the fraction of green vegetation cover and therefore matches the model actively transpiring vegetation fraction  $C_{\text{eff}}$ . However, the non-green vegetated area cover also affects the atmosphere by, e.g., modifying albedo and roughness lengths, which is not considered in the model, as non-green vegetation is represented as bare soil. This is a limitation for the present implementation of the vegetation-specific effective vegetation cover parameterization. The results presented in Fig. 13 showed both increased and reduced RMSE for  $E$  with respect to the reference data in IAKV compared to IAK5. Consistent reductions of  $E$  RMSE in Europe and the eastern US in the JJA period were found. These regions coincide with regions with a high density of FLUXNET tower observations used for generation of the DOLCEv3  $E$  (Hobechi et al., 2021). The lack of tower observations in the tropics, in the Sahel, in southeastern Asia, and at high latitudes



**Figure 15.** Pearson correlation coefficient difference ( $\Delta r$ ) between experiment IAK5 and IAKV (IAKV–IAK5) for (a, c) seasonal anomaly total evaporation ( $E$ ) with respect to DOLCEv3 evaporation for DJF and JJA and (b, d) seasonal anomaly near-surface soil moisture ( $SM_s$ ) with respect to ESA-CCI SM for DJF and JJA. Blue (red) indicates an increased (reduced) correlation in IAKV compared to IAK5, white indicates small and/or insignificant  $\Delta r$ , and grey indicates no data points. The red box is highlighted in Fig. 16. See Table 1 for details of the experiments. Similar figures for all the seasons and monthly anomalies are presented in Figs. S17–S19.

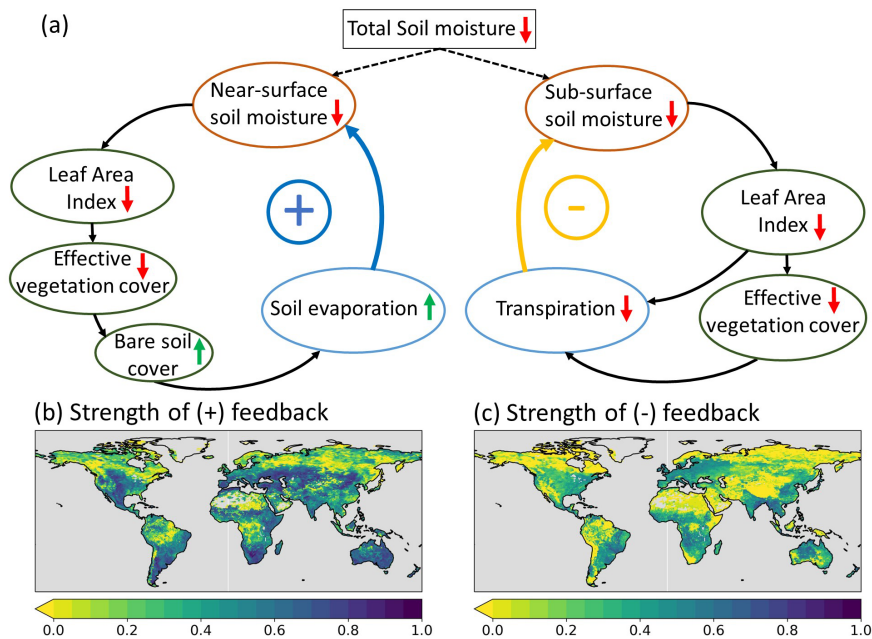


**Figure 16.** Time series of the northeastern Brazil case highlighted in Fig. 15 for (a) LAI anomalies. (b) Effective vegetation cover ( $C_{eff}$ ) anomalies with the CGLS FCover data in black as a reference. (c) Evaporation anomalies with  $E$  total evaporation,  $E_t$  being transpiration,  $E_s$  soil evaporation, and  $E_i$  interception evaporation; DOLCEv3  $E$  is in black as a reference. (d) Soil moisture standardized anomalies with  $SM_s$  near-surface soil moisture and  $SM_{sb}$  subsurface soil moisture; ESA-CCI SM is in black as a reference. Dashed lines in (c) and (d) represent experiment CTR and solid lines IAKV. The shading in (c) and (d) represents the uncertainty associated with the reference data. For this case  $T_L$  is short grass,  $T_H$  is deciduous broadleaf trees,  $A_L = 0.84$ , and  $A_H = 0.16$ .

may potentially explain the mixed RMSE results in these regions presented in Fig. 13. For high latitudes (e.g., northern Canada and eastern Siberia) the RMSE for both  $E$  as  $C_{eff}$  increased and the Pearson correlation was reduced (Fig. S14) in IAKV compared to IAK5 for the JJA period. This might be at least in part related to the poor fit of the parameterization for high LAI values for short grass and tundra, as explained in Sect. 3.3 (Fig. 10).

The interannually varying land cover locally affected the model  $E$  and SM as expected, with reduced (increased)  $E$  driven by corresponding reductions (increases) in high-vegetation cover fraction (Figs. 5 and 6). However, the effects on  $E$  and SM are likely underestimated due to the HTESSEL land cover structure in which the dominant vegetation type and cover fraction are used and vegetation mixing within high- or low-vegetation types is not represented





**Figure 17.** (a) Processes contributing to the anomaly vegetation–soil moisture feedback mechanisms as activated with the improved vegetation variability in IAKV compared to CTR. Upward (downward) arrows indicate positive (negative) change in the involved variables. Positive (blue) arrows indicate positive feedback and negative (yellow) arrows indicate negative feedback. The  $\pm$  symbols refer to the resulting positive or negative feedback loop relative to the sign of the change in the involved variables. The strength of the feedbacks (b, c) is quantified as the absolute correlation between anomaly  $\Delta E_s$  and  $\Delta SM_s$  (b) and between  $\Delta E_t$  and  $\Delta SM_{sb}$  (c), with  $\Delta$  representing the difference between anomaly CTR and IAKV (IAKV–CTR).

(Fig. 1). With this, only major changes in the ESA–CCI vegetation types and fractions are captured by the model. In IALC we evaluated the effects of interannually varying LC individually, but for internal consistency LAI and LC interannual variations should ideally be used together as they are interdependent. The local effects of the interannually varying land cover on the total  $E$  were considerably smaller than on the individual  $E$  fluxes (Fig. 5). The reduced (increased)  $E_t$  and  $E_i$  were compensated for by increased (reduced)  $E_s$ . This compensation is related to the  $C_{\text{eff}}$  parameterization (Eq. 6) and also to the offline setup, which does not allow for couplings with the atmosphere. Reduced  $A_H$  in the Amazon (Fig. 3) led to a reduced  $C_{\text{eff}}$  and an increased bare soil fraction (Sects. 4–7) and therefore reduced  $E_t$  and  $E_i$  as well as increased  $E_s$  in order to fulfill the atmospheric evaporation demand defined by the prescribed atmospheric forcing. Similarly, the on average reduced  $C_{\text{eff}}$  with the vegetation-specific  $C_{\text{eff}}$  parameterization (Fig. 11) introduced in experiment IAKV led to a shift in annual mean individual  $E$  fluxes, with increased  $E_s$  and reduced  $E_t$  and  $E_i$  (Fig. S6).

It is important to note that the partitioning of evaporation into the three individual components  $E_t$ ,  $E_s$ , and  $E_i$  in the model remains problematic to compare with observations. There is widespread consensus that, globally averaged, transpiration is the largest land evaporation flux component, followed by soil evaporation and interception evaporation (Mi-

ralles et al., 2011; Wei et al., 2017; Nelson et al., 2020). However, estimates of the average global  $E_t$  contribution to total terrestrial evaporation are subject to major uncertainties, with the global  $E_t$  contribution quantified in the range of 35%–80% (Schlesinger and Jasechko, 2014; Coenders-Gerrits et al., 2014). The global mean modeled partitioning of evaporation in our study is on the low end of these estimates with 39%  $E_t$ , 38%  $E_s$ , and 20%  $E_i$  in CTR and 38%  $E_t$ , 41%  $E_s$ , and 20%  $E_i$  in IAKV (the values do not add to 100% due to open-water evaporation). Despite the consistent improvements in anomaly correlation coefficients of  $E$  and  $SM_s$  found in IAKV compared to CTR (Fig. 15), the apparently low contribution of  $E_t$  to total  $E$  needs further evaluation, which was out of scope in this study.

#### 4.2 Methodological limitations

Our model experiments were performed in an offline mode with prescribed atmospheric forcing, which allowed us to analyze individual hydrological processes in detail. However, the fixed atmospheric model input considerably constrains changes in model surface fluxes. Moreover, the ERA5 forcing used here is based on an LSM that does not represent land cover and vegetation variability, which is partially corrected for by data assimilation of observations (Hersbach et al., 2020; Nogueira et al., 2021). The potential mismatch between our LSM and the ERA5 atmospheric forcing may

have also influenced the observed model effects. Another possible limitation is the absence of recalibration of model parameters, such as roughness lengths and minimum stomatal resistances. Fixed model parameters were originally calibrated using the lookup-table  $C_{\text{eff}}$  parameterization, MODIS LAI, and GLCC LC, and they have not been adjusted for the three new model scenarios tested here (IALC, IAK5, and IAKV). This was also emphasized by Johannsen et al. (2019), Nogueira et al. (2020, 2021), and Boussetta et al. (2021), who concluded that model vegetation changes should be implemented in an integral context and recalibration of model parameters is needed.

This study emphasizes the importance of realistic representation of vegetation variability for modeling land surface–atmosphere interactions. However, for further applications exploring how the vegetation variability influences atmospheric variables in a coupled model setup is needed. The availability of reliable reference data is therefore fundamental to properly understand and model the processes of relevance for the land surface and interaction with the atmosphere. Here, the evaluation of model performance was limited to total evaporation and near-surface soil moisture. The evaluated performances of model  $E$  and  $SM_s$  need to be interpreted in a careful way, bearing in mind the uncertainties. For total evaporation we used the DOLCEv3 evaporation data that merge FLUXNET tower observations with evaporation from FLUXCOM-RSMETEO, GLEAM v3.5a and v3.5b, and ERA5-land, which all include very specific model assumptions on vegetation representations. Although these data are considered suitable for time series and trend analyses, the associated uncertainty estimates are large (Hobeichi et al., 2021) (Fig. 16). Figure 16c shows that the DOLCEv3 interannual variability is systematically smaller than the modeled variability. This limited interannual variability in DOLCEv3 could be at least in part related to the combination of several products because the averaging based on FLUXNET towers unavoidably dampens the anomalies, reducing the interannual variability. Evaluation of the modeled near-surface soil moisture was limited by missing data due to dense forests or snow cover and the lack of information on the representative soil depth. While the ESA-CCI combined active–passive SM product was generated using the absolute values and the dynamic range of GLDAS-Noah soil moisture, preserving the dynamics and trends of the original retrievals (Liu et al., 2012), it is important to note that during dry-downs the soil moisture dynamics can also be impacted to some extent, as highlighted by Raoult et al. (2022). However, we still find the ESA-CCI SM to be the best-suited globally available reference data for our study because they represent a direct product of remote sensing observations, without directly blending land surface model dynamics as done for DOLCEv3.

## 5 Conclusions

This study aimed to address the limitations of state-of-the-art land surface models in representing spatial and temporal vegetation dynamics. We evaluated the effects of improving the representation of land cover and vegetation variability based on satellite observational products in the HTESSEL land surface model. Specifically, we directly integrated satellite-based interannually varying land cover and seasonally and interannually varying LAI. In addition, we formulated and integrated an effective vegetation cover parameterization that can distinguish between different vegetation types. The effects of these three implementations were analyzed for soil moisture and evaporation in offline experiments forced with ERA5 atmospheric forcing.

The interannually varying land cover locally altered the model evaporation and soil moisture. In regions with major land cover changes, such as the Amazon, the model evaporation fluxes and soil moisture responded consistently, capturing the effects of increased or decreased high or low vegetation cover. The interannually varying LAI led to significant improvements of the correlation coefficients computed with the available reference data on near-surface soil moisture and evaporation. This was specifically true in semiarid regions with predominantly low vegetation during the dry season. The interannually varying LAI and effective vegetation cover allow for an adequate representation of soil moisture–evaporation feedback by activating the couplings with vegetation during vegetation–water–stressed periods (Fig. 17). From these results, we conclude that it is essential to realistically represent interannual variability of LAI and to include the exponential relation between LAI and effective vegetation cover to correctly capture land–atmospheric feedbacks during droughts in HTESSEL. The developments of the effective vegetation cover parameterization considerably improved the spatial and temporal variability of the model effective vegetation cover and regionally reduced the model errors of evaporation and near-surface soil moisture. Overall, our results emphasize the need to represent spatial and temporal vegetation variability in LSMs used for climate re-analyses and near-term climate predictions. In climate predictions, we obviously cannot rely on satellite retrievals, and therefore the development and validation of dynamical or statistical models able to reliably predict vegetation dynamics, from leaf to ecosystem scales, remain an important challenge for the future in the land surface modeling community.

**Code and data availability.** ESA-CCI land cover data were taken from <https://doi.org/10.24381/cds.006f2c9a> (Copernicus Climate Change Service, 2019). CGLS LAI and FCover data were downloaded from <https://land.copernicus.eu/global/products/> (last access: January 2022) (Copernicus Global Land Service, 2022). AVHRR LAI was accessed from <https://www.ncei.noaa.gov/access/metadata/landing-page/bin/iso?id=gov.noaa.ncdc:C01559> (last ac-

cess: June 2021). The preparation of the ESA-CCI land cover, CGLS LAI, and AVHRR LAI for use in HTESSEL is explained in Boussetta and Balsamo (2021). DOLCEv3 was accessed through <https://doi.org/10.25914/606e9120c5ebe> (last access: May 2022) (Hobeichi et al., 2021) and ESA-CCI SM v06.1 through <https://esa-soilmoisture-cci.org/data> (last access: May 2022). The offline HTESSEL model was provided by EC-EARTH, together with the ERA5 forcing data as well as vegetation and soil data. The scripts underlying this <https://doi.org/10.5281/zenodo.8254556> (van Oorschot, 2023b). Data underlying this publication are available at <https://doi.org/10.5281/zenodo.8307861> (van Oorschot, 2023a).

**Supplement.** The supplement related to this article is available online at: <https://doi.org/10.5194/esd-14-1239-2023-supplement>.

**Author contributions.** The study was conceived by AA and amended with input from all co-authors. FvO carried out the study and prepared the original draft with supervision of AA, RJvdE, and MH. SB and GB prepared some of the data used for modeling. All authors contributed to review and editing the final draft.

**Competing interests.** The contact author has declared that none of the authors has any competing interests.

**Disclaimer.** Publisher's note: Copernicus Publications remains neutral with regard to jurisdictional claims made in the text, published maps, institutional affiliations, or any other geographical representation in this paper. While Copernicus Publications makes every effort to include appropriate place names, the final responsibility lies with the authors.

**Acknowledgements.** Acknowledgement is given for the use of ECMWF's computing and archive facilities for this research, which were provided by CNR-ISAC and by ECMWF in the framework of the special project SPITALEs.

**Financial support.** This work was supported by the European Union's Horizon 2020 research and innovation program under grant agreement no. 101004156 (CONFESS project) and by ICSC – Centro Nazionale di Ricerca in High-Performance Computing, Big Data and Quantum Computing, funded by the European Union – NextGenerationEU (project code: CN00000013; CUP: B93C22000620006). Ruud J. van der Ent received funding from the Netherlands Organization for Scientific Research (NWO) under project number 016.Veni.181.015.

**Review statement.** This paper was edited by Anping Chen and reviewed by two anonymous referees.

## References

- Albergel, C., Munier, S., Leroux, D. J., Dewaele, H., Fairbairn, D., Barbu, A. L., Gelati, E., Dorigo, W., Faroux, S., Meurey, C., Le Moigne, P., Decharme, B., Mahfouf, J.-F., and Calvet, J.-C.: Sequential assimilation of satellite-derived vegetation and soil moisture products using SURFEX\_v8.0: LDAS-Monde assessment over the Euro-Mediterranean area, *Geosci. Model Dev.*, 10, 3889–3912, <https://doi.org/10.5194/gmd-10-3889-2017>, 2017.
- Albergel, C., Munier, S., Bocher, A., Bonan, B., Zheng, Y., Draper, C., Leroux, D. J., and Calvet, J.-C.: LDAS-Monde Sequential Assimilation of Satellite Derived Observations Applied to the Contiguous US: An ERA-5 Driven Reanalysis of the Land Surface Variables, *Remote Sens.*, 10, 2072–4292, <https://doi.org/10.3390/rs10101627>, 2018.
- Alessandri, A., Gualdi, S., Polcher, J., and Navarra, A.: Effects of land surface–vegetation on the boreal summer surface climate of a GCM, *J. Climate*, 20, 255–278, <https://doi.org/10.1175/JCLI3983.1>, 2007.
- Alessandri, A., Catalano, F., De Felice, M., Van Den Hurk, B., Doblas Reyes, F., Boussetta, S., Balsamo, G., and Miller, P. A.: Multi-scale enhancement of climate prediction over land by increasing the model sensitivity to vegetation variability in EC-Earth, *Clim. Dynam.*, 49, 1215–1237, <https://doi.org/10.1007/s00382-016-3372-4>, 2017.
- Anderson, M. C., Norman, J. M., Kustas, W. P., Li, F., Prueger, J. H., and Mecikalski, J. R.: Effects of vegetation clumping on two-source model estimates of surface energy fluxes from an agricultural landscape during SMACEX, *J. Hydrometeorol.*, 6, 892–909, <https://doi.org/10.1175/JHM465.1>, 2005.
- Balsamo, G., Viterbo, P., Beijaars, A., van den Hurk, B., Hirschi, M., Betts, A. K., and Scipal, K.: A revised hydrology for the ECMWF model: Verification from field site to terrestrial water storage and impact in the integrated forecast system, *J. Hydrometeorol.*, 10, 623–643, <https://doi.org/10.1175/2008JHM1068.1>, 2009.
- Balsamo, G., Agusti-Panareda, A., Albergel, C., Arduini, G., Beljaars, A., Bidlot, J., Bousserez, N., Boussetta, S., Brown, A., Buizza, R., Buontempo, C., Chevallier, F., Choulga, M., Cloke, H., Cronin, M. F., Dahoui, M., Rosnay, P. D., Dirmeyer, P. A., Drusch, M., Dutra, E., Ek, M. B., Gentile, P., Hewitt, H., Keeley, S. P., Kerr, Y., Kumar, S., Lupu, C., Mahfouf, J. F., McNorton, J., Mecklenburg, S., Mogensen, K., Muñoz-Sabater, J., Orth, R., Rabier, F., Reichle, R., Ruston, B., Pappenberger, F., Sandu, I., Seneviratne, S. I., Tietsche, S., Trigo, I. F., Uijlenhoet, R., Wedi, N., Woolway, R. I., and Zeng, X.: Satellite and in situ observations for advancing global earth surface modelling: A review, *Remote Sens.*, 10, 1–72, <https://doi.org/10.3390/rs10122038>, 2018.
- Boussetta, S. and Balsamo, G.: Vegetation dataset of Land Use/Land Cover and Leaf Area Index Land Use/Land Cover and Leaf Area Index, *Tech. Rep.*, <https://confess-h2020.eu/results/deliverables/> (last access: January 2022), 2021.
- Boussetta, S., Balsamo, G., Beljaars, A., Kral, T., and Jarlan, L.: Impact of a satellite-derived leaf area index monthly climatology in a global numerical weather prediction model, *Int. J. Remote Sens.*, 34, 3520–3542, <https://doi.org/10.1080/01431161.2012.716543>, 2013.
- Boussetta, S., Balsamo, G., Arduini, G., Dutra, E., Choulga, M., Agustí-panareda, A., Beljaars, A., Wedi, N., Muñoz-sabater,

- J., Rosnay, P. D., Sandu, I., Hadade, I., Mazzetti, C., Prudhomme, C., Yamazaki, D., and Zsoter, E.: ECLand: The ECMWF Land surface modelling system, *Atmosphere*, 12, 723, <https://doi.org/10.3390/atmos12060723>, 2021.
- Chen, B., Lu, X., Wang, S., Chen, J. M., Liu, Y., Fang, H., Liu, Z., Jiang, F., Arain, M. A., Chen, J., and Wang, X.: Evaluation of Clumping Effects on the Estimation of Global Terrestrial Evapotranspiration, *Remote Sens.*, 13, 1–24, 2021.
- Chen, J. M., Menges, C. H., and Leblanc, S. G.: Global mapping of foliage clumping index using multi-angular satellite data, *Remote Sens. Environ.*, 97, 447–457, <https://doi.org/10.1016/j.rse.2005.05.003>, 2005.
- Coenders-Gerrits, A. M., Van Der Ent, R. J., Bogaard, T. A., Wang-Erlandsson, L., Hrachowitz, M., and Savenije, H. H.: Uncertainties in transpiration estimates, *Nature*, 506, 2013–2015, <https://doi.org/10.1038/nature12925>, 2014.
- Copernicus Climate Change Service: Land cover classification gridded maps from 1992 to present derived from satellite observations, [data set], <https://doi.org/10.24381/cds.006f2c9a>, 2019.
- Copernicus Global Land Service: Fraction of green vegetation Cover and Leaf Area Index, [data set], <https://land.copernicus.eu/global/products/> (last access: January 2022), 2022.
- Defourny, P., Lamarche, C., Bontemps, S., de Maet, T., Van Bogaert, E., Moreau, I., Brockmann, C., Boettcher, M., Kirches, G., Wevers, J., and Santoro, M.: Land Cover CCI Product User Guide, Tech. Rep., European Space Agency, [https://maps.elie.ucl.ac.be/CCI/viewer/download/ESACCI-LC-Ph2-PUGv2\\_2.0.pdf](https://maps.elie.ucl.ac.be/CCI/viewer/download/ESACCI-LC-Ph2-PUGv2_2.0.pdf) (last access: January 2022), 2017.
- Di Gregorio, A. and Jansen, L. J. M.: Land cover classification system: LCCS: Classification concepts and user manual, Food & Agriculture Organization of the United Nations, ISBN 95-5-105327-8, 2005.
- Dong, J., Crow, W. T., Tobin, K. J., Cosh, M. H., Bosch, D. D., Starks, P. J., Seyfried, M., and Collins, C. H.: Comparison of microwave remote sensing and land surface modeling for surface soil moisture climatology estimation, *Remote Sens. Environ.*, 242, 111756, <https://doi.org/10.1016/j.rse.2020.111756>, 2020.
- Dorigo, W., Wagner, W., Albergel, C., Albrecht, F., Balsamo, G., Brocca, L., Chung, D., Ertl, M., Forkel, M., Gruber, A., Haas, E., Hamer, P. D., Hirschi, M., Ikonen, J., de Jeu, R., Kidd, R., Lahoz, W., Liu, Y. Y., Miralles, D., Mistelbauer, T., Nicolai-Shaw, N., Parinussa, R., Pratola, C., Reimer, C., van der Schalie, R., Seneviratne, S. I., Smolander, T., and Lecomte, P.: ESA CCI Soil Moisture for improved Earth system understanding: State-of-the-art and future directions, *Remote Sens. Environ.*, 203, 185–215, <https://doi.org/10.1016/j.rse.2017.07.001>, 2017.
- Döscher, R., Acosta, M., Alessandri, A., Anthoni, P., Arsouze, T., Bergman, T., Bernardello, R., Boussetta, S., Caron, L.-P., Carver, G., Castrillo, M., Catalano, F., Cvijanovic, I., Davini, P., Dekker, E., Doblaz-Reyes, F. J., Docquier, D., Echevarria, P., Fladrich, U., Fuentes-Franco, R., Gröger, M., v. Hardenberg, J., Hieronymus, J., Karami, M. P., Keskinen, J.-P., Koenigk, T., Makkonen, R., Massonnet, F., Ménégoz, M., Miller, P. A., Moreno-Chamarro, E., Nieradzki, L., van Noije, T., Nolan, P., O'Donnell, D., Olinaho, P., van den Oord, G., Ortega, P., Prims, O. T., Ramos, A., Reerink, T., Rousset, C., Ruprich-Robert, Y., Le Sager, P., Schmith, T., Schrödner, R., Serva, F., Sicardi, V., Sloth Madsen, M., Smith, B., Tian, T., Tourigny, E., Uotila, P., Vancoppenolle, M., Wang, S., Wårlind, D., Willén, U., Wyser, K., Yang, S., Yepes-Arbós, X., and Zhang, Q.: The EC-Earth3 Earth system model for the Coupled Model Intercomparison Project 6, *Geosci. Model Dev.*, 15, 2973–3020, <https://doi.org/10.5194/gmd-15-2973-2022>, 2022.
- ECMWF: Part IV: Physical Processes, IFS Documentation – Cy43r1, 1–223, <https://www.ecmwf.int/sites/default/files/elibrary/2016/17117-part-iv-physical-processes.pdf> (last access: June 2022), 2015.
- European Space Agency: Soil Moisture CCI Data v07.1, <https://www.esa-soilmoisture-cci.org/data> (last access: May 2022), 2022.
- Faroux, S., Kaptué Tchuenté, A. T., Roujean, J.-L., Masson, V., Martin, E., and Le Moigne, P.: ECOCLIMAP-II/Europe: a twofold database of ecosystems and surface parameters at 1 km resolution based on satellite information for use in land surface, meteorological and climate models, *Geosci. Model Dev.*, 6, 563–582, <https://doi.org/10.5194/gmd-6-563-2013>, 2013.
- Fisher, R. A. and Koven, C. D.: Perspectives on the Future of Land Surface Models and the Challenges of Representing Complex Terrestrial Systems, *J. Adv. Model. Earth Sy.*, 12, e2018MS001453, <https://doi.org/10.1029/2018MS001453>, 2020.
- Gruber, A., Dorigo, W. A., Crow, W., and Wagner, W.: Triple Collocation-Based Merging of Satellite Soil Moisture Retrievals, *IEEE T. Geosci. Remote*, 55, 6780–6792, <https://doi.org/10.1109/TGRS.2017.2734070>, 2017.
- Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., De Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Fleming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R. J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., de Rosnay, P., Rozum, I., Vamborg, F., Villaume, S., and Thépaut, J. N.: The ERA5 global reanalysis, *Q. J. Roy. Meteor. Soc.*, 146, 1999–2049, <https://doi.org/10.1002/qj.3803>, 2020.
- Hobeichi, S., Abramowitz, G., Evans, J., and Ukkola, A.: Derived Optimal Linear Combination Evapotranspiration (DOLCE): a global gridded synthesis ET estimate, *Hydrol. Earth Syst. Sci.*, 22, 1317–1336, <https://doi.org/10.5194/hess-22-1317-2018>, 2018.
- Hobeichi, S., Abramowitz, G., and Evans, J. P.: Robust historical evapotranspiration trends across climate regimes, *Hydrol. Earth Syst. Sci.*, 25, 3855–3874, <https://doi.org/10.5194/hess-25-3855-2021>, 2021.
- Johannsen, F., Ermida, S., Martins, J. P., Trigo, I. F., Nogueira, M., and Dutra, E.: Cold bias of ERA5 summertime daily maximum land surface temperature over Iberian Peninsula, *Remote Sens.*, 11, 2570, <https://doi.org/10.3390/rs11212570>, 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 Biogeochem. Cy.*, 19, 1–33, <https://doi.org/10.1029/2003GB002199>, 2005.
- Kumar, S. V., Mocko, D. M., Wang, S., Peters-Lidard, C. D., and Borak, J.: Assimilation of remotely sensed leaf area index into the noah-mp land surface model: Impacts on water and car-

- bon fluxes and states over the continental United States, *J. Hydrometeorol.*, 20, 1359–1377, <https://doi.org/10.1175/JHM-D-18-0237.1>, 2019.
- Le Moigne, P.: SURFEX scientific documentation, Groupe de météorologie a moyenne échelle, p. 237, [http://www.cnrm.meteo.fr/surfex/IMG/pdf/surfex\\_scidoc\\_v2.pdf](http://www.cnrm.meteo.fr/surfex/IMG/pdf/surfex_scidoc_v2.pdf) (last access: October 2022), 2012.
- Ling, X. L., Fu, C. B., Guo, W. D., and Yang, Z. L.: Assimilation of Remotely Sensed LAI Into CLM4CN Using DART, *J. Adv. Model. Earth Sy.*, 11, 2768–2786, <https://doi.org/10.1029/2019MS001634>, 2019.
- Liu, Y., Dorigo, W., Parinussa, R., de Jeu, R., Wagner, W., McCabe, M., Evans, J., and van Dijk, A.: 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.
- Lo, M. H., Famiglietti, J. S., Yeh, P. J., and Syed, T. H.: Improving parameter estimation and water table depth simulation in a land surface model using GRACE water storage and estimated base flow data, *Water Resour. Res.*, 46, 1–15, <https://doi.org/10.1029/2009WR007855>, 2010.
- Loveland, T. R., Reed, B. C., Ohlen, D. O., Brown, J. F., Zhu, Z., Yang, L., and Merchant, J. W.: Development of a global land cover characteristics database and IGBP DISCover from 1 km AVHRR data, *Int. J. Remote Sens.*, 21, 1303–1330, <https://doi.org/10.1080/014311600210191>, 2000.
- MacBean, N., Maignan, F., Peylin, P., Bacour, C., Bréon, F.-M., and Ciais, P.: Using satellite data to improve the leaf phenology of a global terrestrial biosphere model, *Biogeosciences*, 12, 7185–7208, <https://doi.org/10.5194/bg-12-7185-2015>, 2015.
- Miralles, D. G., De Jeu, R. A. M., Gash, J. H., Holmes, T. R. H., and Dolman, A. J.: Magnitude and variability of land evaporation and its components at the global scale, *Hydrol. Earth Syst. Sci.*, 15, 967–981, <https://doi.org/10.5194/hess-15-967-2011>, 2011.
- Miralles, D. G., Brutsaert, W., Dolman, A. J., and Gash, J. H.: On the Use of the Term “Evapotranspiration”, *Water Resour. Res.*, 56, e2020WR028055, <https://doi.org/10.1029/2020WR028055>, 2020.
- Nelson, J. A., Pérez-Priego, O., Zhou, S., Poyatos, R., Zhang, Y., Blanken, P. D., Gimeno, T. E., Wohlfahrt, G., Desai, A. R., Gioli, B., Limousin, J. M., Bonal, D., Paul-Limoges, E., Scott, R. L., Varlagin, A., Fuchs, K., Montagnani, L., Wolf, S., Delpierre, N., Berveiller, D., Gharun, M., Beletti Marchesini, L., Gianelle, D., Šigut, L., Mammarella, I., Siebicke, L., Andrew Black, T., Knohl, A., Hörtnagl, L., Magliulo, V., Besnard, S., Weber, U., Carvalhais, N., Migliavacca, M., Reichstein, M., and Jung, M.: Ecosystem transpiration and evaporation: Insights from three water flux partitioning methods across FLUXNET sites, *Global Change Biol.*, 26, 6916–6930, <https://doi.org/10.1111/gcb.15314>, 2020.
- Nogueira, M., Albergel, C., Boussetta, S., Johannsen, F., Trigo, I. F., Ermida, S. L., Martins, J. P. A., and Dutra, E.: Role of vegetation in representing land surface temperature in the HTESSEL (CY45R1) and SURFEX-ISBA (v8.1) land surface models: a case study over Iberia, *Geosci. Model Dev.*, 13, 3975–3993, <https://doi.org/10.5194/gmd-13-3975-2020>, 2020.
- Nogueira, M., Boussetta, S., Balsamo, G., Albergel, C., Trigo, I. F., Johannsen, F., Miralles, D. G., and Dutra, E.: Upgrading Land-Cover and Vegetation Seasonality in the ECMWF Coupled System: Verification With FLUXNET Sites, ME-TEOSAT Satellite Land Surface Temperatures, and ERA5 Atmospheric Reanalysis, *J. Geophys. Res.-Atmos.*, 126, 1–26, <https://doi.org/10.1029/2020JD034163>, 2021.
- Orth, R., Dutra, E., Trigo, I. F., and Balsamo, G.: Advancing land surface model development with satellite-based Earth observations, *Hydrol. Earth Syst. Sci.*, 21, 2483–2495, <https://doi.org/10.5194/hess-21-2483-2017>, 2017.
- Pacholczyk, P. and Verger, A.: GEOV2 AVHRR Product User Manual, Tech. Rep., <https://www.theia-land.fr/wp-content/uploads/2020/11/THEIA-MU-44-0369-CNES-GEOV2-AVHRR-Product-User-Manual-V2.pdf> (last access: January 2022), 2020.
- Pitman, A. J., De Noblet-Ducoudré, N., Cruz, F. T., Davin, E. L., Bonan, G. B., Brovkin, V., Claussen, M., Delire, C., Ganzeveld, L., Gayler, V., Van Den Hurk, B. J., Lawrence, P. J., Van Der Molen, M. K., Müller, C., Reick, C. H., Seneviratne, S. I., Strengen, B. J., and Voldoire, A.: Uncertainties in climate responses to past land cover change: First results from the LUCID intercomparison study, *Geophys. Res. Lett.*, 36, 1–6, <https://doi.org/10.1029/2009GL039076>, 2009.
- Poulter, B., MacBean, N., Hartley, A., Khlystova, I., Arino, O., Betts, R., Bontemps, S., Boettcher, M., Brockmann, C., Defourny, P., Hagemann, S., Herold, M., Kirches, G., Lamarche, C., Lederer, D., Otlé, C., Peters, M., and Peylin, P.: Plant functional type classification for earth system models: results from the European Space Agency’s Land Cover Climate Change Initiative, *Geosci. Model Dev.*, 8, 2315–2328, <https://doi.org/10.5194/gmd-8-2315-2015>, 2015.
- Rahman, A., Zhang, X., Xue, Y., Houser, P., Sauer, T., Kumar, S., Mocko, D., and Maggioni, V.: A synthetic experiment to investigate the potential of assimilating LAI through direct insertion in a land surface model, *J. Hydrol.*, 9, 100063, <https://doi.org/10.1016/j.hydroa.2020.100063>, 2020.
- Rahman, A., Zhang, X., Houser, P., Sauer, T., and Maggioni, V.: Global Assimilation of Remotely Sensed Leaf Area Index: The Impact of Updating More State Variables Within a Land Surface Model, *Front. Water*, 3, 1–15, <https://doi.org/10.3389/frwa.2021.789352>, 2022.
- Raoult, N., Ruscica, R. C., Salvia, M. M., and Sörensson, A. A.: Soil Moisture Drydown Detection Is Hindered by Model-Based Rescaling, *IEEE Geosci. Remote Sens.*, 19, 1–5, <https://doi.org/10.1109/LGRS.2022.3178685>, 2022.
- Ryu, Y., Nilson, T., Kobayashi, H., Sonnentag, O., Law, B. E., and Baldocchi, D. D.: On the correct estimation of effective leaf area index: Does it reveal information on clumping effects?, *Agr. Forest Meteorol.*, 150, 463–472, <https://doi.org/10.1016/j.agrformet.2010.01.009>, 2010.
- Savenije, H. H.: The importance of interception and why we should delete the term evapotranspiration from our vocabulary, *Hydrol. Proc.*, 18, 1507–1511, <https://doi.org/10.1002/hyp.5563>, 2004.
- Schlesinger, W. H. and Jasechko, S.: Transpiration in the global water cycle, *Agr. Forest Meteorol.*, 189, 115–117, <https://doi.org/10.1016/j.agrformet.2014.01.011>, 2014.
- Schulzweida, U.: CDO User Guide, Max Planck Institute for Meteorology, <https://doi.org/10.5281/zenodo.3539275>, 2022.
- 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.
- Trigo, I. F., Boussetta, S., Viterbo, P., Balsamo, G., Beljaars, A., and Sandu, I.: Comparison of model land skin temperature with remotely sensed estimates and assessment of surface-atmosphere coupling, *J. Geophys. Res.-Atmos.*, 175, 238, <https://doi.org/10.1038/175238c0>, 2015.
- Ukkola, A. M., De Kauwe, M. G., Pitman, A. J., Best, M. J., Abramowitz, G., Haverd, V., Decker, M., and Haughton, N.: Land surface models systematically overestimate the intensity, duration and magnitude of seasonal-scale evaporative droughts, *Environ. Res. Lett.*, 11, 104012, <https://doi.org/10.1088/1748-9326/11/10/104012>, 2016.
- van Oorschot, F.: Dataset underlying van Oorschot et al. (2023), Zenodo [data set], <https://doi.org/10.5281/zenodo.8307861>, 2023a.
- van Oorschot, F.: Python scripts van Oorschot et al. (2023), Zenodo [code], <https://doi.org/10.5281/zenodo.8254556>, 2023b.
- van Oorschot, F., van der Ent, R. J., Hrachowitz, M., and Alessandri, A.: Climate-controlled root zone parameters show potential to improve water flux simulations by land surface models, *Earth Syst. Dynam.*, 12, 725–743, <https://doi.org/10.5194/esd-12-725-2021>, 2021.
- Verger, A., Baret, F., and Weiss, M.: Near real-time vegetation monitoring at global scale, *IEEE J. Sel. Top. Appl.*, 7, 3473–3481, <https://doi.org/10.1109/JSTARS.2014.2328632>, 2014.
- Verger, A., Baret, F., and Weiss, M.: Vegetation and Energy Algorithm Theoretical Basis Document, Copernicus Global Land Operations, 1–93, [https://land.copernicus.eu/global/sites/cgls.vito.be/files/products/CGLOPS1\\_ATBD\\_LAI1km-V2\\_I1.41.pdf](https://land.copernicus.eu/global/sites/cgls.vito.be/files/products/CGLOPS1_ATBD_LAI1km-V2_I1.41.pdf) (last access: January 2022), 2019.
- Wei, Z., Yoshimura, K., Wang, L., Miralles, D. G., Jasechko, S., and Lee, X.: Revisiting the contribution of transpiration to global terrestrial evapotranspiration, *Geophys. Res. Lett.*, 44, 2792–2801, <https://doi.org/10.1002/2016GL072235>, 2017.
- Yang, K., Zhu, L., Chen, Y., Zhao, L., Qin, J., Lu, H., Tang, W., Han, M., Ding, B., and Fang, N.: Land surface model calibration through microwave data assimilation for improving soil moisture simulations, *J. Hydrol.*, 533, 266–276, <https://doi.org/10.1016/j.jhydrol.2015.12.018>, 2016.
- Zhang, L., Hu, Z., Fan, J., Zhou, D., and Tang, F.: A meta-analysis of the canopy light extinction coefficient in terrestrial ecosystems, *Front. Earth Sci.*, 8, 599–609, <https://doi.org/10.1007/s11707-014-0446-7>, 2014.