Impact of Updating Vegetation Information on Land Surface Model Performance

Melissa Ruiz-Vásquez[1,2] , **Sungmin O**[3] , **Gabriele Arduini**[4] , **Souhail Boussetta**[4], **Alexander Brenning**[2] , **Ana Bastos**[1] , **Sujan Koirala**[1] , **Gianpaolo Balsamo**[4] , **Markus Reichstein**[1] , and **René Orth**[1]

[1]Department of Biogeochemical Integration, Max Planck Institute for Biogeochemistry, Jena, Germany, [2]Department of Geography, Friedrich Schiller University Jena, Jena, Germany, [3]Department of Climate and Energy System Engineering, Ewha Womans University, Seoul, South Korea, [4]Research Department, European Centre for Medium-Range Weather Forecasts, Reading, UK

Abstract Vegetation plays a fundamental role in modulating the exchange of water, energy, and carbon fluxes between the land and the atmosphere. These exchanges are modeled by Land Surface Models (LSMs), which are an essential part of numerical weather prediction and data assimilation. However, most current LSMs implemented specifically in weather forecasting systems use climatological vegetation indices, and land use/land cover data sets in these models are often outdated. In this study, we update land surface data in the European Centre for Medium-range Weather Forecast (ECMWF) land surface modeling system (ECLand) using Earth observation-based time varying leaf area index and land use/land cover data, and evaluate the impact of vegetation dynamics on model performance. The performance of the simulated latent heat flux and soil moisture is then evaluated against global gridded observation-based data sets. Updating the vegetation information does not always yield better model performances because the model's parameters are adapted to the previously employed land surface information. Therefore we recalibrate key soil and vegetation-related parameters at individual grid cells to adjust the model parameterizations to the new land surface information. This substantially improves model performance and demonstrates the benefits of updated vegetation information. Interestingly, we find that a regional parameter calibration outperforms a globally uniform adjustment of parameters, indicating that parameters should sufficiently reflect spatial variability in the land surface. Our results highlight that newly available Earth-observation products of vegetation dynamics and land cover changes can improve land surface model performances, which in turn can contribute to more accurate weather forecasts.

Plain Language Summary The accuracy of weather forecasts relies critically on the accurate modeling of the exchange of water and energy between the land surface and the atmosphere. The latent heat flux and the soil moisture are two important land surface variables in this exchange through the net balances of water and energy. The accurate simulation of these variables is challenging in most land surface models specifically used for numerical weather prediction due to (a) outdated land surface cover information and/or (b) neglecting the role of short-term anomalies in vegetation functioning, for example, related to droughts. This study quantifies the benefits of including up-to-date land use/land cover information and an explicit consideration of the current vegetation state on the prediction of latent heat flux and soil moisture. We find that model simulation performance can only benefit from updated land surface information through further adjustments to key soil and vegetation related parameters in the model. Overall, we demonstrate that the new Earth observation data sets can help to improve land surface model performance, which then contributes to more accurate weather forecasts.

1. Introduction

The atmosphere is sensitive to variations in land surface processes, and such covariability between the land and atmosphere states is described as the land-atmosphere coupling (Quillet et al., 2010; Santanello et al., 2009, 2018). The land surface characteristics, for example, vegetation state, albedo, and soil moisture, play important roles in this coupling as they modulate the exchange of water, energy, and carbon between the land surface and the atmosphere (Balsamo et al., 2011; de Rosnay et al., 2013; Dirmeyer et al., 2018). Accordingly, an adequate representation of land surface properties in the land surface models that are specifically used in numerical

Methodology: Melissa Ruiz-Vásquez, Sungmin O, Alexander Brenning, Ana Bastos, Gianpaolo Balsamo, Markus Reichstein, René Orth
Resources: Gianpaolo Balsamo
Software: Gabriele Arduini, Souhail Boussetta
Supervision: Sungmin O, Alexander Brenning, Gianpaolo Balsamo, Markus Reichstein, René Orth
Validation: Melissa Ruiz-Vásquez
Visualization: Melissa Ruiz-Vásquez
Writing – original draft: Melissa Ruiz-Vásquez, Sungmin O, Alexander Brenning, Ana Bastos, Sujan Koirala, René Orth
Writing – review & editing: Melissa Ruiz-Vásquez, Sungmin O, Gabriele Arduini, Souhail Boussetta, Alexander Brenning, Ana Bastos, Sujan Koirala, René Orth

weather prediction (hereafter LSMs) contributes to improved forecast skills from short-range weather forecasts to long-range seasonal predictions (Dirmeyer & Halder, 2017; Guo et al., 2011; Nogueira et al., 2020), helping to better predict extreme events like heat waves or droughts (Hirsch et al., 2019; Meng et al., 2014; Miralles et al., 2019; Zhang et al., 2008).

As LSMs are an essential component of the models that are typically used for weather forecasting systems, there have been considerable efforts in recent decades to improve LSM performance (Dutra et al., 2010; Fisher & Koven, 2020; Laguë et al., 2019; Wipfler et al., 2011). The constantly increasing computing power allows us to include more realistic descriptions of relevant processes and their interactions with the atmosphere, including soil thermodynamics, vegetation dynamics, and land cover and management (González-Rouco et al., 2021; Nemunaitis-Berry et al., 2017; Steinert et al., 2021). The increasing computing power also allows the use of big Earth observation data to characterize surface properties and better constrain model simulations (Balsamo et al., 2018; Ghilain et al., 2012; Hawkins et al., 2019; Orth et al., 2017). For LSMs that employ data assimilation, such as the Carbon Cycle Data Assimilation System (CCDAS) (Rayner et al., 2005) and the Organizing Carbon and Hydrology In Dynamic Ecosystems (ORCHIDEE) (Santaren et al., 2007), Earth observation constitutes an important data source for key land surface variables including soil moisture, vegetation state, albedo, and land use/land cover (Guillevic et al., 2002; Meng et al., 2014; Seneviratne et al., 2010). However, exploiting these new data streams for enhanced land surface model performance is not straightforward (Wulfmeyer et al., 2018).

Traditional LSMs used for weather forecasting incorporate the effect of vegetation on simulated land surface meteorology through lookup tables providing different parameter values depending on the biome type (Boussetta et al., 2013; Duveiller et al., 2022; Johannsen et al., 2019). This requires up-to-date information on land cover described through the considered biome types. Furthermore, state-of-the-art LSMs use satellite-observed vegetation indices such as the leaf area index (LAI) to describe vegetation greening, maturity, and senescence (Boussetta et al., 2013; Stevens et al., 2020). However, in most LSMs, the vegetation state is represented only through climatological seasonality, neglecting possible impacts of anomalies in vegetation functioning on the weather (Duveiller et al., 2022). Although other studies have contributed to LSM development applying high-resolution and/or more accurate Earth observation data, the full potential of LSMs in the face of the newly available data sets is not yet well exploited, resulting in opportunities for further improving weather prediction accuracy.

Different studies have demonstrated the benefits of considering spatial heterogeneity and high-resolution Earth observations in LSMs at local or regional scale. Chaney et al. (2018) accounted for the spatial heterogeneity using subgrid land surface information (300 tiles) within a 0.25° grid cell in California, US, and they analyzed its effects on the water, energy and carbon fluxes. A highlight from their results is that the accurate representation of topographic gradients has implications in the simulated interannual variability of these fluxes and, therefore, in depicting their extremes. Complementarily, L. Li et al. (2022) investigated the impacts of different spatial heterogeneity sources on the simulated water and energy fluxes over the continental US. They showed that the atmospheric forcing and the land use/land cover information are the dominant heterogeneous aspects in determining the spatial variability of the fluxes and therefore have potential to improve LSM performance. Yet another informative example applied over the continental US is from Torres-Rojas et al. (2022) with the use of subgrid tiles to represent spatial heterogeneity and to resolve nonlinear hydrological processes. They accounted for different scales in spatial heterogeneity, resulting in a very good model agreement with observations of water and energy fluxes. A similar study over continental China (Ji et al., 2023) showed that a high resolution representation of land surface parameters (30 m) substantially reduces errors in simulated hydrological variables. The results of these four studies offer insights for future developments in LSM to focus on land surface parameterizations, predefined accurate topography, soil moisture subgrid representation and urban dynamics.

As seen with the previous studies, there is ample potential for further developing LSMs in the face of the newly available Earth observation data at the local scale. Yet, this is not fully exploited globally, providing opportunities to continuously improve weather prediction accuracy. In this study, we use the ECMWF land surface modeling system (ECLand) based on the previous Hydrology Tiled ECMWF Surface Scheme for Exchange over Land (HTESSEL) to investigate the impact of updating vegetation and land cover information on model performance (Boussetta et al., 2021). Previous studies have found that updating the vegetation information in HTESSEL enhances the performance of simulated soil moisture and energy fluxes thanks to a more accurate representation of (a) the soil moisture uptake and (b) the modulation of evapotranspiration in response to soil moisture changes (Boussetta et al., 2013, 2015; Nogueira et al., 2020; O et al., 2020; Orth et al., 2017; Stevens et al., 2020). More

recent studies that use the coupled version of HTESSEL within the Integrated Forecasting System (IFS) show the subsequent effect of updated land surface information on the forecast skill. For instance, Johannsen et al. (2019) showed that large biases in temperature simulated by the IFS strongly relate to the outdated land cover representation within HTESSEL. Further, Nogueira et al. (2021) showed that it is necessary to adapt the model to the new data, that is, to perform a recalibration of model parameters. This recalibration is an important step in the process of exploiting the potential of updated land surface information since the model is well adapted to the previously used data. However, most existing studies overlook the importance of model recalibration, partially due to the lack of land observations to constrain the model parameters (Orth et al., 2016).

Even though there have been considerable efforts to exploit additional Earth observations with HTESSEL, they have never brought together all updates in one single study, nor have they performed this in combination with a parameter recalibration. Building upon the most recent HTESSEL model performance studies, we perform a comprehensive analysis with updated land surface information in ECLand as follows: (a) we update the land use/land cover information using the European Space Agency Climate Change Initiative and the Copernicus Climate Change Service (ESA-CCI/C3S) data set; (b) we introduce interannual variability of LAI and land cover fraction from the Sentinel-3 and THEA GEOV2 satellite observations; (c) we perform a recalibration of key model parameters to adjust the model parameterizations to the newly updated land cover and vegetation information. This way, we explore the contribution of near-real time land surface information and model calibration to model performance.

2. Data and Methods

2.1. Model Description

ECLand is the LSM from the operational IFS used at ECMWF (Boussetta et al., 2021). It describes the land surface fluxes of energy, water and carbon. At the interface between the surface and the atmosphere, each grid-box is divided into fractions (tiles). Each fraction has its own properties defining separate heat and water fluxes. Current Land Use/Land Cover data (LU/LC) are from the United States Geological Survey Global Land Cover Characterization (GLCC) v1.2 (Loveland et al., 2000) which is based on observations from the Advanced Very High Resolution Radiometer (AVHRR) covering 1992–1993. The nominal resolution is 1 km. The data set provides for each pixel a biome classification based on the Biosphere-Atmosphere Transfer Scheme (BATS) (Table S1 in Supporting Information S1).

Vegetation is represented by six static parameters: vegetation cover of low vegetation, vegetation cover of high vegetation, low vegetation type, high vegetation type, LAI for low vegetation and LAI for high vegetation. The first four parameters are derived from the GLCCv1.2 data by aggregating over the target grid squares. The fractional covers for low and high vegetation are obtained by combining the fractions from all the low and high vegetation types. The model phenology and its seasonality is represented by the LAI, split into low and high vegetation tiles. The LAI in the model is prescribed with a lookup table value that represents the yearly maximum LAI per vegetation type and applies a seasonal cycle based on a MODIS LAI climatology.

2.2. Modeling Experiments

We perform multiple uncoupled model experiments while continuously updating the land and vegetation information of ECLand, as listed in Table 1. We use meteorological forcing from the ECMWF atmospheric Reanalysis ERA5 (Hersbach et al., 2020) at a reduced Gaussian grid of approximately 0.5° spatial resolution and hourly temporal resolution, from 1 January 1995 to 31 December 2019 (available at: https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-era5-single-levels?tab=overview). The temperature, humidity, surface pressure and wind fields are instantaneous values and representative of the atmospheric layer at 2 m (the first two variables) and at 10 m (the last two variables) above the surface. The incoming shortwave and longwave radiation at the surface, rainfall and snowfall are provided as hourly accumulations (Boussetta et al., 2015). We use a spin-up period from 1995 to 1999, and all results shown do not include these five years.

For each of the first five modeling experiments in Table 1, we update one aspect of the LSM, that is, land cover, cover fraction, or LAI. We start from a baseline simulation (CONTROL) which is based on the land cover data set from the GLCCv1.2, static cover fraction and a prescribed LAI climatology (through lookup tables) and default

Table 1
Modeling Experiments With ECLand

Experiment	Land cover data set	Cover fraction dynamics	LAI dynamics	Land surface parameters
CONTROL	GLCC	Climatology	Climatology	Default
LC	ESA-CCI/C3S	Climatology	Climatology	Default
LC_COV	ESA-CCI/C3S	Interannual variability	Climatology	Default
LC_LAI	ESA-CCI/C3S	Climatology	Interannual variability	Default
LC_COV_LAI	ESA-CCI/C3S	Interannual variability	Interannual variability	Default
Global calibration	ESA-CCI/C3S	Interannual variability	Interannual variability	Spatially constant calibration
Regional calibration	ESA-CCI/C3S	Interannual variability	Interannual variability	Regionally varying calibration

model parameters, until we perform the LC_COV_LAI experiment in which we update all aspects including: the land cover data set using information with a 300m spatial resolution from ESA-CCI/C3S (Bontemps et al., 2017), the cover fraction interannual variability and the LAI interannual variability using 10-daily data from Sentinel-3 (Verger et al., 2022) and THEA GEOV2 (Verger et al., 2020) at 1 and 4 km spatial resolution respectively, but with default model parameters. The cover fraction and LAI interannual variability refers to monthly values that vary every year, in contrast to climatological monthly means, based on the monthly mean calculated over the period 1993–2019, that is, we do not use new lookup tables for LAI and cover fraction but we apply observational data directly to each grid cell.

The new land cover maps have a total of 22 land cover classes based on the Land Cover Classification System developed by the United Nations Food and Agriculture Organization. To be used in the model, the 22 classes are converted to the BATS classes (Figure S1 in Supporting Information S1). The comparison between ESA-CCI and GLCCv1.2 based vegetation cover fraction maps shows an increase (decrease) of low (high) vegetation cover in forest areas and a decrease of low vegetation cover favoring more bareground (Figure S2 in Supporting Information S1). The vegetation types show a split of the mixed and interrupted forest types between pure forest and low vegetation types. The new LAI interannual variability (Figure S3 in Supporting Information S1) extracts a seasonally varying LAI lookup table per vegetation type. It defines the ratio of low and high LAI components which is then applied to disaggregate the satellite based LAI.

For the last two modeling experiments (last two rows in Table 1) we recalibrate six soil- and vegetation-related model parameters which are listed in Table 2. We select these key parameters based on (a) previous parameter calibration studies with HTESSEL (MacLeod et al., 2016; O et al., 2020; Orth et al., 2016) and (b) the selection of parameters that are theoretically strongly related to the vegetation type, and thus, expected to affect the performance of ECLand. Specifically, the minimum stomatal resistance is considered in combination with LAI as what matters for evapotranspiration stress is the ratio between minimum stomatal resistance and LAI. The recalibration experiments are as follows: first, we perform a global calibration in which we search a unique parameter set that works best overall for all selected grid cells (i.e., spatially constant calibration); second, we perform a regional calibration in which we define the best parameter set individually for each grid cell (i.e., regionally varying calibration). We use Latin hypercube sampling (McKay et al., 1979) to select 1000 random combinations

Table 2
Model Parameters Considered for the Recalibration Experiments

Model parameter	Units	Range of default values	Range of perturbation factors
Hydraulic conductivity	m s^{-1}	0.83–3.83	0.01–100.0
Humidity stress function	m s^{-1} hPa	0.00–0.03	0.25–4.0
Minimum stomatal resistance	s m^{-1}	80–250	0.25–4.0
Soil moisture stress function	–	–	0.25–4.0
Total soil depth	cm	1–800	0.5–2.0
Transmission of net solar radiation through vegetation	–	0.03–0.05	0.1–10.0

Table 3
Reference Data Sets for Model Performance Evaluation

Output variable	Reference data set	Source of information
Near-surface soil moisture normalized anomalies	SoMo.ml 0–10 cm soil layer (upscaled in situ observations)	O and Orth (2021)
	GLEAM 0–10 cm soil layer (physical-based model)	Martens et al. (2017)
	MERRA-2 0–5 cm soil layer (reanalysis)	Gelaro et al. (2017)
Deep soil moisture normalized anomalies	SoMo.ml 10–50 cm soil layer (upscaled in situ observations)	O and Orth (2021)
	GLEAM 0–250 cm soil layer (physical-based model)	Martens et al. (2017)
	MERRA-2 0–100 cm soil layer (reanalysis)	Gelaro et al. (2017)
Surface latent heat flux	FLUXCOM RS V6 (upscaled in situ observations)	Jung et al. (2019)
	GLEAM (physical-based model)	Martens et al. (2017)
	MERRA-2 (reanalysis)	Gelaro et al. (2017)

of perturbation factors independently chosen for each parameter within a specified range. The selection of the range for each parameter follows previously used ranges in recent literature about parameter sensitivity analysis and recalibration of similar parameters in HTESSEL (MacLeod et al., 2016; O et al., 2020; Orth et al., 2016, 2017; Johannsen et al., 2019; Stevens et al., 2020).

The original parameters have been calibrated internally by ECMWF for operational Numerical Weather Prediction purposes, given the static maps, model physics and other settings of the ECLand cycle used in this study. The vegetation-related parameters are dependent on the vegetation type through lookup tables, that is, the default values change in response to the new vegetation information that we provide to the model. Then, during our recalibration experiments, we do not modify the values of the lookup tables directly but we apply perturbation factors to them, to either amplify (factors >1) or to reduce (factors <1) the effect of each parameter in the water and energy exchange of the land surface with the atmosphere. We apply the range of perturbation factors equally to all vegetation types and they multiply the default value of the parameter in each grid cell given by the lookup tables. In other words, the perturbation factors do not vary between vegetation types but the resulting parameter values (=perturbation factor × lookup table value) are different for each grid cell.

For computational efficiency, we perform the parameter calibration experiments at 230 randomly chosen grid cells across the globe (their location is shown in global maps in Section 3.2.2). We only consider grid cells with a long-term mean Enhanced Vegetation Index (EVI) greater than 0.2 to exclude regions with scarce vegetation. The EVI data are derived from the satellite MODIS V6 (Didan, 2015). First, we select 30 grid cells to run the 1,000 simulations (one for each parameter set), and we select the best 100 parameter sets according to the model performance metric introduced in Section 2.3. Second, we run the best 100 parameter sets in the remaining 200 grid cells and we again evaluate their performance to find the best-performing parameters that work over a wider range of climate regimes. The selection of the first 30 grid cells is a calibration strategy for computation time efficiency, to narrow down the parameter ranges. Our results are based on the 230 grid cells sampled from diverse regions, to sufficiently represent spatial heterogeneities.

2.3. Model Evaluation

Not only are novel Earth observations beneficial for modeling directly as in the case of the LAI and land cover used here, but furthermore the increasing suite of Earth observations allows a more comprehensive model evaluation. This way, as reference data sets in this study we have chosen gridded data sets based on interpolating and upscaling station measurements through machine learning methods, which are independent from model-based data sets and hence well suited to evaluate model simulations. Additionally, we employ reanalysis(-like) data sets which benefit from a growing number of available satellite observations which are assimilated into the underlying model simulations.

For each model experiment, we compare simulated latent heat flux and soil moisture with respective global gridded observation-based data sets listed in Table 3. In Table S2 in Supporting Information S1, we present a brief description on the different reference data sets and their advantages and limitations. We use all the data sets at a 0.5° spatial resolution and at a daily temporal resolution. In the case of the surface latent heat flux from FLUXCOM, we linearly interpolate the original 8-daily values to compute daily values. While we use absolute values for latent heat flux, for near-surface and deep soil moisture we analyze normalized anomalies to account for different systematic errors in ECLand and in each reference data set. To compute normalized anomalies for each soil moisture variable and data set (a) we subtract the linear long-term trend from the time series, (b) we remove the mean seasonal cycle calculated at daily time steps over the period 2000–2019, and (c) we divide by the standard deviation of the resulting time series.

Since we use three different reference data sets per output variable and they are all derived with different approaches (e.g., machine learning-based or physical-based), we need to account for their uncertainties and their biases in order to compute a robust model performance. For this, we

use censored RMSE (cenRMSE) as a performance metric, which is based on modified root mean squared error (RMSE). The term "censored" refers to a value that occurs outside the range of a measuring instrument (Fridley & Dixon, 2007). We compute the cenRMSE as follows:

$$\text{cenRMSE} = \sqrt{\sum_{i=1}^{n} dy_i^2} \tag{1}$$

$$dy_i = \min(|\hat{y}_i - y_{i,r}|), r = 1, 2, 3 \tag{2}$$

$\hat{y}_i$ is the model value in time step i and $y_{i,r}$ is the reference data for the three references ($y_{i,1}$, $y_{i,2}$, $y_{i,3}$). $dy_i = 0$ if $\hat{y}_i$ is in the interval defined by the range of the reference values, that is, if the value of the model output in the time step i falls within the range of the reference values ($\min(y_{i,1}, y_{i,2}, y_{i,3}) < \hat{y}_i < \max(y_{i,1}, y_{i,2}, y_{i,3})$), otherwise the minimum is taken to compute the cenRMSE. The cenRMSE behaves like RMSE outside the interval and is 0 if all predictions are within the range of reference values.

Specifically for the parameter calibration experiments, we combine the cenRMSE performance metric of the three target variables (i.e., near-surface soil moisture, deep soil moisture and surface latent heat flux). We rank the 1,000 perturbation factors individually for each variable and then we add the individual ranks up. The lowest (highest) sums constitute the best (poorest) perturbation factors in terms of model performance.

2.4. Spatial Variability of Regional Parameters

We extend our analysis to the spatial features of calibrated model parameters (Table 2). We employ random forest models (Breiman, 2001; Molnar, 2020) (hereafter RF) to predict each of the six calibrated parameter values across grid pixels (six RF models are used). As predictor variables we use (a) long-term mean climatic and land surface characteristics such as aridity, temperature and EVI, (b) differences in high and low vegetation cover between the two land cover data sets used in the modeling experiments (ESA-CCI/C3S and GLCC) (Boussetta et al., 2021), and (c) the values of the remaining five parameters (other than the target parameter). As hyperparameters we choose 1,000 trees, a maximum depth of 25 and oob_score = True to estimate the generalization score; otherwise we apply default settings of scikit-learn RF regressor (i.e., max_features = "auto," min_samples_split = 2, min_samples_leaf = 1, bootstrap = True). Previous studies suggest that RF performance is relatively insensitive to different hyperparameter settings and generally works well with default settings (Martínez-Muñoz & Suárez, 2010; Probst et al., 2019; Schratz et al., 2019), therefore we do not perform a hyperparameter tuning here.

We use information from the 230 grid cells for the RF training. We assess the performance of the RF models by computing the R^2 between the predicted and the observed target variables for out-of-bag (OOB) data that was not used for training (hereafter referred to as estimate of R^2) (W. Li et al., 2021). We infer the relative importance of each predictor variable to the target (i.e., how strongly related is each predictor to the target according to the model, while accounting for the other predictors in the model) from SHapley Additive exPlanations (SHAP) feature importance which is based on the average marginal contribution of each predictor to the modeled target variable (Lundberg & Lee, 2017; Sundararajan & Najmi, 2020). We look for spatial correlation or co-variation as expressed by the SHAP values within the RF models. We chose this set of predictors because they have a physical link with the targets and therefore the spatial co-variation between targets and predictors allows identifying which processes are more influential in the parameter calibration. The process-related justifications for the inclusion of each of the predictor variables are: (a) the climatological properties of a region, such as aridity (which is a function of precipitation and radiation) and temperature partially determine the vegetation type within a region and therefore the surface latent heat flux and the soil moisture which are strongly linked with EVI; (b) the ratio between high and low vegetation cover in a region affects the surface latent heat flux and the soil moisture variability; and (c) the degree of interdependence of the model parameters might influence the parameter recalibration results.

We note a potential caveat in our approach with the RF due to existing relationships among our selected set of predictors. Accordingly, we compute individual Spearman correlations (Wilks, 2011) among the predictors to report the magnitude of these associations and to identify the most affected variables.
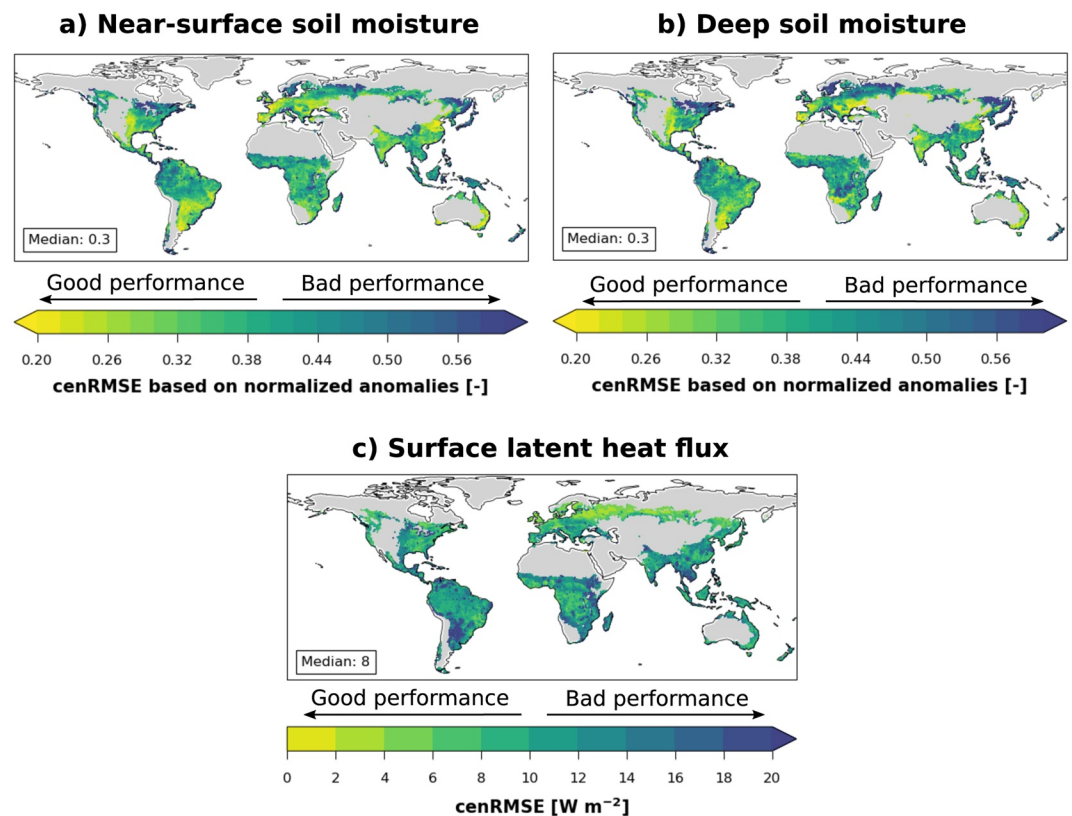
Figure 1. cenRMSE performance metric of CONTROL simulation for (a) near-surface soil moisture, (b) deep soil moisture and (c) surface latent heat flux. cenRMSE is computed based on absolute values for latent heat flux, while normalized anomalies are used for soil moisture. Numbers in the textboxes represent the global median. Gray areas are masked as their long-term mean EVI is lower than 0.2.

3. Results and Discussion

3.1. Impact of Updated Land Surface Information on Model Performance

Figure 1 shows ECLand's model performance in the CONTROL experiment. In general, the model performance varies considerably across regions. For near-surface and deep soil moisture (Figures 1a and 1b), we see relatively good performance in the mid-latitudes of Europe, North America and southern South America. On the contrary, the model performs poorly in high-latitude regions, possibly due to high uncertainty in soil moisture-related processes, for example, soil freeze/thaw cycles (Diro et al., 2018; Dutra et al., 2010, 2011). In some regions, the model performance for deep soil moisture is slightly poorer than for near-surface soil moisture. This can be due to the high uncertainty among the reference data sets for deep soil moisture values as a consequence of sparse observations (Denissen et al., 2020; Koster et al., 2020; W. Li et al., 2021). For the surface latent heat flux (Figure 1c) the cenRMSE is relatively good in central and eastern Europe and North America, which might be related to the high density of observations that can support model development and parameter calibration dedicated to these regions (Stegehuis et al., 2013).

Figure 2 shows the performance of the experiment with the most updated land information (LC_COV_LAI) compared to the performance of the CONTROL experiment. We find a general deterioration of model performance (red color) for all three variables considered which is related to the high sensitivity of the RMSE-based metrics to outliers. Recomputing the cenRMSE without the 10% largest disagreements between LC_COV_LAI and CONTROL simulation (Figure S4 in Supporting Information S1) shows, however, that the percentage difference in cenRMSE improves in most regions over Figure 1. Therefore, on average, an update of the land surface information in ECLand has positive impacts on the prediction of surface latent heat flux and near-surface and deep soil moisture for most of the time, while creating stronger deviations at particular times.
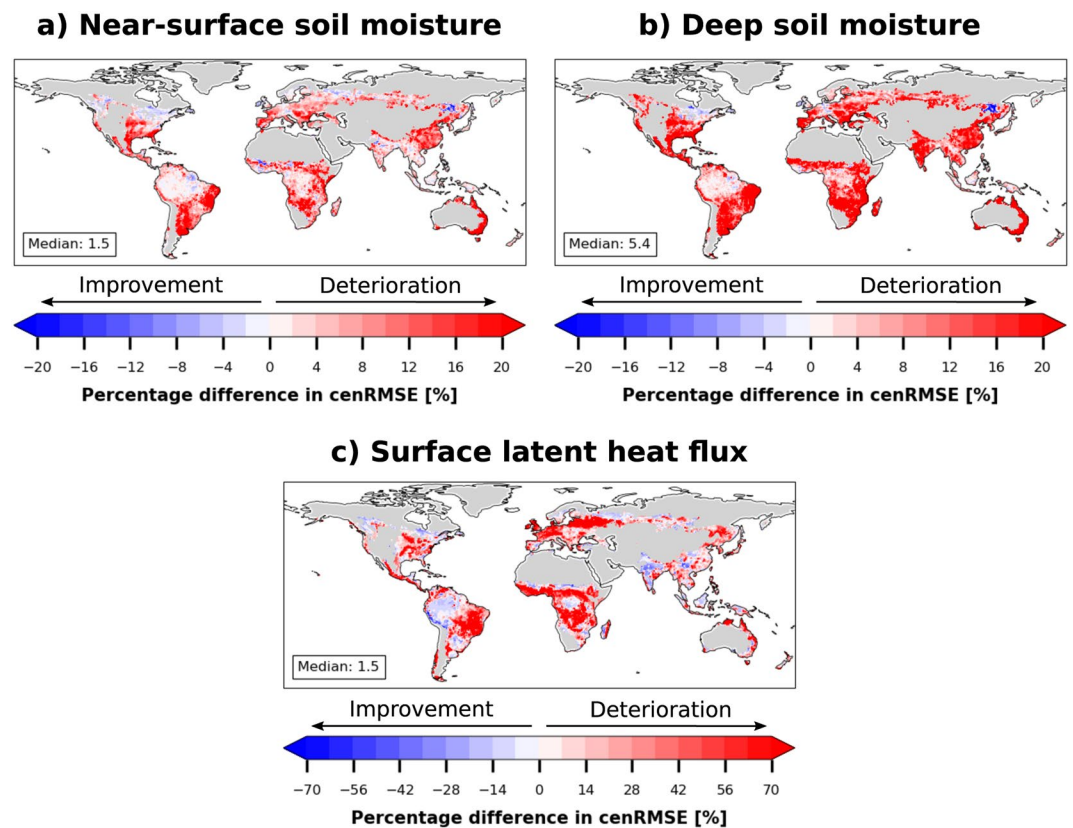
Figure 2. Similar to Figure 1, but for percentage differences in performance: LC_COV_LAI minus CONTROL divided by CONTROL.

We assume that the introduction of the interannual variability of LAI is the main cause of the strongest outliers since the magnitude of the errors increase substantially after we perform the LC_LAI model experiment (Figures S5c, S6c, and S7c in Supporting Information S1). We see a high interannual variability of LAI for low vegetation type in the Amazon river basin (Figure S3 in Supporting Information S1) that was neglected in the CONTROL simulation, but up to the LC_LAI experiment the model parameters do not respond to this variability (i.e., they are not yet calibrated), as also stated by Nogueira et al. (2021). As a second cause for the outliers, the new high and low vegetation cover data sets from ESA-CCI/C3S have an increase of low vegetation cover at the expense of the high vegetation cover in forest areas (Figure S2 in Supporting Information S1), like the Amazon river basin and the Congo river basin where we see the strongest differences between the cenRMSE (Figure 2) and the cenRMSE without outliers (Figure S4 in Supporting Information S1).

The updated land surface information has a much clearer impact on the simulation of latent heat flux compared to soil moisture, as indicated by a larger magnitude of percentage changes in surface latent heat flux. Also, the spatial patterns of improvement/deterioration are not always consistent between latent heat flux and soil moisture; for instance, in southern South America and India, there are improvements in most areas for surface latent heat flux but for both near-surface and deep soil moisture we find deterioration. This points to possible weaknesses in the representation of the coupling between latent heat flux and soil moisture in the model, as also stated in other studies (Dirmeyer & Halder, 2017; Fairbairn et al., 2019; Meng et al., 2014; Quillet et al., 2010; Santanello et al., 2009; Wulfmeyer et al., 2018; Zhang et al., 2008). The improvement in surface latent heat flux in South America coincides with the region where we find more interannual variability in LAI for low vegetation type (Figure S3b in Supporting Information S1) corresponding mostly to deciduous shrubs and short grass (Figure S1b in Supporting Information S1). This is likely linked with a more accurate description of the LAI interannual variability in this region, which was neglected with the prescribed climatology in the outdated version of ECLand.

We also look at the model performance of each individual experiment in terms of the three considered output variables (Figures S4, S5, and S6 in Supporting Information S1). In general, the spatial patterns of improvement
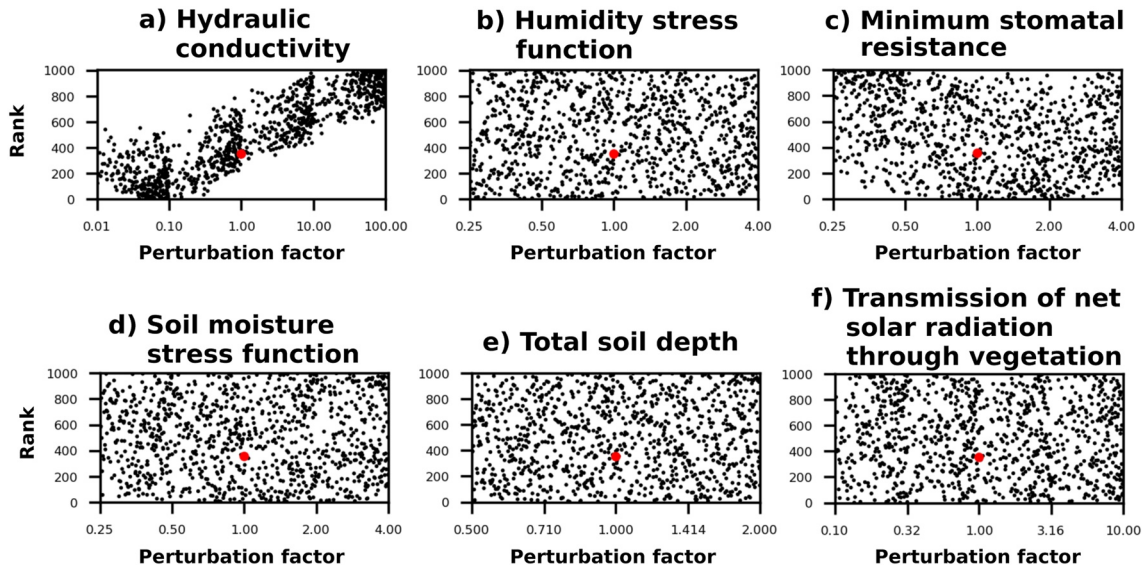
Figure 3. Relating model performance to perturbations in the considered individual ECLand parameters: (a) hydraulic conductivity, (b) humidity stress function, (c) minimum stomatal resistance, (d) soil moisture stress function, (e) total soil depth and (f) transmission of net solar radiation through vegetation. Red dots indicate the performance of the default parameterizations (i.e., no perturbation). A rank value of 0 (1,000) in the *Y*-axis indicates the best (poorest) perturbation factor in terms of model performance.

and deterioration are similar to the results in Figure 2. Comparing the magnitudes of the changes we find that the strongest effect on the model performance is exerted by the land cover type update, which is present in all experiments. The LAI interannual variability update has the second strongest effect on the model performance (Boussetta et al., 2013, 2015; Duveiller et al., 2022; Stevens et al., 2020).

3.2. Effect of Recalibration of Model Parameters

3.2.1. Ranks of the Parameter Sets

We rank the 1,000 model simulations with perturbed parameter values according to the cenRMSE performance metric of the three target variables (see Section 2.3), and relate the ranking to individual parameter perturbations in Figure 3 in order to assess their individual contribution. Table S3 in Supporting Information S1 shows the individual optimal perturbation factors for the model parameters. Hydraulic conductivity and minimum stomatal resistance show the strongest systematic influence on model performance, similar to the results from Orth et al. (2016, 2017).

Hydraulic conductivity governs the water transport in the soil and is therefore directly linked to soil moisture and evapotranspiration (latent heat flux). We find that a substantial reduction of the hydraulic conductivity from its default value improves model performance. This reduces percolation of infiltrated water and therefore increases near-surface soil moisture and ultimately latent heat flux (O et al., 2020). If the model with the new land surface information displays a general dry bias in soil moisture, a lower hydraulic conductivity would help in retaining more water into the soil matrix. One possible way for the introduction of a dry bias with the new land surface information is an increase in the LAI. Actually, Boussetta et al. (2021) reports that the LAI component for high vegetation tends to increase by up to 80% in the summer months. An increase in LAI compared to the CONTROL experiment should be related to an increase in evapotranspiration and therefore a decrease in the soil moisture.

In the case of the minimum stomatal resistance it strongly relates to evapotranspiration as it modulates the exchange of moisture from vegetated surfaces (Orth et al., 2016). Our results suggest that there is an optimum perturbation value for the minimum stomatal resistance between 1 and 2, that is, close to the default parameterization, thus, modifying it has little potential to improve the model. The increase in stomatal resistance should be related to an excess of evapotranspiration with the new land surface information, for instance, with an increase of LAI, compared to the CONTROL experiment.

Table 4
Range of Observed Values of the Considered Model Parameters in the Experiments Reported in Literature

Model parameter	Units	Range of recalibrated values	Range of experimental values	Reference
Hydraulic conductivity	m s^{-1}	0.0166–3.83	10^{-7}–10	Medici and West (2021)
			8.20E^{-5}–2.88E^{-4}	Verbist et al. (2009)
			5.5E^{-9}–60.87E^{-7}	Sihag et al. (2019)
			2E^{-9}–2.26E^{-3}	Araya and Ghezzehei (2019)
			6.46E^{-8}–8.13E^{-5}	Coutadeur et al. (2002)
Minimum stomatal resistance	s m^{-1}	27.2–827.5	39.8–78.1	Kardel et al. (2010)
			ca. 100–700	D. S. Niyogi and Raman (1997)
			20.4–41.3 (one site)	D. S. Niyogi et al. (1997)
			ca. 100–1,000	Mehrez et al. (1992)
			ca. 1,000–15,000 (one plant)	Cockburn et al. (1979)
			ca. 100–6,000	Turner (1974)
			ca. 10–250	Monteith et al. (1965)
Transmission of net solar radiation through vegetation	–	0.003–0.495	ca. 0.2–0.6	Hardy et al. (2004)

We also analyze the influence of parameter perturbations on model performance in terms of the considered individual variables (Figures S8, S9, and S10 in Supporting Information S1). The clear pattern of better model performance in the case of lower hydraulic conductivity found in Figure 3 is mainly related to an improvement of the soil moisture performance, especially for the near-surface layer (Figure S8 in Supporting Information S1) and in a second extent for the deep soil moisture (Figure S9 in Supporting Information S1). For the minimum stomatal resistance the pattern found in Figure 3 is related to variations in the simulation performance of latent heat flux (Figure S10 in Supporting Information S1), since this parameter is a dominant driver in the evapotranspiration process. Additionally, the total soil depth is relevant for the simulation performance of deep soil moisture (Figure S9 in Supporting Information S1), as also found in a similar study by Hawkins et al. (2019). This illustrates that different parameters matter for different land surface variables, as well as that different observational data sets are needed to constrain different parameters.

It is relevant for LSM development that model parameters still have physical meaning, rather than becoming solely effective numbers (i.e., parameter values that reduce the model errors but that don't have any physical reasoning behind them). It is difficult to prove that the parameters are theoretically plausible because model parameters often represent physical quantities that are not directly observable. Even when the parameters can be observed, they are based on small soil samples that are location-specific, whereas the LSM is applied globally. When performing this study, there were no data sets to our knowledge that provided global estimates of these parameters, and even if there were, they could also vary in the same site due to the heterogeneity of the soil. Nevertheless, for those parameters that can be measured (like hydraulic conductivity, minimum stomatal resistance and transmission of net solar radiation through vegetation), we look for typical observational values found in literature and we compare them with our calibrated values. In Table 4 we add the observational values found in different field experiments available in the literature. The values of these parameters vary greatly among the different studies. Even for the same study, the variability of one parameter is substantial and depends on the time of the day that the experiments were carried out, the level of the soil that the samples were taken from, among other factors. Yet, our recalibrated values are within these different ranges of observations, so at least they seem physically plausible and hence not necessarily entirely effective parameters.

3.2.2. Model Performance in Parameter Calibration Experiments

Figures 4 and 5 show the model performance changes relative to CONTROL after the global and the regional recalibration across 230 grid cells, respectively. Generally, for the global calibration (Figure 4) we find inconsistent results with improved or deteriorated model performance depending on the grid cell. This suggests that there is no one(calibration)-fits-all(regions) solution, probably related to the spatial heterogeneity in climate along with different land surface characteristics, or its insufficient representation in the current default values in the model (like for specific vegetation types, soil textures, etc.) (Laguë et al., 2019; Nogueira et al., 2021), as can be seen
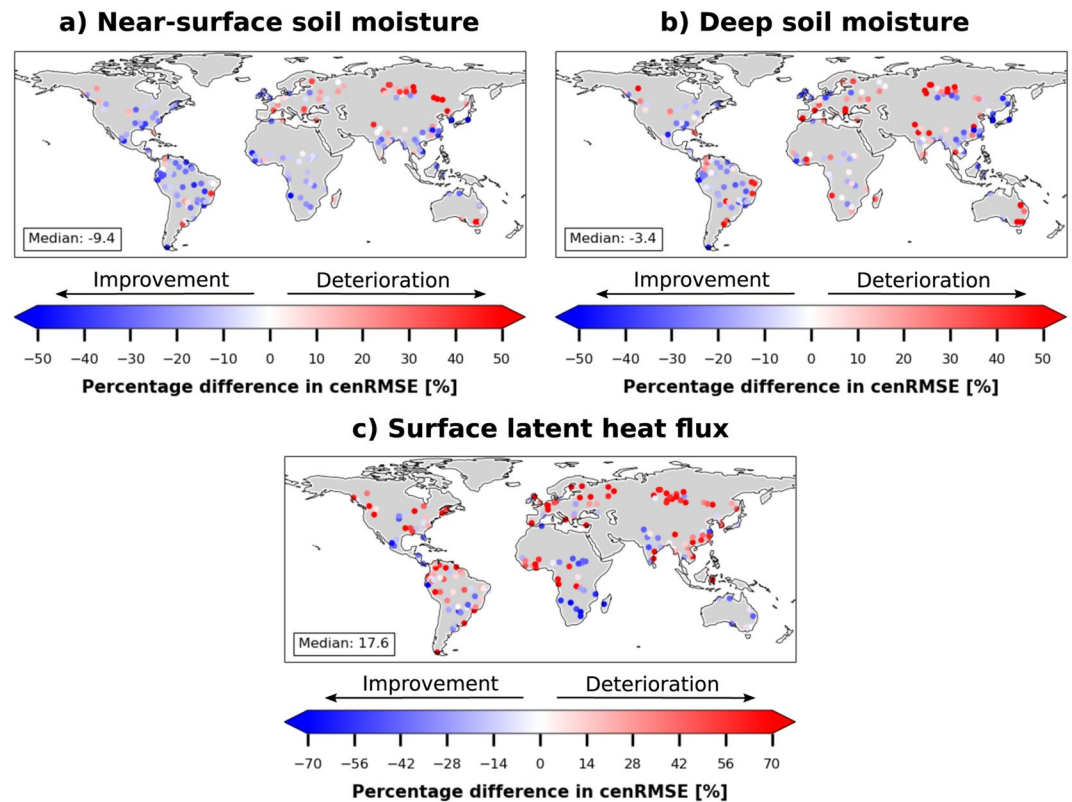
Figure 4. Similar to Figure 1, but for percentage differences in performance: Global calibration minus CONTROL divided by CONTROL.

from the spatial distribution of the calibrated parameter values in Figure S11 in Supporting Information S1. After the global calibration we already see an improvement in both soil moisture variables but it is not always the case for the surface latent heat flux, probably due to compensation in model performance between variables (McCabe et al., 2005). This is expected as the newly applied datastreams are related to land cover and vegetation structure. Specifically, the model performance in the grid cells in northern Asia always degrades from a global calibration, whereas for the other regions we see mixed results.

After the regional calibration, we find substantial improvement in model performance for all three variables as shown in Figure 5. See also Figure S12 in Supporting Information S1 for comparisons of model performance between the regional and global calibrations. In a similar study for another LSM, Xie et al. (2007) found an improvement in model performance after a regional calibration of model parameters. This suggests that parameters should sufficiently reflect land surface heterogeneity, different climate zones, different biome types, etc. The regional calibration leads to better model performance for most grid pixels, except for high latitudes in Northern Asia, possibly due to high uncertainty in the representation of soil freeze processes, as found in other studies (Diro et al., 2018; Dutra et al., 2010, 2011). Given that the regional calibration exceeds the model improvement of the global calibration by 20% or more in some of the tested grid cells (Figure S12 in Supporting Information S1), there should always be a motivation to calibrate each grid cell individually instead of seeking a single globally perturbation factor for the parameters.

To aggregate our main findings, Figure 6 shows the median global change in model performance for each experiment and variable. Most of the experiments do not show clear model performance improvement with regards to the CONTROL simulation before recalibration. Only the regional calibration experiment shows improvement in all output variables, which calls for parameter recalibration after updating land surface information on LSMs to exploit the benefits of Earth observation developments (Nogueira et al., 2021). This is specifically the case for a regional (spatially varying) as opposed to the global (spatially constant) calibration as this can better account for spatial heterogeneities, and compensate for potentially related shortcomings in the model structure
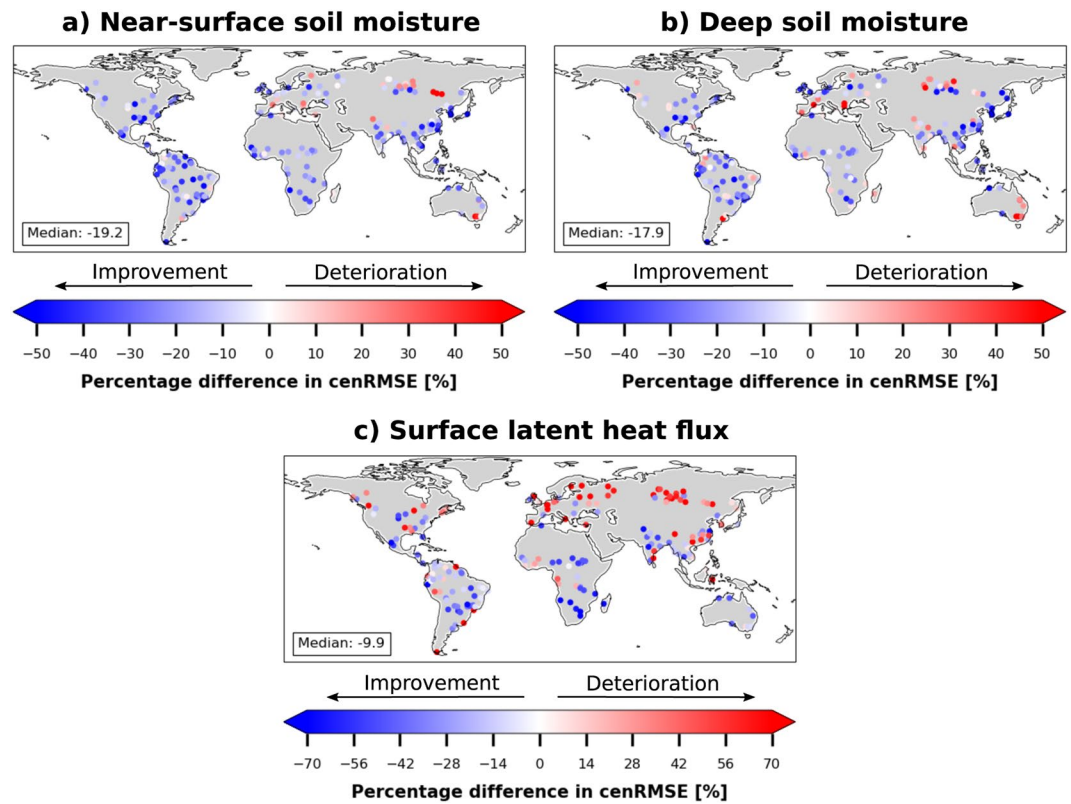
Figure 5. Similar to Figure 1, but for percentage differences in performance: Regional calibration minus CONTROL divided by CONTROL.

(Xie et al., 2007). The variability of the experiments (represented by the error bars in Figure 6) for the surface latent heat flux is higher than for the two soil moisture variables. We attribute this to a direct effect on latent heat flux from the perturbation of the selected parameters because these are mostly related with evapotranspiration, whereas they have an indirect effect on soil moisture (Jefferson et al., 2017; Montzka et al., 2017).
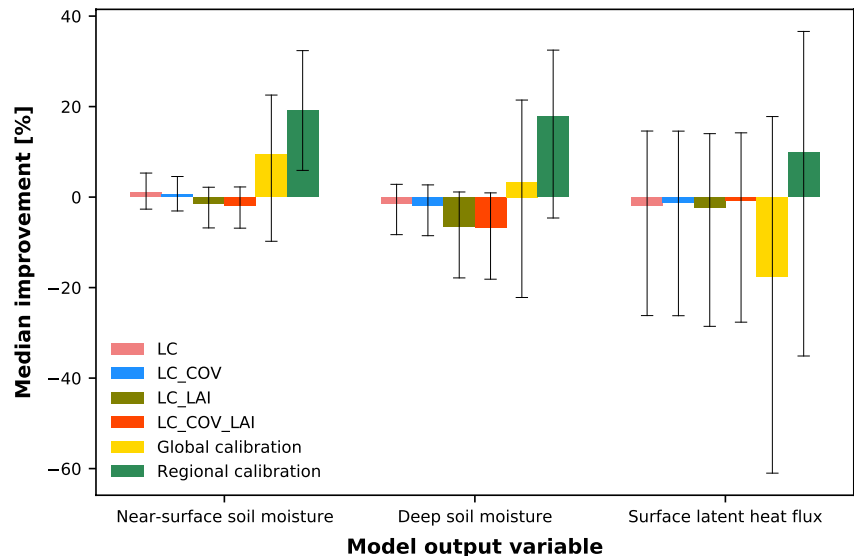

Figure 6. Summary of ECLand performance for each experiment compared to the CONTROL simulation. Medians of percentage change of cenRMSE across 230 grid cells are shown. The error bars represent the 25th and 75th percentile.

In a final step, we study model performance changes in wet versus dry regions by producing Figure 6 for such regions separately (Figure S13 in Supporting Information S1). The effect of updating land surface information in ECLand on model performance is generally stronger in dry grid cells than in wet grid cells. This is expected since vegetation plays a more important role for modulating the exchange of water and energy in dry-to-transitional regions, whereas the role of the vegetation and relevant land processes in comparison to the effect of atmospheric dynamics is less prominent in wet regions (Denissen et al., 2020; Miralles et al., 2019; Seneviratne et al., 2010).

3.3. Attribution Analysis of Spatial Patterns of Regional Parameter Calibration

In a final step, we analyze the spatial patterns of the optimal parameter perturbations determined in the grid cell-wise model calibration shown in Figure S11 in Supporting Information S1. In order to explain the spatial pattern of each parameter we consider several predictors including climate and vegetation characteristics, as well as the calibrated values of the other considered parameters. This attribution analysis is done separately for each parameter (target in the regional calibration). Figure 7 shows that overall for each of the modeled parameters, the remaining parameters are the best factors to predict the values of the target: the parameter values depend most strongly on each other; this emphasizes the important role of equifinality, and it means that parameter sets need to be chosen consistently and parameters cannot be calibrated individually. The motivation behind this analysis was to see if an extrapolation of the calibrated values of parameters to other grid cells (beyond the 230 selected grid cells) was possible using climatological and land/vegetation characteristics of the grid cells, instead of individual grid cell calibration which is computationally expensive. Our results suggest that this is not possible since the most influential predictors are the calibrated parameters and therefore we cannot extrapolate their values simply using climatological and land/vegetation information. Only for the humidity stress function (Figure 7b) and for the transmission of net solar radiation through vegetation (Figure 7f) the difference in vegetation type and the temperature are important predictors (other than the remaining model parameters) in the RF models, respectively. We attribute this to an equifinality problem in the model and accept it as a caveat in our analysis: we select only the best parameter sets while other sets might perform almost as good as the best set (Williams et al., 2009).

The RF models have in general a good model performance (Figure S14 in Supporting Information S1), meaning that the considered factors can explain the spatial patterns of model parameters. The hydraulic conductivity calibration has the best RF model performance due to the clear systematic pattern in the parameter set ranks (Figure 3a), specially given by the dependence of the near-surface soil moisture model performance on this parameter (Figure S8 in Supporting Information S1).

The relative importance is analyzed here for correlation or co-variation and not causation. We acknowledge that some of the selected factors are highly correlated (Figure S15 in Supporting Information S1) and their actual relative importance might be reduced by the co-variation (Ghosh & Maiti, 2021). The most cross-correlated ones are: hydraulic conductivity and total soil depth; minimum stomatal resistance and soil moisture stress function; EVI and aridity; EVI and temperature; and the differences in high and low vegetation cover. Although most pairs of factors show correlations lower than 0.2, we accept collinearities as a caveat in our analysis since independence is not the usual case in Earth system sciences (W. Li et al., 2021; Silva et al., 2022; He et al., 2023; Wadoux et al., 2023).

4. Implications for Future Model Calibration Strategies

This study provides an approach on how to perform parameter recalibration in the ECLand model following the inclusion of newly available Earth observations. We find a clear added value of parameter recalibration for model performance which allows the model to benefit (more) from the accuracy of the newly included Earth observations. Therefore, we feel that our approach and its limitations are relevant to the broader modeling community as well as to forecasting centers as also coupled forecasts may be impacted by emerging Earth observations and related recalibration strategies for the LSM. For this reason, we want to reflect on some aspects related to model recalibration in this section.

Most parameters in LSMs originally had an actual physical meaning such as hydraulic conductivity or stomatal resistance. As models evolve, process representations are revised and expanded and underlying soil and vegetation characteristics are updated. As a result of such changes, the original parameter values may be inconsistent with the updated model structure and data. Hence, they need to be recalibrated, making them "effective" rather
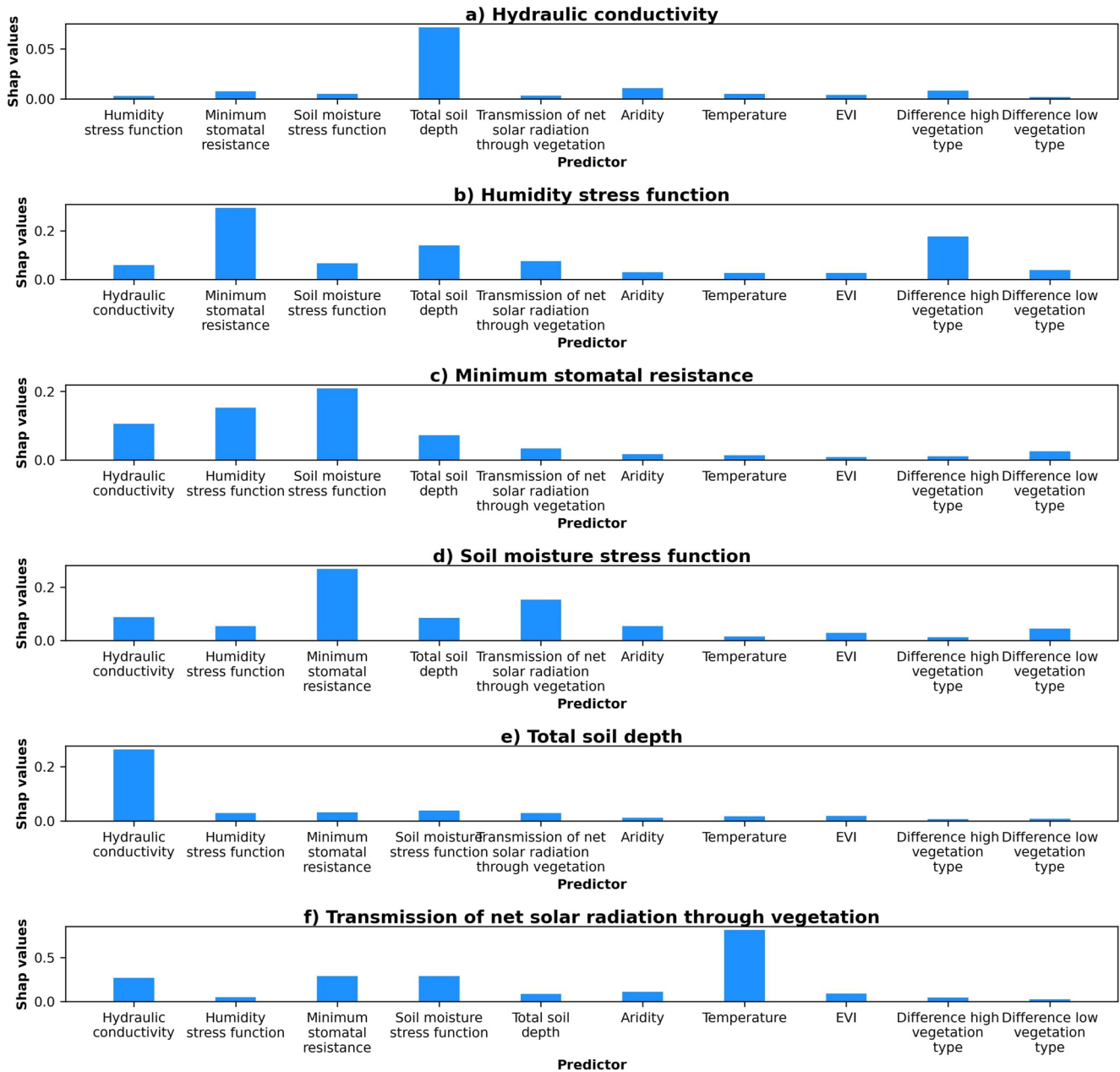
Figure 7. Relative importance (SHAP values) of multiple factors to explain the spatial patterns of regionally calibrated model parameters for (a) hydraulic conductivity, (b) humidity stress function, (c) Minimum stomatal resistance, (d) soil moisture stress function, (e) total soil depth and (f) transmission of net solar radiation through vegetation. Note that the *Y*-axes have different ranges.

than physical parameters. Comparing the recalibrated parameter values with respective observations from soil samples (Table 4) allows us to assess the extent to which parameters have deviated from physically meaningful values through the model development and recalibration process. In our case, as the range of reported values is large as a consequence of the heterogeneity of the soil, parameter values are still physically plausible.

Generally, it is beneficial to aim for physically interpretable parameters as their comparison with lab or field observations provides an opportunity to independently validate model calibration results. This can also reveal structural inconsistencies in a model if the parameters need to change beyond a physically plausible range in order to maintain reasonable model performance. In this context, novel high-resolution data sets about land surface characteristics are valuable, such as the 1 km resolution data set of global land surface parameters (L. Li

et al., 2023) derived through the development of machine-learning models utilizing the latest and most accurate available data sets related to land use and land cover, vegetation, soil, and topography. Higher resolution means smaller grid cells within which soil and vegetation types are often more homogenous such that model parameters can better reflect land surface conditions compared with larger grid cells with mixed composition where the resulting behavior in the land climate system is an emerging property. As an outlook, using higher resolution data sets in future work will enable us to derive more physically meaningful calibration results and to check the calibrated parameter values against observations and thereby have another type of validation of the model structure and implemented processes.

Although the current manuscript is focused on offline assessment, the ultimate goal of model development and parameter recalibration is to yield more accurate weather forecasts and climate projections which are done with coupled models. In such complex schemes it is not obvious whether and how changes in the land surface model affect the interaction with other model components and the overall model simulation. Evidence from previous studies is mixed, showing that land surface model improvements can benefit the coupled system (Orth et al., 2016), while in other cases this was not the case due to compensating errors and uncertainties in parameters and physics schemes (Santanello et al., 2013, 2019). One way to address this is to revise model formulations affecting the land-atmosphere coupling and/or including more processes (D. Niyogi, 2019). A comprehensive revision of the land-atmosphere coupling should include all land surface variables that describe the water, energy and carbon balances (like latent and sensible heat flux, soil moisture, runoff, Gross Primary Productivity, among others) and a broad list of vegetation-related parameters (like hydraulic conductivity, skin conductivity, minimum stomatal resistance, rooting depth, hydraulic diffusivity, among others). Furthermore, also a recalibration of the entire coupled model, instead of the land surface model alone, could help while this is a considerable effort.

Model calibration is challenged by the limited availability of observations to validate model performance in terms of a comprehensive set of variables. This can lead to the problem that model performance is optimized for considered variables while degrading it for others. It has been shown that this risk decreases the more variables are considered in the calibration (Orth et al., 2017). Therefore, the increasing suite of Earth observations offers the opportunity to calibrate models more robustly. This is even more the case as various (partly) independent data sets describing the same variable can be used in order to consider the observational uncertainty in the model calibration, as done in this study.

With regards to global and regional recalibration exercises our study reveals how successful or not is the adaptation of parameters to land surface characteristics, that is, if this was working perfectly, the local recalibration should yield no improvements over the global recalibration. This has potential implications for the creation of lookup tables in LSM and forecasting systems, which do not properly translate globally valid parameters to different soil and vegetation types prevailing locally. An update of these lookup tables could allow that future recalibrations can be done globally only.

Another important aspect is the timing of performing parameter recalibration, that is, when this effort is justified and can be expected to yield significant improvements in model performance. We think that it would be adequate to spend the effort of a parameter recalibration as soon as the model performance is found to degrade with updated land surface data. If the model performance remains similar, while we expect that it should improve because it includes a more close-to-reality land surface, then the recalibration step seems less urgent. Another aspect to consider is the duration of cycles of model development, and of satellite generations and data sets resulting from them. As the completion of such cycles can be expected to involve major model adaptations, it may be worth considering to perform recalibration to harvest the full potential of the model updates.

5. Summary and Conclusion

Recent studies performed substantial efforts for exploiting additional Earth observations in ECLand model validation (Boussetta et al., 2013, 2015; Orth et al., 2017; Nogueira et al., 2020; O et al., 2020; Stevens et al., 2020). However these experiments have never included all updates in one single study. Neither have they performed a follow-up recalibration of the model to exploit the benefits of including more accurate land surface information. Here we make a step in this direction with our comprehensive modeling experiments (Gupta et al., 1999), not only updating land cover type but also including interannual variability of LAI and cover fraction.

We find a substantial impact of updating land and vegetation information from newly available Earth observations on the simulated surface latent heat flux and near-surface and deep soil moisture. However, these modifications do not always show positive impacts on the model performance. The changes in model performance vary between regions and considered variables, indicating the need for model evaluation based on multivariable analysis to make conclusive remarks on model performance (McCabe et al., 2005). Further, this shows that ingesting novel Earth observation data streams into current LSMs is not automatically leading to improved model performance as the model parameterizations need to be adapted to these updates (Nogueira et al., 2021). By considering several reference data sets, we benefit from the growing suite of global observational products, and manage to incorporate the uncertainty between these products into our evaluation of model performance.

As a further step we also recalibrate the model to adapt it to the new conditions. For the model recalibration we follow two approaches: global calibration and regional calibration (Xie et al., 2007). We find that the regional calibration yields substantial better agreement between model simulations and reference data sets, suggesting it may be beneficial to revise the spatial variability of model parameters which so far is based on soil and vegetation types that is, lookup tables. An update of those lookup tables and/or the consideration of more aspects of spatial heterogeneity may be a way forward in this context. This would allow that future calibrations can be done globally only.

We suggest that one reason for the lack of improvement in the model performance after updating land surface information with state-of-the-art observations is attributed to the fact that the previously calibrated parameter values are then inconsistent with the new vegetation data sets in the model. The model shows substantial improvement when adjusting parameters, particularly through the regional calibration, indicating that land information updates in the model cannot be treated independently from model parameterization. Future work should consider the impact of the improved and calibrated ECLand performance within a coupled model system.

Data Availability Statement

The meteorological forcing for ECLand from ERA5 is available at https://cds.climate.copernicus.eu/ (ECMWF & Service, 2018). The EVI data from MODIS are available through NASA's data catalogue at https://lpdaac.usgs.gov/products/mod13c1v006/ (EOSDIS, 2015). Both the evaporative fraction data from FLUXCOM and the soil moisture data from SoMo.ml are available at the Data Portal of the Max Planck Institute for Biogeochemistry at https://www.bgc-jena.mpg.de/geodb/projects/Data.php (Max Planck Institute for Biogeochemistry, 2019, 2021). The output data from the ECLand modelling experiments are available in the Zenodo repository at https://doi.org/10.5281/zenodo.7823893.

References

Araya, S. N., & Ghezzehei, T. A. (2019). Using machine learning for prediction of saturated hydraulic conductivity and its sensitivity to soil structural perturbations. *Water Resources Research*, *55*(7), 5715–5737. https://doi.org/10.1029/2018WR024357

Balsamo, G., Agusti-Panareda, A., Albergel, C., Arduini, G., Beljaars, A., Bidlot, J., et al. (2018). Satellite and in situ observations for advancing global Earth surface modelling: A review. *Remote Sensing*, *10*(12), 2038. https://doi.org/10.3390/rs10122038

Balsamo, G., Pappenberger, F., Dutra, E., Viterbo, P., & vanden Hurk, B. (2011). A revised land hydrology in the ECMWF model: A step towards daily water flux prediction in a fully-closed water cycle. *Hydrological Processes*, *25*(7), 1046–1054. https://doi.org/10.1002/hyp.7808

Bontemps, S., Radoux, J., de Maet, T., Lamarche, C., Moreau, I., Vittek, M., & Defourny, P. (2017). Land cover CCI algorithm theoretical basis document part III: Classification year 2-1.2 (Tech. Rep. No. 1.2). *ESA Climate Change Initiative*. Retrieved from https://www.esa-landcover-cci.org/?q=webfm_send/139

Boussetta, S., Balsamo, G., Arduini, G., Dutra, E., McNorton, J., Choulga, M., et al. (2021). ECLand: The ECMWF land surface modelling system. *Atmosphere*, *12*(6), 723. https://doi.org/10.3390/atmos12060723

Boussetta, S., Balsamo, G., Beljaars, A., Kral, T., & Jarlan, L. (2013). Impact of a satellite-derived leaf area index monthly climatology in a global numerical weather prediction model. *International Journal of Remote Sensing*, *34*(9–10), 3520–3542. https://doi.org/10.1080/01431161.2012.716543

Boussetta, S., Balsamo, G., Dutra, E., Beljaars, A., & Albergel, C. (2015). Assimilation of surface albedo and vegetation states from satellite observations and their impact on numerical weather prediction. *Remote Sensing of Environment*, *163*, 111–126. https://doi.org/10.1016/j.rse.2015.03.009

Breiman, L. (2001). Random forests. *Machine Learning*, *45*(1), 5–32. https://doi.org/10.1023/A:1010933404324

Chaney, N. W., Van Huijgevoort, M. H. J., Shevliakova, E., Malyshev, S., Milly, P. C. D., Gauthier, P. P. G., & Sulman, B. N. (2018). Harnessing big data to rethink land heterogeneity in Earth system models. *Hydrology and Earth System Sciences*, *22*(6), 3311–3330. https://doi.org/10.5194/hess-22-3311-2018

Cockburn, W., Ting, I. P., & Sternberg, L. O. (1979). Relationships between stomatal behavior and internal carbon dioxide concentration in crassulacean acid metabolism plants. *Plant Physiology*, *63*(6), 1029–1032. https://doi.org/10.1104/pp.63.6.1029

Coutadeur, C., Coquet, Y., & Roger-Estrade, J. (2002). Variation of hydraulic conductivity in a tilled soil. *European Journal of Soil Science*, *53*(4), 619–628. https://doi.org/10.1046/j.1365-2389.2002.00473.x

Denissen, J. M., Teuling, A. J., Reichstein, M., & Orth, R. (2020). Critical soil moisture derived from satellite observations over Europe. *Journal of Geophysical Research: Atmospheres*, *125*(6), e2019JD031672. https://doi.org/10.1029/2019JD031672

de Rosnay, P., Drusch, M., Vasiljevic, D., Balsamo, G., Albergel, C., & Isaksen, L. (2013). A simplified extended Kalman filter for the global operational soil moisture analysis at ECMWF. *Quarterly Journal of the Royal Meteorological Society*, *139*(674), 1199–1213. https://doi.org/10.1002/qj.2023

Didan, K. (2015). MOD13Q1 MODIS/Terra vegetation indices 16-day L3 global 250m SIN grid v006. *NASA EOSDIS Land Processes DAAC*. https://doi.org/10.5067/MODIS/MOD13Q1.006

Dirmeyer, P. A., & Halder, S. (2017). Application of the land–atmosphere coupling paradigm to the operational coupled forecast system, version 2 (CFSv2). *Journal of Hydrometeorology*, *18*(1), 85–108. https://doi.org/10.1175/JHM-D-16-0064.1

Dirmeyer, P. A., Halder, S., & Bombardi, R. (2018). On the harvest of predictability from land states in a global forecast model. *Journal of Geophysical Research: Atmospheres*, *123*(23), 13111–13127. https://doi.org/10.1029/2018JD029103

Diro, G. T., Sushama, L., & Huziy, O. (2018). Snow-Atmosphere coupling and its impact on temperature variability and extremes over North America. *Climate Dynamics*, *50*(7), 2993–3007. https://doi.org/10.1007/s00382-017-3788-5

Dutra, E., Balsamo, G., Viterbo, P., Miranda, P. M. A., Beljaars, A., Schär, C., & Elder, K. (2010). An improved snow scheme for the ECMWF land surface model: Description and offline validation. *Journal of Hydrometeorology*, *11*(4), 899–916. https://doi.org/10.1175/2010JHM1249.1

Dutra, E., Kotlarski, S., Viterbo, P., Balsamo, G., Miranda, P. M. A., Schär, C., et al. (2011). Snow cover sensitivity to horizontal resolution, parameterizations, and atmospheric forcing in a land surface model. *Journal of Geophysical Research*, *116*(D21), D21109. https://doi.org/10.1029/2011JD016061

Duveiller, G., Pickering, M., Muñoz Sabater, J., Caporaso, L., Boussetta, S., Balsamo, G., & Cescatti, A. (2022). Getting the leaves right matters for estimating temperature extremes. *Geoscientific model development discussions, 2022* (pp. 1–26). https://doi.org/10.5194/gmd-2022-216

ECMWF, & Service, C. C. C. (2018). ERA5 hourly data on single levels from 1959 to present. Retrieved from https://cds.climate.copernicus.eu/

EOSDIS, N. (2015). NASA EOSDIS land processes DAAC: Mod13c1 MODIS/TERRA vegetation indices 16-day L3 global 0.05deg CMG v006. Retrieved from https://lpdaac.usgs.gov/products/mod13c1v006/

Fairbairn, D., de Rosnay, P., & Browne, P. A. (2019). The new stand-alone surface analysis at ecmwf: Implications for land–atmosphere da coupling. *Journal of Hydrometeorology*, *20*(10), 2023–2042. https://doi.org/10.1175/JHM-D-19-0074.1

Fisher, R. A., & Koven, C. D. (2020). Perspectives on the future of land surface models and the challenges of representing complex terrestrial systems. *Journal of Advances in Modeling Earth Systems*, *12*(4), e2018MS001453. https://doi.org/10.1029/2018MS001453

Fridley, B. L., & Dixon, P. (2007). Data augmentation for a Bayesian spatial model involving censored observations. *Environmetrics*, *18*(2), 107–123. https://doi.org/10.1002/env.806

Gelaro, R., McCarty, W., Suárez, M. J., Todling, R., Molod, A., Takacs, L., et al. (2017). The modern-era retrospective analysis for research and applications, version 2 (MERRA-2). *Journal of Climate*, *30*(14), 5419–5454. https://doi.org/10.1175/JCLI-D-16-0758.1

Ghilain, N., Arboleda, A., Sepulcre-Cantò, G., Batelaan, O., Ardö, J., & Gellens-Meulenberghs, F. (2012). Improving evapotranspiration in a land surface model using biophysical variables derived from MSG/SEVIRI satellite. *Hydrology and Earth System Sciences*, *16*(8), 2567–2583. https://doi.org/10.5194/hess-16-2567-2012

Ghosh, A., & Maiti, R. (2021). Soil erosion susceptibility assessment using logistic regression, decision tree and random forest: Study on the Mayurakshi river basin of eastern India. *Environmental Earth Sciences*, *80*(8), 328. https://doi.org/10.1007/s12665-021-09631-5

González-Rouco, J. F., Steinert, N. J., García-Bustamante, E., Hagemann, S., de Vrese, P., Jungclaus, J. H., et al. (2021). Increasing the depth of a land surface model. Part I: Impacts on the subsurface thermal regime and energy storage. *Journal of Hydrometeorology*, *22*(12), 3211–3230. https://doi.org/10.1175/JHM-D-21-0024.1

Guillevic, P., Koster, R. D., Suarez, M. J., Bounoua, L., Collatz, G. J., Los, S. O., & Mahanama, S. P. P. (2002). Influence of the interannual variability of vegetation on the surface energy balance—A global sensitivity study. *Journal of Hydrometeorology*, *3*(6), 617–629. https://doi.org/10.1175/1525-7541(2002)003⟨0617:IOTIVO⟩2.0.CO;2

Guo, Z., Dirmeyer, P. A., & DelSole, T. (2011). Land surface impacts on subseasonal and seasonal predictability. *Geophysical Research Letters*, *38*(24), L24812. https://doi.org/10.1029/2011GL049945

Gupta, H. V., Bastidas, L. A., Sorooshian, S., Shuttleworth, W. J., & Yang, Z. L. (1999). Parameter estimation of a land surface scheme using multicriteria methods. *Journal of Geophysical Research*, *104*(D16), 19491–19503. https://doi.org/10.1029/1999JD900154

Hardy, J., Melloh, R., Koenig, G., Marks, D., Winstral, A., Pomeroy, J., & Link, T. (2004). Solar radiation transmission through conifer canopies. *Agricultural and Forest Meteorology*, *126*(3), 257–270. https://doi.org/10.1016/j.agrformet.2004.06.012

Hawkins, L. R., Rupp, D. E., McNeall, D. J., Li, S., Betts, R. A., Mote, P. W., et al. (2019). Parametric sensitivity of vegetation dynamics in the TRIFFID model and the associated uncertainty in projected climate change impacts on Western U.S. forests. *Journal of Advances in Modeling Earth Systems*, *11*(8), 2787–2813. https://doi.org/10.1029/2018MS001577

He, Z., Yang, Y., Fang, R., Zhou, S., Zhao, W., Bai, Y., et al. (2023). Integration of Shapley additive explanations with random forest model for quantitative precipitation estimation of mesoscale convective systems. *Frontiers in Environmental Science*, *10*. https://doi.org/10.3389/fenvs.2022.1057081

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., et al. (2020). The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, *146*(730), 1999–2049. https://doi.org/10.1002/qj.3803

Hirsch, A. L., Evans, J. P., Di Virgilio, G., Perkins-Kirkpatrick, S. E., Argüeso, D., Pitman, A. J., et al. (2019). Amplification of Australian heatwaves via local Land-Atmosphere coupling. *Journal of Geophysical Research: Atmospheres*, *124*(24), 13625–13647. https://doi.org/10.1029/2019JD030665

Jefferson, J. L., Maxwell, R. M., & Constantine, P. G. (2017). Exploring the sensitivity of photosynthesis and stomatal resistance parameters in a land surface model. *Journal of Hydrometeorology*, *18*(3), 897–915. https://doi.org/10.1175/JHM-D-16-0053.1

Ji, P., Yuan, X., Shi, C., Jiang, L., Wang, G., & Yang, K. (2023). A long-term simulation of land surface conditions at high resolution over continental China. *Journal of Hydrometeorology*, *24*(2), 285–314. https://doi.org/10.1175/JHM-D-22-0135.1

Johannsen, F., Ermida, S., Martins, J. P. A., Trigo, I. F., Nogueira, M., & Dutra, E. (2019). Cold bias of ERA5 summertime daily maximum land surface temperature over Iberian peninsula. *Remote Sensing*, *11*(21), 2570. https://doi.org/10.3390/rs11212570

Jung, M., Koirala, S., Weber, U., Ichii, K., Gans, F., Camps-Valls, G., et al. (2019). The FLUXCOM ensemble of global Land-Atmosphere energy fluxes. *Scientific Data*, *6*(1), 74. https://doi.org/10.1038/s41597-019-0076-8

Kardel, F., Wuyts, K., Babanezhad, M., Vitharana, U., Wuytack, T., Potters, G., & Samson, R. (2010). Assessing urban habitat quality based on specific leaf area and stomatal characteristics of Plantago lanceolata L. *Environmental Pollution*, *158*(3), 788–794. https://doi.org/10.1016/j.envpol.2009.10.006

Koster, R. D., Schubert, S. D., DeAngelis, A. M., Molod, A. M., & Mahanama, S. P. (2020). Using a simple water balance framework to quantify the impact of soil moisture initialization on subseasonal evapotranspiration and air temperature forecasts. *Journal of Hydrometeorology*, *21*(8), 1705–1722. https://doi.org/10.1175/JHM-D-20-0007.1

Laguë, M. M., Bonan, G. B., & Swann, A. L. S. (2019). Separating the impact of individual land surface properties on the terrestrial surface energy budget in both the coupled and uncoupled Land–Atmosphere System. *Journal of Climate*, *32*(18), 5725–5744. https://doi.org/10.1175/JCLI-D-18-0812.1

Li, L., Bisht, G., Hao, D., & Leung, L. R. (2023). Global 1km land surface parameters for kilometer scale Earth system modeling. *Earth system science data discussions, 2023* (pp. 1–47). https://doi.org/10.5194/essd-2023-242

Li, L., Bisht, G., & Leung, L. R. (2022). Spatial heterogeneity effects on land surface modeling of water and energy partitioning. *Geoscientific Model Development*, *15*(14), 5489–5510. https://doi.org/10.5194/gmd-15-5489-2022

Li, W., Migliavacca, M., Forkel, M., Walther, S., Reichstein, M., & Orth, R. (2021). Revisiting global vegetation controls using multi-layer soil moisture. *Geophysical Research Letters*, *48*(11), e2021GL092856. https://doi.org/10.1029/2021GL092856

Loveland, T. R., Reed, B. C., Brown, J. F., Ohlen, D. O., Zhu, Z., Yang, L., & Merchant, J. W. (2000). Development of a global land cover characteristics database and IGBP DISCover from 1 km AVHRR data. *International Journal of Remote Sensing*, *21*(6–7), 1303–1330. https://doi.org/10.1080/014311600210191

Lundberg, S., & Lee, S.-I. (2017). A unified approach to interpreting model predictions.

MacLeod, D. A., Cloke, H. L., Pappenberger, F., & Weisheimer, A. (2016). Improved seasonal prediction of the hot summer of 2003 over Europe through better representation of uncertainty in the land surface. *Quarterly Journal of the Royal Meteorological Society*, *142*(694), 79–90. https://doi.org/10.1002/qj.2631

Martens, B., Miralles, D. G., Lievens, H., vander Schalie, R., de Jeu, R. A. M., Fernández-Prieto, D., et al. (2017). GLEAM v3: Satellite-based land evaporation and root-zone soil moisture. *Geoscientific Model Development*, *10*(5), 1903–1925. https://doi.org/10.5194/gmd-10-1903-2017

Martínez-Muñoz, G., & Suárez, A. (2010). Out-of-bag estimation of the optimal sample size in bagging. *Pattern Recognition*, *43*(1), 143–152. https://doi.org/10.1016/j.patcog.2009.05.010

Max Planck Institute for Biogeochemistry. (2019). Fluxcom [Dataset]. BGC-JENA. Retrieved from https://www.bgc-jena.mpg.de/geodb/projects/Data.php

Max Planck Institute for Biogeochemistry. (2021). Somo.ml [Dataset]. BGC-JENA. Retrieved from https://www.bgc-jena.mpg.de/geodb/projects/Data.php

McCabe, M., Franks, S., & Kalma, J. (2005). Calibration of a land surface model using multiple data sets. *Journal of Hydrology*, *302*(1), 209–222. https://doi.org/10.1016/j.jhydrol.2004.07.002

McKay, M. D., Beckman, R. J., & Conover, W. J. (1979). A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics*, *21*(2), 239–245. https://doi.org/10.2307/1268522

Medici, G., & West, L. J. (2021). Groundwater flow velocities in karst aquifers; importance of spatial observation scale and hydraulic testing for contaminant transport prediction. *Environmental Science and Pollution Research*, *28*(32), 43050–43063. https://doi.org/10.1007/s11356-021-14840-3

Mehrez, M., Taconet, O., Vidal-Madjar, D., & Valencogne, C. (1992). Estimation of stomatal resistance and canopy evaporation during the HAPEX-MOBILHY experiment. *Agricultural and Forest Meteorology*, *58*(3), 285–313. https://doi.org/10.1016/0168-1923(92)90066-D

Meng, X. H., Evans, J. P., & McCabe, M. F. (2014). The impact of observed vegetation changes on Land–Atmosphere feedbacks during drought. *Journal of Hydrometeorology*, *15*(2), 759–776. https://doi.org/10.1175/JHM-D-13-0130.1

Miralles, D. G., Gentine, P., Seneviratne, S. I., & Teuling, A. J. (2019). Land-atmospheric feedbacks during droughts and heatwaves: State of the science and current challenges. *Annals of the New York Academy of Sciences*, *1436*(1), 19–35. https://doi.org/10.1111/nyas.13912

Molnar, C. (2020). *Interpretable machine learning*. Lulu.com. Retrieved from https://christophm.github.io/interpretable-ml-book/

Monteith, J. L., Szeicz, G., & Waggoner, P. E. (1965). The measurement and control of stomatal resistance in the field. *Journal of Applied Ecology*, *2*(2), 345–355. https://doi.org/10.2307/2401484

Montzka, C., Herbst, M., Weihermüller, L., Verhoef, A., & Vereecken, H. (2017). A global data set of soil hydraulic properties and sub-grid variability of soil water retention and hydraulic conductivity curves. *Earth System Science Data*, *9*(2), 529–543. https://doi.org/10.5194/essd-9-529-2017

Nemunaitis-Berry, K. L., Klein, P. M., Basara, J. B., & Fedorovich, E. (2017). Sensitivity of predictions of the urban surface energy balance and heat island to variations of urban canopy parameters in simulations with the WRF model. *Journal of Applied Meteorology and Climatology*, *56*(3), 573–595. https://doi.org/10.1175/JAMC-D-16-0157.1

Niyogi, D. (2019). Land surface processes. In D. A. Randall, J. Srinivasan, R. S. Nanjundiah, & P. Mukhopadhyay (Eds.), *Current trends in the representation of physical processes in weather and climate models* (pp. 349–370). Springer Singapore. https://doi.org/10.1007/978-981-13-3396-5_17

Niyogi, D. S., & Raman, S. (1997). Comparison of four different stomatal resistance schemes using FIFE observations. *Journal of Applied Meteorology*, *36*(7), 903–917. https://doi.org/10.1175/1520-0450(1997)036⟨0903:COFDSR⟩2.0.CO;2

Niyogi, D. S., Raman, S., Prabhu, A., Kumar, U., & Joshi, S. S. (1997). Direct estimation of stomatal resistance for meteorological applications. *Geophysical Research Letters*, *24*(14), 1771–1774. https://doi.org/10.1029/97GL01790

Nogueira, M., Albergel, C., Boussetta, S., Johannsen, F., Trigo, I. F., Ermida, S. L., et al. (2020). Role of vegetation in representing land surface temperature in the CHTESSEL (CY45R1) and SURFEX-ISBA (v8.1) land surface models: A case study over Iberia. *Geoscientific Model Development*, *13*(9), 3975–3993. https://doi.org/10.5194/gmd-13-3975-2020

Nogueira, M., Boussetta, S., Balsamo, G., Albergel, C., Trigo, I. F., Johannsen, F., et al. (2021). Upgrading land-cover and vegetation seasonality in the ECMWF coupled system: Verification with FLUXNET sites, METEOSAT satellite land surface temperatures, and ERA5 atmospheric reanalysis. *Journal of Geophysical Research: Atmospheres*, *126*(15), e2020JD034163. https://doi.org/10.1029/2020JD034163

Orth, R., Dutra, E., & Pappenberger, F. (2016). Improving weather predictability by including land surface model parameter uncertainty. *Monthly Weather Review*, *144*(4), 1551–1569. https://doi.org/10.1175/MWR-D-15-0283.1

Orth, R., Dutra, E., Trigo, I. F., & Balsamo, G. (2017). Advancing land surface model development with satellite-based Earth observations. *Hydrology and Earth System Sciences*, *21*(5), 2483–2495. https://doi.org/10.5194/hess-21-2483-2017

O, S., Dutra, E., & Orth, R. (2020). Robustness of process-based versus data-driven modeling in changing climatic conditions. *Journal of Hydrometeorology*, *21*(9), 1929–1944. https://doi.org/10.1175/JHM-D-20-0072.1

O, S., & Orth, R. (2021). Global soil moisture data derived through machine learning trained with *in-situ* measurements. *Scientific Data*, *8*(1), 170. https://doi.org/10.1038/s41597-021-00964-1

Probst, P., Wright, M. N., & Boulesteix, A.-L. (2019). Hyperparameters and tuning strategies for random forest. *WIREs Data Mining and Knowledge Discovery*, *9*(3). https://doi.org/10.1002/widm.1301

Quillet, A., Peng, C., & Garneau, M. (2010). Toward dynamic global vegetation models for simulating vegetation–climate interactions and feedbacks: Recent developments, limitations, and future challenges. *Environmental Reviews*, *18*(NA), 333–353. https://doi.org/10.1139/A10-016

Rayner, P. J., Scholze, M., Knorr, W., Kaminski, T., Giering, R., & Widmann, H. (2005). Two decades of terrestrial carbon fluxes from a carbon cycle data assimilation system (CCDAS). *Global Biogeochemical Cycles*, *19*(2), GB2026. https://doi.org/10.1029/2004GB002254

Santanello, J. A., Dirmeyer, P. A., Ferguson, C. R., Findell, K. L., Tawfik, A. B., Berg, A., et al. (2018). Land–atmosphere interactions: The LoCo perspective. *Bulletin of the American Meteorological Society*, *99*(6), 1253–1272. https://doi.org/10.1175/BAMS-D-17-0001.1

Santanello, J. A., Jr., Lawston, P., Kumar, S., & Dennis, E. (2019). Understanding the impacts of soil moisture initial conditions on NWP in the context of land–atmosphere coupling. *Journal of Hydrometeorology*, *20*(5), 793–819. https://doi.org/10.1175/JHM-D-18-0186.1

Santanello, J. A., Kumar, S. V., Peters-Lidard, C. D., Harrison, K., & Zhou, S. (2013). Impact of land model calibration on coupled land–atmosphere prediction. *Journal of Hydrometeorology*, *14*(5), 1373–1400. https://doi.org/10.1175/JHM-D-12-0127.1

Santanello, J. A., Peters-Lidard, C. D., Kumar, S. V., Alonge, C., & Tao, W.-K. (2009). A modeling and observational framework for diagnosing local Land–Atmosphere coupling on diurnal time scales. *Journal of Hydrometeorology*, *10*(3), 577–599. https://doi.org/10.1175/2009JHM1066.1

Santaren, D., Peylin, P., Viovy, N., & Ciais, P. (2007). Optimizing a process-based ecosystem model with eddy-covariance flux measurements: A pine forest in southern France. *Global Biogeochemical Cycles*, *21*(2), GB2013. https://doi.org/10.1029/2006GB002834

Schratz, P., Muenchow, J., Iturritxa, E., Richter, J., & Brenning, A. (2019). Hyperparameter tuning and performance assessment of statistical and machine-learning algorithms using spatial data. *Ecological Modelling*, *406*, 109–120. https://doi.org/10.1016/j.ecolmodel.2019.06.002

Seneviratne, S. I., Corti, T., Davin, E. L., Hirschi, M., Jaeger, E. B., Lehner, I., et al. (2010). Investigating soil moisture–climate interactions in a changing climate: A review. *Earth-Science Reviews*, *99*(3), 125–161. https://doi.org/10.1016/j.earscirev.2010.02.004

Sihag, P., Mohsenzadeh-Karimi, S., & Angelaki, A. (2019). Random forest, M5P and regression analysis to estimate the field unsaturated hydraulic conductivity. *Applied Water Science*, *9*(5), 129. https://doi.org/10.1007/s13201-019-1007-8

Silva, S. J., Keller, C. A., & Hardin, J. (2022). Using an explainable machine learning approach to characterize Earth System Model errors: Application of SHAP analysis to modeling lightning flash occurrence. *Journal of Advances in Modeling Earth Systems*, *14*(4), e2021MS002881. https://doi.org/10.1029/2021MS002881

Stegehuis, A. I., Teuling, A. J., Ciais, P., Vautard, R., & Jung, M. (2013). Future European temperature change uncertainties reduced by using land heat flux observations. *Geophysical Research Letters*, *40*(10), 2242–2245. https://doi.org/10.1002/grl.50404

Steinert, N. J., González-Rouco, J. F., de Vrese, P., García-Bustamante, E., Hagemann, S., Melo-Aguilar, C., et al. (2021). Increasing the depth of a land surface model. Part II: Temperature sensitivity to improved subsurface thermodynamics and associated permafrost response. *Journal of Hydrometeorology*, *22*(12), 3231–3254. https://doi.org/10.1175/JHM-D-21-0023.1

Stevens, D., Miranda, P. M. A., Orth, R., Boussetta, S., Balsamo, G., & Dutra, E. (2020). Sensitivity of surface fluxes in the ECMWF land surface model to the remotely sensed leaf area index and root distribution: Evaluation with tower flux data. *Atmosphere*, *11*(12), 1362. https://doi.org/10.3390/atmos11121362

Sundararajan, M., & Najmi, A. (2020). The many Shapley values for model explanation.

Torres-Rojas, L., Vergopolan, N., Herman, J. D., & Chaney, N. W. (2022). Towards an optimal representation of sub-grid heterogeneity in land surface models. *Water Resources Research*, *58*(12), e2022WR032233. https://doi.org/10.1029/2022WR032233

Turner, N. C. (1974). Stomatal behavior and water status of Maize, Sorghum, and Tobacco under field conditions: II. At low soil water potential. *Plant Physiology*, *53*(3), 360–365. https://doi.org/10.1104/pp.53.3.360

Verbist, K., Baetens, J., Cornelis, W., Gabriels, D., Torres, C., & Soto, G. (2009). Hydraulic conductivity as influenced by stoniness in degraded drylands of Chile. *Soil Science Society of America Journal*, *73*(2), 471–484. https://doi.org/10.2136/sssaj2008.0066

Verger, A., Descals, A., Lacaze, R., & Cherlet, M. (2022). *Copernicus global land operations "vegetation and energy". (Tech. Rep. No. l1.10)*. Copernicus. Retrieved from https://land.copernicus.eu/global/sites/cgls.vito.be/files/products/CGLOPS1_ATBD_LAI300m-V1.1_I1.10.pdf

Verger, A., Weiss, M., Baret, F., Pacholczyk, P., & Leroy, M. (2020). *Algorithm theoretical basis document (Tech. Rep. No. l2.50)*. Théia and CREAF. Retrieved from https://www.theia-land.fr/wp-content/uploads/2022/03/THEIA-SP-44-0207-CREAF_I2.50-1.pdf

Wadoux, A. M. J.-C., Saby, N. P. A., & Martin, M. P. (2023). Shapley values reveal the drivers of soil organic carbon stock prediction. *SOIL*, *9*(1), 21–38. https://doi.org/10.5194/soil-9-21-2023

Wilks, D. (2011). Chapter 3 - Empirical distributions and exploratory data analysis. In D. S. Wilks (Ed.), *Statistical methods in the atmospheric sciences* (Vol. 100, pp. 23–70). Academic Press. https://doi.org/10.1016/B978-0-12-385022-5.00003-8

Williams, M., Richardson, A. D., Reichstein, M., Stoy, P. C., Peylin, P., Verbeeck, H., et al. (2009). Improving land surface models with FLUX-NET data. *Biogeosciences*, *6*(7), 1341–1359. https://doi.org/10.5194/bg-6-1341-2009

Wipfler, E. L., Metselaar, K., van Dam, J. C., Feddes, R. A., van Meijgaard, E., van Ulft, L. H., et al. (2011). Seasonal evaluation of the land surface scheme HTESSEL against remote sensing derived energy fluxes of the Transdanubian region in Hungary. *Hydrology and Earth System Sciences*, *15*(4), 1257–1271. https://doi.org/10.5194/hess-15-1257-2011

Wulfmeyer, V., Turner, D. D., Baker, B., Banta, R., Behrendt, A., Bonin, T., et al. (2018). A new research approach for observing and characterizing Land–Atmosphere feedback. *Bulletin of the American Meteorological Society*, *99*(8), 1639–1667. https://doi.org/10.1175/BAMS-D-17-0009.1

Xie, Z., Yuan, F., Duan, Q., Zheng, J., Liang, M., & Chen, F. (2007). Regional parameter estimation of the VIC land surface model: Methodology and application to river basins in China. *Journal of Hydrometeorology*, *8*(3), 447–468. https://doi.org/10.1175/JHM568.1

Zhang, J., Wang, W.-C., & Leung, L. R. (2008). Contribution of Land-Atmosphere coupling to summer climate variability over the contiguous United States. *Journal of Geophysical Research*, *113*(D22), D22109. https://doi.org/10.1029/2008JD010136