JGR Atmospheres

RESEARCH ARTICLE
10.1029/2020JD034163

Key Points:

- Daily maximum land surface temperature in European Centre for Medium-Range Weather Forecasting (ECMWF) model is affected by large biases associated with the representation of land cover
- An updated representation of vegetation in the ECMWF model based on state-of-the-art Earth observations reduces this bias
- The implemented changes have an overall global positive impact, but a re-calibration of related model parameters is required

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

M. Nogueira,
mdnogueira@fc.ul.pt

Citation:

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*, e2020JD034163. https://doi.org/10.1029/2020JD034163

Received 30 OCT 2020
Accepted 16 JUL 2021

Author Contributions:

Conceptualization: Miguel Nogueira, Souhail Boussetta, Gianpaolo Balsamo, Emanuel Dutra
Data curation: Isabel F. Trigo, Diego G. Miralles
Formal analysis: Miguel Nogueira, Frederico Johannsen, Emanuel Dutra
Funding acquisition: Emanuel Dutra

Upgrading Land-Cover and Vegetation Seasonality in the ECMWF Coupled System: Verification With FLUXNET Sites, METEOSAT Satellite Land Surface Temperatures, and ERA5 Atmospheric Reanalysis

Miguel Nogueira[1], Souhail Boussetta[2], Gianpaolo Balsamo[2], Clément Albergel[3], Isabel F. Trigo[1,4], Frederico Johannsen[1], Diego G. Miralles[5], and Emanuel Dutra[1,4]

[1]Instituto Dom Luiz, IDL, Faculty of Sciences, University of Lisbon, Lisbon, Portugal, [2]ECMWF, Reading, UK, [3]European Space Agency Climate Office, ECSAT, Didcot, UK, [4]Instituto Português do Mar e da Atmosfera, Lisbon, Portugal, [5]Hydro-Climate Extremes Lab (H-CEL)—Ghent University, Ghent, Belgium

Abstract In this study, we show that limitations in the representation of land cover and vegetation seasonality in the European Centre for Medium-Range Weather Forecasting (ECMWF) model are partially responsible for large biases (up to ~10°C, either positive or negative depending on the region) on the simulated daily maximum land surface temperature (LST) with respect to satellite Earth Observations (EOs) products from the Land Surface Analysis Satellite Application Facility. The error patterns were coherent in offline land-surface and coupled land-atmosphere simulations, and in ECMWF's latest generation reanalysis (ERA5). Subsequently, we updated the ECMWF model's land cover characterization leveraging on state-of-the-art EOs—the European Space Agency Climate Change Initiative land cover data set and the Copernicus Global Land Services leaf area index. Additionally, we tested a clumping parameterization, introducing seasonality to the effective low vegetation coverage. The updates reduced the overall daily maximum LST bias and unbiased root-mean-squared errors. In contrast, the implemented updates had a neutral impact on daily minimum LST. Our results also highlighted the complex regional heterogeneities in the atmospheric sensitivity to land cover and vegetation changes, particularly with issues emerging over eastern Brazil and northeastern Asia. These issues called for a re-calibration of model parameters (e.g., minimum stomatal resistance, roughness length, rooting depth), along with a revision of several model assumptions (e.g., snow shading by high vegetation).

1. Introduction

The land surface plays an important role in the Earth climate system, controlling the energy, water and carbon storage, and fluxes over a wide range of temporal and spatial scales (Bonan, 2008; Pitman et al., 2009; Seneviratne et al., 2010). Correspondingly, land surface models (LSMs) are an essential part of numerical weather prediction (NWP) and climate models (Delworth & Manaba, 1993; Dirmeyer, 2003; Koster et al., 2004; Shukla & Mintz, 1982). Over the past few years, sensitivity studies have highlighted the importance of an accurate description of land surface processes for the realistic simulation of multiple relevant atmospheric climate variables, such as temperature, wind, and precipitation (Beljaars et al., 1996; Hurk et al., 2010; Koster et al., 2004; Wang & Kumar, 1998). However, state-of-the-art LSMs are still not able to accurately represent the full range of relevant processes affecting the land surface, resulting in well-known uncertainties in the partitioning of surface turbulent heat fluxes, and in water and carbon storage, flow, and exchanges (Best et al., 2015; Gevaert et al., 2018; Nogueira, Albergel, et al., 2020; Orth et al., 2017; Trigo et al., 2015).

Significant efforts have been devoted to enhancing the representation of the different components in LSMs through the integration of Earth Observations (EOs). The most straightforward way to achieve such an integration involves verification and optimization methods that improve parameterization schemes to reduce model-data discrepancies at the surface. For example, EO has been used to calibrate model parameters representing relevant land surface properties, such as land cover fraction or soil properties (e.g., Balsamo et al., 2018; Johannsen et al., 2019; Nogueira, Albergel, et al., 2020; Orth et al., 2017; Trigo et al., 2015). Remote sensing observations are particularly useful in this context, given their ability to provide estimates

Investigation: Miguel Nogueira
Methodology: Miguel Nogueira, Clément Albergel, Isabel F. Trigo, Emanuel Dutra
Project Administration: Emanuel Dutra
Writing – original draft: Miguel Nogueira
Writing – review & editing: Miguel Nogueira, Souhail Boussetta, Gianpaolo Balsamo, Clément Albergel, Isabel F. Trigo, Diego G. Miralles, Emanuel Dutra

for multiple land surface variables with high spatial resolution and coverage. Data assimilation techniques provide an alternative path to improve the forecasting accuracy at short to sub-seasonal scales (e.g., Albergel et al., 2019; Boussetta et al., 2015; de Rosnay et al., 2013; Ghent et al., 2010; Massari et al., 2018; Pipunic et al., 2008; Schlosser & Dirmeyer, 2001).

The recent work by Johannsen et al. (2019) (henceforth JO19) and Nogueira, Albergel, et al. (2020) (henceforth NO20) provides good examples of LSM validation and development leveraging on the last generation EO products. These studies showed that the European Centre for Medium-Range Weather Forecasting (ECMWF) LSM, that is, Carbon Hydrology Tiled ECMWF Scheme for Surface Exchanges over Land (CHTESSEL; Balsamo et al., 2009; Boussetta, Balsamo, Beljaars, Panareda, et al., 2013; Dutra et al., 2010; van den Hurk et al., 2000), misrepresents vegetation coverage and its seasonality over Iberia with respect to the Copernicus Global Land Service (CGLS) leaf area index (LAI) and fraction of green vegetation coverage. This was shown to result in large systematic errors in daily maximum land surface temperature (LST), reaching magnitudes around 10°C when compared against the LST products from Satellite Application Facility on Land Surface Analysis (LSA-SAF). Additionally, JO19 and NO20 showed that the misrepresentation of vegetation coverage over Iberia and the resulting systematic errors in daily maximum LST were also present in the fifth (and latest) generation ECMWF reanalysis (ERA5). More generally, the importance of the representation of Land Use and Land Cover (LULC) changes in climate models has been addressed by several previous research efforts. For example, the LUCID (Land-Use and Climate, Identification of Robust Impacts) model intercomparison project reported a cooling effect over North America and Eurasia in general circulation models (GCMs) associated with LULC changes of similar magnitude to the warming induced by greenhouse gas emissions since the preindustrial period (de Noblet-Ducoudré et al., 2012). The results from LUCID also evidenced significant uncertainties in land-use change impacts amongst the different GCMs, particularly for the simulated regional response (de Noblet-Ducoudré et al., 2012; Kumar et al., 2013; Pitman et al., 2009). The errors in the models' representation of the LULC effects on the lower troposphere have also been shown to limit the progress in weather and climate predictability (Guo et al., 2011, 2012; Orth et al., 2016), and in the accurate representation of extreme events (Hauser et al., 2016; Miralles et al., 2019; Nogueira & Soares, 2019; Nogueira, Lima, & Soares, 2020; Rasmijn et al., 2018; Stegehuis et al., 2015).

In the standard CHTESSEL formulation, the vegetation cover is derived from the Global Land Cover Characteristics data set (GLCC, Loveland et al., 2000), which does not benefit from the developments of vegetation data sets during the past 20 years. NO20 proposed an updated vegetation characterization for CHTESSEL, including a revision of the land cover (vegetation types and fractional coverage), based on the European Space Agency Climate Change Initiative (ESA-CCI) data set (Defourny et al., 2014), and a monthly LAI climatology based on the CGLS data set (Verger et al., 2014), implemented along with a clumping parameterization to introduce the effect of vegetation seasonality. This package resulted in an improved characterization of vegetation patterns and seasonality compared to the CHTESSEL standard formulation. The NO20 updates resulted in large changes to the patterns of high and low vegetation fractions, which affect the regional vegetation density, surface roughness length for momentum and heat, and rooting depth, amongst others. In turn, these changes had important impacts on the simulated partitioning of latent and sensible heat fluxes during summer, resulting in a large reduction of daily maximum LST errors. However, NO20 results were based only on uncoupled land surface simulations (without atmospheric feedbacks) and limited to the Iberian Peninsula.

The present study applied the updates to CHTESSEL vegetation by NO20 to the full globe. It was hypothesized that the better characterization of global LULC patterns may result in an overall improvement of the simulated LST. These potential improvements can have important implications for ECMWF weather forecasts and reanalysis products. This hypothesis was explored by extending the assessment of the LST simulated by offline CHTESSEL simulations to the full Meteosat Second Generation (MSG) disc, for which the LSA-SAF LST product was available. Additionally, LST error assessment was extended to ERA5 and to coupled land-atmosphere simulations generated using the ECMWF Integrated Forecast Systems (IFS), using CHTESSEL as LSM. Subsequently, the impact of the implemented updates on the simulated LST was assessed in uncoupled CHTESSEL and coupled IFS simulations. Additionally, a global sensitivity analysis was performed to evaluate the influence of the vegetation cover updates on the atmospheric temperature,

wind, and humidity simulated by IFS, considering ERA5 as a baseline. Furthermore, the impact on simulated surface latent and sensible heat fluxes was evaluated over 51 FLUXNET locations.

The present manuscript is organized as follows. Section 2 presents data and methods used in this study. Section 3 includes the evaluation of LST and surface latent and sensible heat fluxes simulated by CHTESSEL, respectively, against the LSA-SAF product and FLUXNET in situ observations (Pastorello et al., 2020). The added value of the updated vegetation formulation proposed in NO20 for the global simulations is also assessed in Section 3, along with an analysis of the impacts of the vegetation update on different surface and tropospheric variables, thus evaluating potential gains in coupled land-atmosphere simulations. The results are summarized and discussed in Section 4.

2. Data and Methods

2.1. EO and Reanalysis Products

ESA-CCI provides consistent maps of land cover, based on the 22 classes from the land cover classification system developed by the United Nations Food and Agriculture Organization. They are derived by combining remotely sensed surface reflectance and in situ observations (Defourny et al., 2014). This data set was available at 300 m spatial resolution, on an annual basis, from 1992 to 2015 from ESA-CCI; the years 2016–2018 were delivered by the Copernicus Climate Change Service (C3S). A constant land cover taking 2010 data was chosen, as the effects of land cover evolution are beyond the scope of this study. The LAI estimates were obtained from CGLS at 1 km resolution covering the entire globe. The product is derived from the SPOT-VEGETATION and PROBA-V satellite observations using the algorithm described by Verger et al. (2014).

The LSA-SAF LST estimates were available every 15 min since 2004 to present over land pixels within the MSG disk, comprising satellite zenith view angles between 0° and 80°, with a 3 km resolution at the nadir. They are derived from the outgoing thermal infrared radiation (TIR) measured at top-of-atmosphere by the Spinning Enhanced Visible and InfraRed Imager, employing a generalized "split-window" technique (Freitas et al., 2010). The TIR spectral band (8–13 μm) was particularly appropriate as it presents relatively weak atmospheric attenuation under clear sky conditions and includes the peak of the Earth's spectral radiance (Ermida et al., 2019; Li et al., 2013). On the negative side, LSA-SAF LST estimates derived from TIR were mostly limited to clear-sky observations, which poses a significant limitation to their coverage (e.g., Ermida et al., 2019; Li et al., 2013; Trigo et al., 2011). In this study, we opted to use LST from a single sensor/geostationary platform, ensuring as much as possible that the satellite LST product was of homogeneous quality over the period and area of study.

ERA5 reanalysis is based on IFS (cycle Cy41r2) and uses CHTESSEL as LSM (Hersbach et al., 2020). It assimilates a large number of observations, including surface pressure, soil moisture, near-surface temperature and humidity over land, near-surface wind over oceans, and upper-air wind, temperature and humidity. But it does not assimilate LST, which is relevant to the analysis presented in Section 3. ERA5 provided the global forcing fields, with 31 km horizontal resolution and hourly timestep, for the offline CHTESSEL experiments, along with the initial, and boundary conditions for the coupled atmosphere simulations. Additionally, the upper atmospheric fields in the coupled simulations were nudged (relaxed) toward ERA5 value (see Section 2.2.3 for details). Total cloud cover, LST and air temperature fields at different vertical levels were also extracted from the ERA5 data set, providing a baseline to our simulation experiments. Recent works have demonstrated the overall better accuracy of ERA5 compared to previous reanalyses in representing a wide variety of relevant climate variables, including rainfall, wind, radiation, LST, surface turbulent fluxes, and near-surface air temperature (e.g., Beck et al., 2019; Hersbach et al., 2020; JO19; Martens et al., 2020; Nogueira, 2020; Rivas & Stoffelen, 2019; Urraca et al., 2018).

2.2. Description of IFS and CHTESSEL

IFS is the spectral NWP model developed at ECMWF in cooperation with Météo-France. It is the cornerstone for the ECMWF's operational weather forecasts and reanalysis products, including ERA5. IFS includes a set of physical sub-models of the atmosphere, sea-ice, ocean, land surface and ocean-waves, as well

as the infrastructure required to perform data assimilation and ensemble forecasts. The IFS integrations presented here used the release Cy46r1, which was the operational model cycle at ECMWF between June 2019 and June 2020; a detailed technical and scientific documentation can be found in ECMWF (2019).

CHTESSEL computes the land surface response to atmospheric forcing, estimating the surface energy and water fluxes, as well as the evolution of soil temperature, moisture, and snowpack. The surface-atmosphere exchanges take place in a surface skin layer separating the sub-soil from the lower troposphere. The skin layer has no heat capacity, and its temperature is used to compute the upwelling longwave radiation, making it directly comparable with LST. Each grid-box can represent different types of land cover using multiple tiles, including bare ground, dominant high vegetation type, dominant low vegetation type, intercepted water (on the canopy), and shaded and exposed snow. The CHTESSEL LST considered in the present study corresponded to the area-weighted average LST over all tiles for each grid-point. A detailed description of the model can be found in chapter 8 of ECMWF (2019) along with a schematic representation of the model structure (*ibid*. Figure 8.1).

2.2.1. Updated Land Cover and Vegetation

The vegetation coverage in CHTESSEL is represented in terms of land cover types (see Table S1), which can be grouped into high vegetation (trees and forests) and low vegetation (grass, shrubs, crops, deserts, and tundra). This information is provided to the model as two-dimensional static input fields with the fractions of low and high vegetation (respectively, CV_l and CV_h, ranging between 0 and 1), the dominant vegetation type (TV_l and TV_h for the low, and high vegetation fractions, respectively), and the vegetation density parameters for low and high vegetation (respectively, $cveg_l$ and $cveg_h$, ranging between 0 and 1). The effective fractions of low and high vegetation are given, respectively, by:

$$C_l = CV_l \times cveg_l \tag{1}$$

$$C_h = CV_h \times cveg_h$$

and the effective grid-box total vegetation coverage is given by $C_l + C_h$.

In the original CHTESSEL formulation, the vegetation types and fractions are derived from the GLCC data set (Loveland et al., 2000), while the vegetation density parameters for low and high vegetation are obtained from lookup tables as a function of the respective dominant vegetation type. In this study, the representation of vegetation coverage in CHTESSEL was revised following the procedure proposed by NO20, extending its application to the full globe. This included updating the high and low vegetation fractions and types for all global land grid-points using information derived from the ESA-CCI land cover data set. First, two-dimensional distributions of the fraction of each of the 22 land cover classes in ESA-CCI were generated by spatial aggregation of the original 300 m data to the model resolution. The spatial aggregation was performed by counting the number of 300 m pixels of each class occurring within each model grid-cell. The derived fractional classes were then converted to the CHTESSEL land cover classes using a cross-walking table adapted from Poulter et al. (2015), from which revised CV_l, CV_h and TV_l, TV_h distributions were obtained (see NO20 for a more detailed description). Updating the vegetation fractions and types impacted several model parameters that modulate land-atmosphere interactions, including the roughness lengths for momentum and heat ($z0_m$ and $z0_h$), vegetation density ($cveg$), and minimum stomatal resistance ($rsmin$). This is because, for each tile in each grid-point, these parameters are prescribed in the CHTESSEL model using a lookup table for each vegetation type (cf. Table S1). The roughness lengths enter the calculation of the turbulent exchange coefficients. Following Jarvis (1976), the canopy resistance is a function of the rsmin normalized by the LAI (rsmin/LAI, see Equation 8.9 in ECMWF, 2019). In favorable atmospheric and soil moisture conditions, the rsmin/LAI ratio provides the minimum canopy resistance for evaporation, which adds to the aerodynamic resistance. CHTESSEL computes the surface energy balance for each tile independently. The turbulent fluxes are then aggregated, weighted by tile fraction, for the coupling with the atmosphere.

The LAI in the standard CHTESSEL formulation is obtained from a monthly climatology computed using the 2000–2008 period from the Moderate Resolution Imaging Spectroradiometer (MODIS). To be used in CHTESSEL, the total MODIS LAI was rescaled into high and low vegetation components with a static LAI data set previously used in the ECMWF system as described in Boussetta, Balsamo, Beljaars, Kral, and Jarlan (2013).

Here the LAI was revised processing the 10-daily CGLS LAI version 2 following NO20. To this end, a mean monthly CGLS LAI climatology was computed for each calendar month from the 1999–2018 period at 1 km resolution. The total LAI climatology at 1 km was then disaggregated to high- and low-vegetation LAI at the model resolution by averaging the LAI at 1 km in the pixels identified as dominant high or low vegetation. For this identification, the ESA-CCI land cover was used by generating the land-cover map at the 1 km resolution from the original 300 m. To generate a spatially consistent data set for high and low LAI, model grid points without any dominant type were filled with the total LAI.

The snow-free albedo in CHTESSEL is prescribed as a monthly climatology derived from MODIS. Therefore, despite the changes in land cover and LAI, all simulations share the same surface albedo in regions without seasonal snow cover. When snow is present, the grid-box surface albedo (α) is a weighted average on snow cover fraction (f_{sn}) between the snow-free (α_{sf}) and the snow albedo (α_{sn}):

$$\alpha = f_{sn}\alpha_{sn} + \left(1 - f_{sn}\right)\alpha_{sf} \tag{2}$$

In turn, α_{sn} is computed as a weighted average on effective grid-box high vegetation coverage between the exposed snow albedo (snow over bare ground and low vegetation, α_{esn}) and shaded snow albedo (snow under high vegetation, α_{ssn}):

$$\alpha_{sn} = \left(1 - C_h\right)\alpha_{esn} + C_h\alpha_{ssn} \tag{3}$$

The exposed snow albedo has a prognostic evolution depending on snow age varying between 0.85 for fresh snow and 0.5 for old snow. The shaded snow albedo is a function of vegetation type varying between 0.27 for evergreen needleleaf trees to 0.33 for deciduous needleleaf trees (Dutra et al., 2010). Snow-covered regions with high vegetation have a lower albedo due to the exposed snow-free canopy (ranging between 0.27 and 0.33) when compared with exposed snow (ranging between 0.5 and 0.85). Therefore, a reduction of high vegetation cover in snow-covered areas will cause a surface albedo increase (see Figure 1a for the changes in high vegetation cover between ESA-CCI and the default IFS land cover). Moreover, if snow is present, changes in the dominant type of high vegetation will also impact surface albedo.

2.2.2. Simulation Experiments

In the present study, six different experiments were considered to investigate the sensitivity of different variables to the model representation of the vegetation cover over the full globe—see Table 1 for a summarized description of the different experiments.

In the first experiment, denoted _CTR, the standard representation of the vegetation cover in CHTESSEL was considered. In the second experiment, denoted _CCI, the vegetation fractions and types were revised based on ESA-CCI land-cover data set, as presented in the previous section. The resulting changes to the vegetation fraction and type are presented in Figure 1. In the third experiment, denoted _LAI, the input LAI fields provided to CHTESSEL were updated based on the CGLS data set (cf. Figures 1c and 1d), as described in Section 2.2.1, in conjunction with the ESA-CCI-based update to the vegetation fractions and types. In the experiments _CCI and _LAI, the vegetation density parameters (and other parameters associated with each dominant vegetation type) were obtained from the CHTESSEL lookup tables (given in Table S1), analogous to the procedure in _CTR, but taking the updated vegetation types and fractions into account.

The fourth experiment was denoted _LCLUP. It used the same setup as experiment _LAI, including the revised LAI and vegetation fractions and types. However, in this case, $\mathrm{cveg_l}$ is computed based on LAI using a clumping approach for low vegetation following a Lambert-Beer law (Alessandri et al., 2017; Anderson et al., 2005):

$$\mathrm{cveg_l} = 1 - \exp\left(-k_l \times LAI\right) \tag{4}$$

The LAI dependence in Equation 4 introduced a representation of the seasonal cycle on the low vegetation density and coverage. A similar clumping approach has previously shown to improve the simulation of observed vegetation coverage seasonality over Iberia (NO20). Moreover, the Météo-France surface modeling framework (SURFEX) employs a similar approach to estimate $\mathrm{cveg_l}$ (Le Moigne, 2018). In this study, as in NO20, $k_l = 0.6$ is the same value adopted for SURFEX (see Le Moigne, 2018). In contrast, Alessandri et al. (2017) used $k_l = 0.5$ but NO20 found that the impact of using 0.5 or 0.6 in the LST simulation was

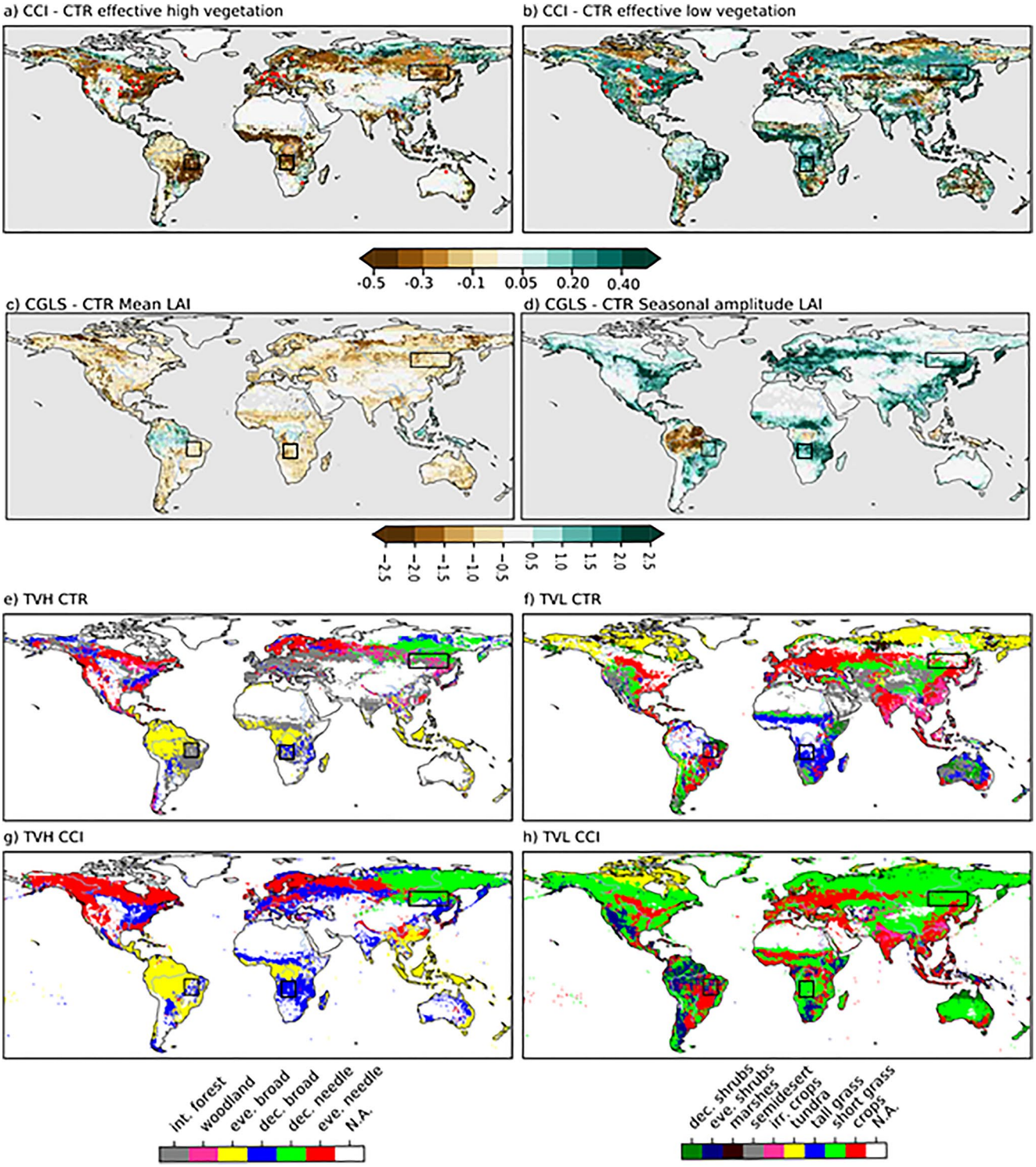

Figure 1. Global maps of change in effective (a) high (Ch) and (b) low (Cl) vegetation coverage between Climate Change Initiative (CCI) and CTR experiments. Differences between Copernicus Global Land Services (CGLS) and CTR mean Leaf Area Index (LAI) (c) and seasonal amplitude of LAI (d). Global maps of dominant type of high vegetation in experiment (e) CTR and (f) CCI, and dominant type of low vegetation in experiment (g) CTR, and (h) CCI. The red markers in panels (a and b) mark the location of the 51 FLUXNET stations considered in the present study.

Table 1
List of Experiments

Experiment	Offline	Coupled	Description
CTR	O_CTR	N_CTR	Control CHTESSEL experiment
CCI	O_CCI	N_CCI	Same as CTR experiment but with ESA-CCI land cover updates
LAI	O_LAI	N_LAI	Same as CTR experiment but with ESA-CCI land cover updates and CGLS LAI update
LCLUP	O_LCLUP	N_LCLUP	Same as CTR experiment but with ESA-CCI land cover updates, CGLS LAI update, and low vegetation clumping parameterization implemented
TG	O_TG	N_TG	Same as CTR, but with ESA-CCI land cover updates, CGLS LAI update, and replacing "Shrubs" vegetation types by "Tall Grass" to investigate biases in Brazil
SN	O_SN	N_SN	Same as CTR, but with ESA-CCI land cover updates, CGLS LAI update, and no shaded snow albedo to investigate the impact of snow albedo parameterization on the northern Eurasia bias during spring

Abbreviations: CGLS, Copernicus Global Land Services; CHTESSEL, Carbon Hydrology Tiled ECMWF Scheme for Surface Exchanges over Land; ESA-CCI, European Space Agency Climate Change Initiative; LAI, leaf area index.

small. In NO20, the clumping was also applied to the high vegetation. However, initial testing in coupled simulations indicated a disproportionate impact of the high vegetation clumping on the Northern Hemisphere circulation that requires further investigation.

The fifth experiment was denoted _TG. It uses the same setup as the LAI experiment, but replacing the "Shrubs" class by "Tall Grass". This change was motivated by the emergence of relevant systematic errors over regions where the "Shrubs" class is introduced by the ESA-CCI-based revision of the vegetation types and fractions, as discussed in Section 5.

Finally, the sixth experiment, denoted _SN, also used the same setup as the _LAI experiment. The difference was the removal of the snow shadowing effect in CHTESSEL by replacing the shaded snow albedo (α_{ssn}) by the snow-free albedo (α_{sn}) in the surface albedo calculation (Equation 3). This change was motivated by a systematic cooling occurring during boreal winter and spring over snow-covered regions, which was shown here to be associated with the snow shadowing effect together with the ESA-CCI-based updates to the vegetation cover (discussed in Section 5).

2.2.3. Simulation Setup

The six experiments described in Section 2.2.2 were performed using two different model setups. In the first setup, the simulations were generated with CHTESSEL forced by ERA5 fields using an offline configuration. The forcing fields included near-surface (10 m) air temperature, wind, and humidity, along with surface pressure, rainfall and solar and thermal downwelling radiative fluxes. The offline simulations spanned over the 1995–2018 period, covering the full global land at TL639 resolution (same resolution as ERA5, i.e., about 0.28° in the grid-point space). In the second model setup, coupled land-atmosphere simulations were performed. None of the two setups used data assimilation. The coupled simulations spanned only from December 2009 to November 2010 due to computational cost. The initial atmospheric conditions were retrieved from ERA5, while the initial conditions for land surface and sub-surface were retrieved from the corresponding offline CHTESSEL simulations to avoid the need for the land surface state spin-up. The coupled setup used the same global TL639 horizontal resolution with 137 vertical levels as ERA5.

The coupled simulations were performed with nudging, relaxing the atmospheric vorticity and temperature fields toward ERA5, to avoid drifts in large-scale flow. For vorticity, nudging was applied in spectral space truncated at T63 with a 12-h timescale on all levels above 800 hPa to allow a free evolving boundary layer. A temperature nudging with a 120-h timescale between 280 and 15 hPa was also applied to prevent drifts around the tropopause. Moisture, clouds, and temperature in the lower troposphere and the divergent flow were not relaxed, being controlled by the model following the large-scale variability (at T63) imposed by the vorticity nudging above 800 hPa. A similar coupled model simulation with atmospheric nudging the

setup in Beljaars (2020) showed comparable skill to day 2–3 short-range forecasts. The full set of simulations considered in the present work, encompassing the six experiments carried out on each model setup (offline and coupled) were summarized in Table 1.

2.3. Evaluation

The systematic errors in LST simulated by the offline CHTESSEL and coupled IFS models were evaluated using the LSA-SAF LST as reference, which limited the error analysis to clear-sky conditions. The model simulations were interpolated from the original resolution TL639 to a regular 0.25° × 0.25°, which is the same grid of the ERA5 data available on the Copernicus Climate Data Store. The LSA-SAF LST hourly data were first coarse-grained from their original 3 km nadir resolution to the 0.25° × 0.25° model grid by computing the mean value of all pixels falling within each grid cell. The fraction of valid pixels was retained for each grid cell and each time-step during the coarse-graining procedure, accounting for cloud contamination and other error sources in the LSA-SAF LST product. Only data points with a fraction of valid pixels greater than 70% were retained. On the other hand, to ensure a comparison between clear-sky conditions, the simulated LST values were discarded when the total cloud cover of the model was above 30%. Notice that the total cloud cover was a model diagnostic variable for IFS simulations, while for CHTESSEL offline simulations the total cloud cover from ERA5 was considered. The threshold percentages of 70% and 30% were based on JO19 and NO20. The bias in the simulated daily maximum and minimum LST was computed for each season for grid cells where both model and observations experienced at least 15% of clear-sky conditions throughout the full (seasonal-averaged) diurnal cycle.

The sensitivity of the systematic errors in simulated daily maximum and minimum LST to the changes introduced in the vegetation cover representation in CHTESSEL was assessed by taking the differences between the absolute value of the bias in the CTR experiment and the other five experiments considered here. This procedure was repeated for each season and for each grid cell of the MSG disk.

The response of LST and 2-m air temperature (T_{2m}) fields to LULC changes in the coupled simulations was also investigated by taking the differences between the N_CTR and the other five coupled simulation experiments. The sensitivity of the simulated fields was also computed for each season but, unlike the bias, it was not constrained to the MSG disk of the LSA-SAF product nor to clear-sky conditions, and it was computed for all global land grid cells under all-sky conditions.

A further assessment of the coupled simulations was carried by computing the root mean square errors (RMSEs) of the simulated temperature at four different pressure levels (500, 700, 850, and 925 hPa) using the ERA5 analysis as reference. ERA5 short-range forecasts—valid for lead times +24 h to +48 h (day 2), +48 to +72 h (day 3), and +72 to +96 h (day 4)—were also included. This comparison provided a reference to the amplitude of the atmospheric response resulting from the perturbations induced by the vegetation changes in terms of forecast skill at different forecast lead times. For this comparison, all data were reprojected to a global regular 1° × 1° grid, to reduce data processing. This differs from the 0.25° × 0.25° grid used in the remaining analysis, but it was justified by the fact that in this case RMSE was averaged over three large regions. These regions were defined to cover the entire globe, separating the tropics from the extra-tropics and the two hemispheres due the different seasonality: extratropical Northern Hemisphere (NH, poleward 20° North), Tropics (TROP, between 20° South and 20° North) and extratropical Southern Hemisphere (SH, poleward 20° South). The RMSE was computed on the daily mean field for each region weighting each grid-box by the cosine of latitude.

Finally, turbulent fluxes from eddy-covariance tower data from the FLUXNET 2015 synthesis data set (Pastorello et al., 2020) for the period 1996–2014 were used to evaluate the offline CHTESSEL latent and sensible heat fluxes. The simulation time series were extracted from the grid cell containing the tower flux location. The tower flux data were processed as by Martens et al. (2017, 2020) and aggregated to daily frequency. Only sites with at least 5 years of daily records were retained, resulting in 51 sites (see Figure 1 for their locations). The requirement of at least 5 years of daily data was imposed to allow the calculation of a daily climatology for each site. These were then used to compute the temporal correlation of the standardized anomalies to remove the influence of the seasonal cycle as in Martens et al. (2020). The standardized anomalies were computed by normalizing the daily data by the mean and standard deviation of the local climatology, which was

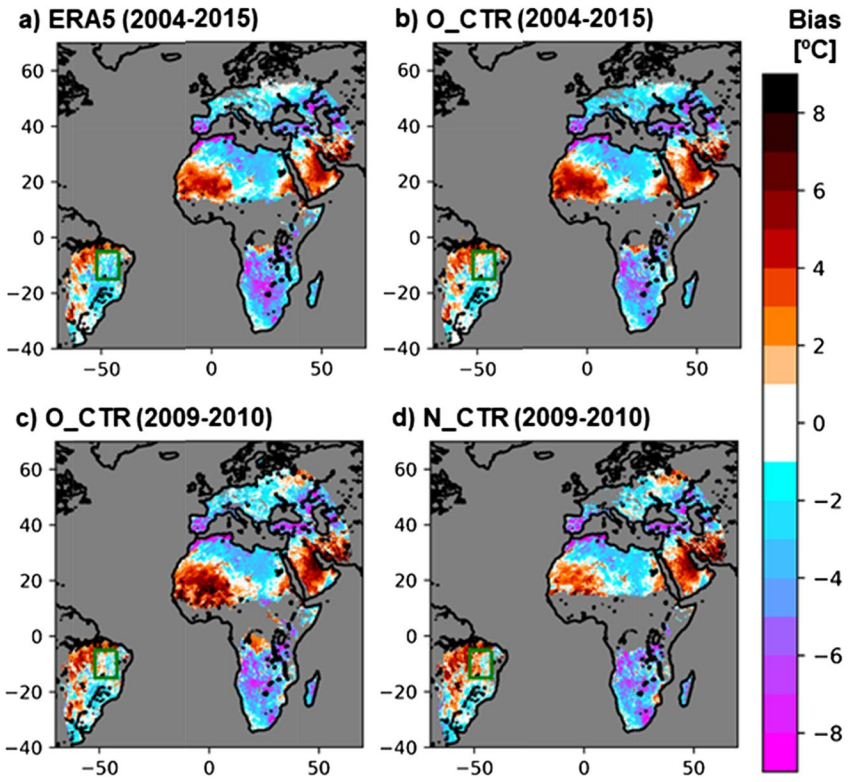

Figure 2. Mean JJA daily maximum LST biases under clear-sky conditions in (a) ERA5, (b) O_CTR for 2004–2015, (c) O_CTR for 2009–2010, and (d) N_CTR for 2009–2010 period. The Land Surface Analysis Satellite Application Facility LST is used as reference. The green box identifies the eBrz domain (see Section 5).

based on a centered 31-day window for each day of the year averaged over the multi-year period. In addition to the temporal correlation (of the standardized daily anomalies, and the raw daily data), the evaluation also included the bias and the unbiased RMSE of the raw data ($\text{ubRMSE} = \sqrt{\left(\left(\text{mod} - \overline{\text{mod}}\right) - \left(\text{obs} - \overline{\text{obs}}\right)\right)^2}$, where mod and obs are the simulated and observed time-series respectively, and the overbar represents temporal averaging); all these metrics were computed per station. The 1-year long coupled simulations were too short and affected by missing data for a robust assessment of the simulated fluxes.

3. Simulation Error Assessment

3.1. LST Systematic Errors

We started our analysis by evaluating the LST simulated in the ERA5 data set considering LSA-SAF LST over 2004–2015 (available over the MSG disk over clear-sky conditions as reference). Emphasis was given to boreal summer (JJA), corresponding to the season with the largest fraction of valid grid points (meeting the clear-sky conditions defined in Section 2.3).

The results showed that daily maximum LST from ERA5 displayed large absolute biases over most land areas during JJA (Figure 2a). Specifically, we found cold biases (ranging between −1°C and −12°C) over southern Europe, large portions of northern, central, and southern Africa, and eastern Brazil. In contrast, the Arabian Peninsula, tropical western Africa, and western Brazil displayed warm bias values (ranging between +1° and +12°C).

The offline CHTESSEL simulation O_CTR reproduced the main features of the systematic error patterns in ERA5 JJA daily maximum LST over the 2004–2015 period (Figure 2b), in agreement with the findings by NO20 and JO19 over Iberia. Figure 2c showed that the identified cold and warm biases were also present

in O_CTR when the errors were computed only over the December 2009 to November 2010 period in which the coupled IFS simulations were performed (Figure 2c). Finally, the coupled simulation N_CTR also reproduced the main features of the JJA LST bias pattern in ERA5 (Figure 2d).

Overall, the main patterns of daily maximum LST systematic errors under clear-sky conditions were alike in ERA5 and offline and coupled simulations throughout all seasons (Figure S1). Thus, both the offline and the coupled simulations could be used to investigate the causes of these errors. The large daily maximum LST cold bias over southern Europe, southern Africa, and coastal northern Africa were present throughout all seasons in ERA5, O_CTR, and N_CTR (Figure S1). In contrast, the warm bias over tropical western Africa and Arabian Peninsula displayed a strong seasonality: during boreal winter (DJF) the sign of the bias switched to negative, being spatially consistent with a general cold bias across the Sahara Desert (Figure S1). The seasonal cycle of the cold and warm bias over Brazil as deduced from LSA-SAF was masked due to the low percentage of clear-sky data points during DJF and SON. Yet, during SON, it was still possible to identify a similar cold bias over eastern Brazil in ERA5, O_CTR, and N_CTR to the one found for JJA (Figure S1).

Finally, we noticed that the updates to CHTESSEL vegetation implemented here had a negligible impact on the strong warm bias in western Africa and the Arabian Peninsula during JJA (cf. Section 3.2), which were likely related to aerosols affecting both satellite retrievals and IFS radiation (see Section. 4 for a more detailed discussion of this issue).

3.2. Impact of Model Changes on LST

The impact of the updated vegetation types and fractions using the ESA-CCI data set on the JJA daily maximum LST simulated by offline CHTESSEL (O_CCI) and coupled IFS runs (N_CCI) compared to the standard CHTESSEL configuration were shown, respectively, in Figures 3a and 3b. The results revealed neutral impacts in over 83% of grid points with valid observations for O_CCI (86% for N_CCI), encompassing large portions of northern Africa, Arabian Peninsula, and southern Europe. Positive impacts emerged in 10% of the grid points for O_CCI with valid LST observations (9% for N_CCI). Finally, negative impacts were found over 7% and 5% of the grid points, respectively, for O_CCI and N_CCI. Here, "positive impacts" refer to reductions in absolute bias greater than 2°C compared to the respective standard simulation (O_CTR for O_CCI and N_CTR for N_CCI), "negative impacts" refer to increases in absolute bias greater than 2°C, and "neutral impacts" refer to differences in absolute bias within 2°C (the 2°C threshold is representative of the uncertainty in LSA-SAF LST estimates provided in Trigo et al., 2015). Despite the relatively low percentages of pixels with positive impacts, a coherent large bias reduction over southern Africa and the Maghreb coast region emerged for O_CCI and N_CCI (up −12°C for in both offline and coupled simulations). The negative impacts were mostly located in southern and central Africa and the Tocantins region in eastern Brazil (increases in absolute bias up to +10°C for O_CCI and +7°C for N_CCI).

The impacts of updating the CHTESSEL LAI using the CGLS product, in addition to the ESA-CCI-based updates, were shown in Figure 3c for the offline simulation (O_LAI), and Figure 3d for the coupled simulation (N_LAI). The impact on JJA daily maximum LST was nearly identical to the results obtained in O_CCI and N_CCI, with positive impacts for 10% (11%) of grid points with valid observations, negative impacts for 8% (5%), and neutral impacts for 82% (84%) for O_LAI (N_LAI). Overall, this suggested a small influence of updating the LAI fields on daily maximum LST systematic errors. The main difference resulting from updating LAI is the expansion and amplification of the bias around the state of Tocantins in Brazil.

The impact of introducing the clumping parameterization for low vegetation, together with the ESA-CCI-based land cover and CGLS-based LAI updates, was shown in Figure 3e for the offline simulation (O_LCLUP) and Figure 3f for the coupled simulation (N_LCLUP). The results showed positive impacts in 16% (17%), negative impacts in 9% (8%), and neutral impacts in 75% (75%) of grid-points with valid observations. Specifically, low vegetation clumping resulted in a strengthened bias reduction over southern Africa, Iberia, southeastern Europe, and the Maghreb coastal region. Neutral impacts concentrated mostly over northern Africa and southwestern Asia. A small region of increased bias emerged over tropical western Africa only in O_LCLUP, which is likely related to misrepresentation of the Intertropical Convergence Zone variability in ERA5 (e.g., Nogueira, 2020), and not driven by LULC changes. Over Brazil, the low vegetation clumping further amplified the bias around the Tocantins state in both offline and coupled simulations.

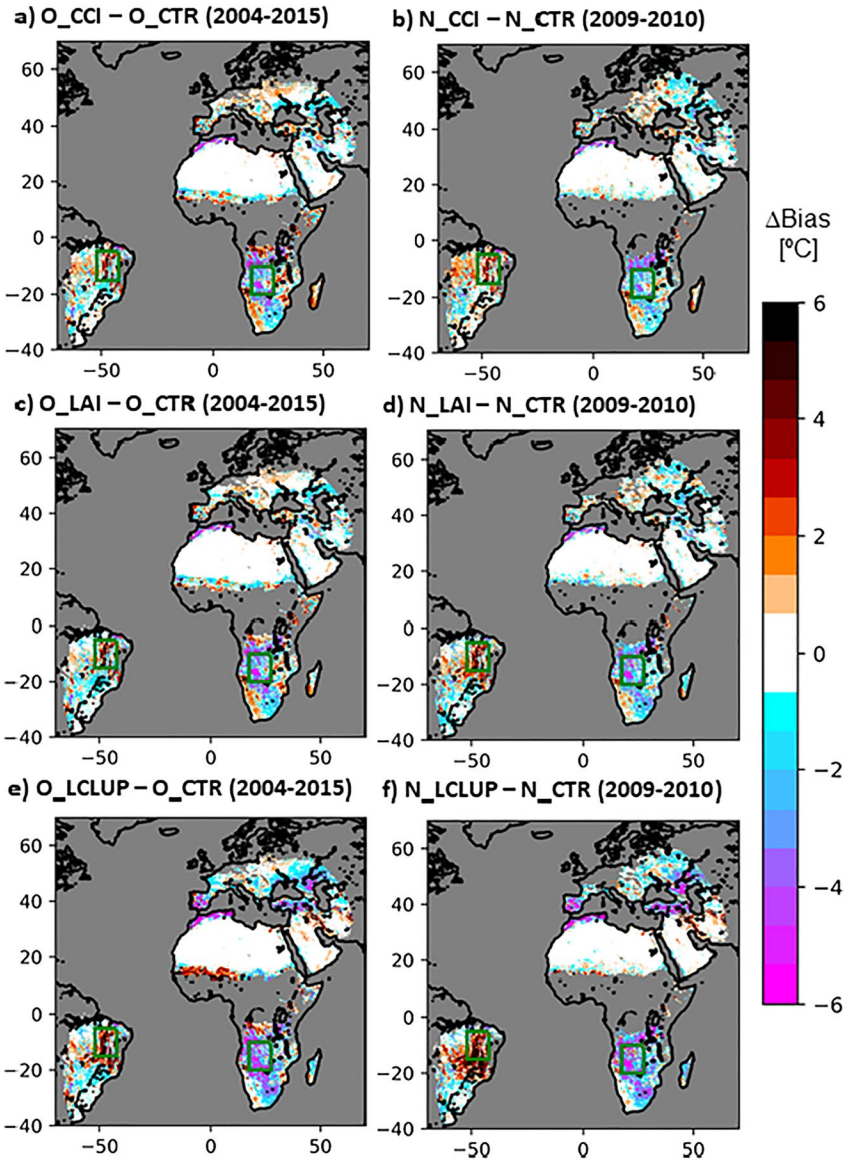

Figure 3. Changes in absolute bias of the JJA daily maximum LST bias between different experiments: (a) O_CCI–O_CTR, (b) N_CCI–N_CTR, (c) O_LAI–O_CTR, (d) N_LAI–N_CTR, (e) O_LCLUP–O_CTR, and (f) N_LCLUP–N_CTR. The green box identifies the eBrz sub-domain (see Section 5 for details).

The above results concern only the JJA period, when the fraction of clear-sky data was maximum. However, Figures S2 and S3 showed that the main daily maximum LST bias changes associated with the proposed model vegetation improvements were representative of all seasons, respectively, for offline and coupled simulations. By "representative" we mean that, in general, the location and sign of the main features identified in Figure 3 were present in all seasons (although differences in magnitude and extension may occur).

In summary, the results indicated an overall improvement in daily maximum LST simulations when all three updates (vegetation cover and fraction, LAI, and clumping) were concurrently implemented, with the main exception of eastern Brazil; the benefit was evident over all seasons, and both for offline and coupled model setups (cf. Figure S4). The percentage of grid points with positive, neutral, and negative impacts for all simulations over all seasons also supported this conclusion (cf. Table S2).

Another important point was the fact that all proposed changes to CHTESSEL had essentially a neutral impact on daily minimum LST for all seasons (cf. Figures S5 and S6, respectively, for offline and coupled

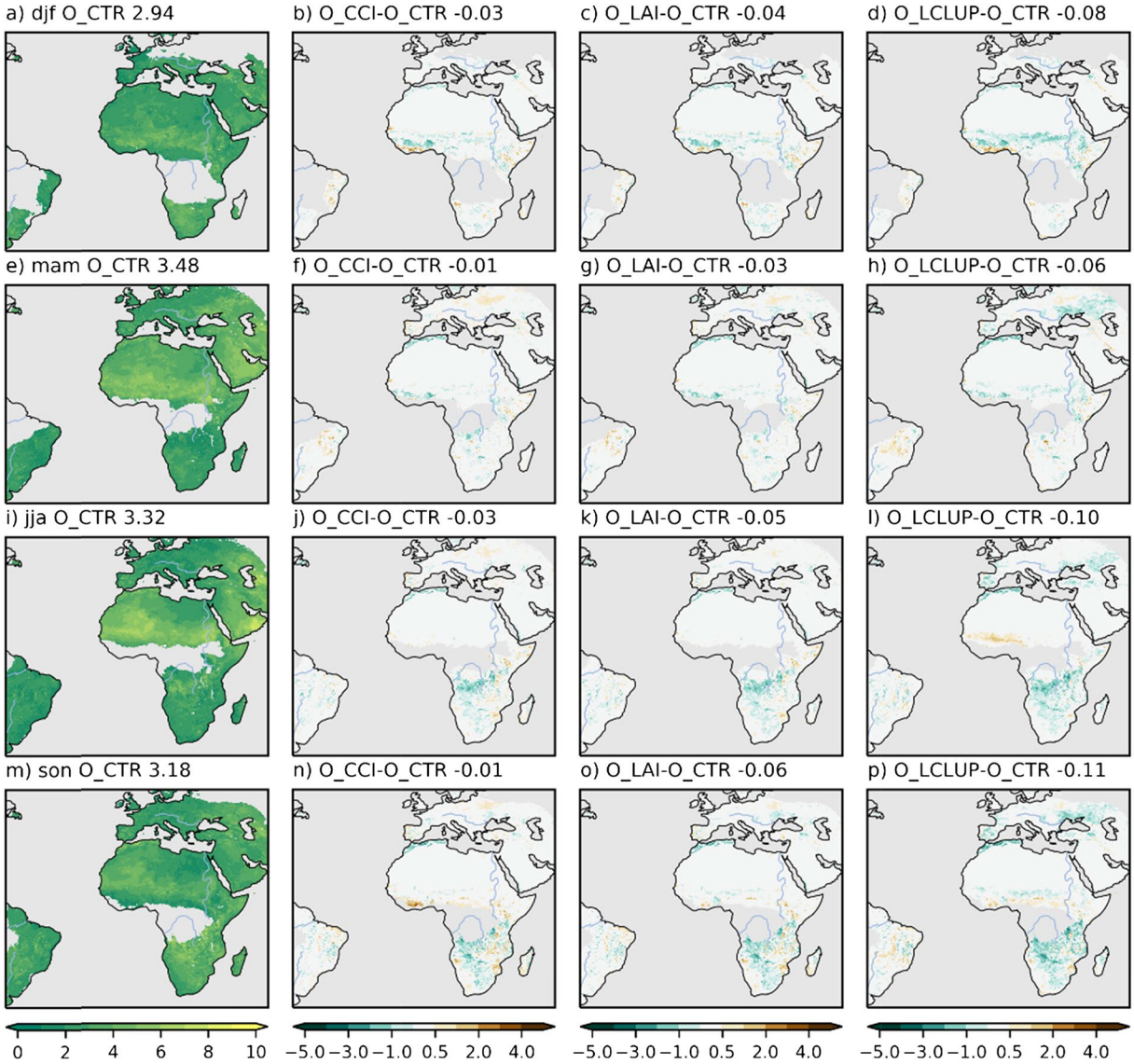

Figure 4. Unbiased root-mean-squared error (ubRMSE) computed for daily maximum LST from O_CTR over the 2004–2015 period during (a) DJF, (e) MAM, (i) JJA, and (m) SON. The Land Surface Analysis Satellite Application Facility observations were considered as reference. The differences in ubRMSE between O_CCI and O_CTR are shown in (b) DJF, (f) MAM, (j) JJA, and (n) SON. The differences in ubRMSE between O_LAI and O_CTR are shown in (c) DJF, (g) MAM, (k) JJA, and (o) SON. Finally, the differences in ubRMSE between O_LCLUP and O_CTR are shown in (d) DJF, (h) MAM, (l) JJA, and (p) SON.

simulations). This fact was also revealed in NO20 over Iberia. However, Figure S7 showed widespread systematic errors in the daily minimum LST from ERA5 and from both offline and coupled simulations. A more detailed analysis of the daily minimum temperature errors in CHTESSEL is needed but falls outside the focus of the present investigation and will be addressed in subsequent work.

We assessed the impact of the proposed updates to CHTESSEL on the unbiased root-mean-squared error (ubRMSE) and Pearson temporal correlation coefficient, using LSA-SAF LST as reference. The O_CTR simulation showed a large range in daily maximum LST ubRMSE (Figures 4a, 4e, 4i and 4m), from low values below 2°C (within observational uncertainty, e.g., over most of South America and Europe) to large errors up to 10°C (e.g., over central and southern Africa and the Arabian Peninsula). The daily maximum LST ubRMSE difference in O_CCI compared to O_CTR, averaged over all grid-points with valid data, ranged between −0.01°C and −0.03°C for all seasons, representing an overall negligible improvement. Yet, larger magnitude ubRMSE reductions were found over southern and central Africa in all seasons (mostly around

−1°C, but with values up to −5°C over specific regions, Figures 4b, 4f, 4j and 4n), concurrent with daily maximum LST bias reductions. Some localized increases to ubRMSE also emerged, particularly in a narrow band over central western Africa region associated with the ITCZ.

The patterns of difference in daily maximum LST ubRMSE between O_LAI and O_CTR were nearly identical to the ones found for O_CCI (Figures 4c, 4g, 4k and 4o), but with a slight increase in the domain-averaged improvements (between −0.03°C and −0.06°C for all seasons). Introducing the low vegetation clumping resulted in a further slight reduction of seasonal domain-averaged daily maximum LST ubRMSE compared to O_CCI (between −0.06°C and −0.11°C for all seasons, Figures 4d, 4h, 4l and 4p). Regionally, the improvements associated with O_LCLUP (between −1°C and −2°C compared to O_LAI) occurred mostly over central Africa, southern Europe, Eastern Europe and western Russia.

We found strong temporal correlations (>0.8) for daily maximum LST between O_CTR and LSA-SAF LST over most grid-points with available data during all seasons, the main exception being northern Africa and the Arabian Peninsula where slightly lower correlations (0.6–0.8) emerged. However, the results also showed that the proposed updates to CHTESSEL had negligible impacts on these correlations patterns during all seasons (cf. Figure S8).

Similarly, the impact of the proposed updates to CHTESSEL vegetation cover were negligible for daily minimum LST temporal correlations (Figure S9). Finally, the seasonal domain-averaged differences in daily minimum LST ubRMSE were small (between 0.00°C and −0.03°C) when comparing O_CCI, O_LAI, and O_LCLUP against O_CTR (cf. Figure S10). Regionally, the proposed updates to CHTESSEL vegetation cover resulted in daily minimum LST ubRMSE increases (mostly below +2°C, i.e., within observational uncertainty) over a narrow band in central western Africa, and daily minimum LST ubRMSE decreases (mostly below −2°C) over southern western Africa.

3.3. Impact of Model Changes on Surface Fluxes

Overall, we found little impact of the proposed vegetation updates to the simulated turbulent fluxes at the available station locations. All four offline simulations (O_CTR, O_CCI, O_LAI, and O_LCLUP) show very similar systematic errors: a positive latent heat bias (median of all four simulations: 12.5 Wm^{-2}, Figure 5a) along with a negative sensible heat bias which is lower in magnitude (−7.0 Wm^{-2}, Figure 5b).

The ubRMSE of the raw data was also similar among the different simulations and for both fluxes (23.2 Wm^{-2} for latent heat and 25.9 Wm^{-2} for sensible heat, Figures 5c and 5d, respectively). The temporal correlations showed higher values for latent heat flux (0.8, Figure 5e) compared with sensible heat flux (0.7, Figure 5f). As expected, lower temporal correlations were seen when removing the seasonality, with lower median for latent heat flux (0.39 Wm^{-2}, Figure 5g) than for sensible heat flux (0.54 Wm^{-2}, Figure 5h). The latter had a less pronounced seasonality, which explained the lower drop in correlations when factoring out the seasonal cycle (Martens et al., 2020).

The neutral impact at the FLUXNET tower station locations does not contradict the overall improvements in simulated daily maximum LST. Indeed, a comparison of Figures 1 and 3 showed that the neutral impact of the proposed updates to CHTESSEL vegetation cover on the latent and sensible heat fluxes at the FLUXNET locations was coherent with the small changes to LST introduced by _CCI, _LAI, and _LCLUP experiments over the tower locations. This issue is further discussed in Section 4.1.

3.4. Global Sensitivity of LST to CHTESSEL Vegetation Updates

In this section, we extended the analysis on the impact of the vegetation updates implemented in CHTESSEL to the LST over the full globe. This was not possible for the LSA-SAF LST which was restricted to the MSG disk. Hence, we focused on the differences between the experiments with updated vegetation against the baseline CTR simulation. Moreover, to assess the atmospheric response to the vegetation changes, we analyzed the near surface air temperature fields generated by the coupled simulations.

Overall, the coupled simulations revealed the sensitivity of daily maximum LST to the ESA-CCI-based updates in vegetation types and fractions, as well as the CGLS-based updates in LAI and the introduction of

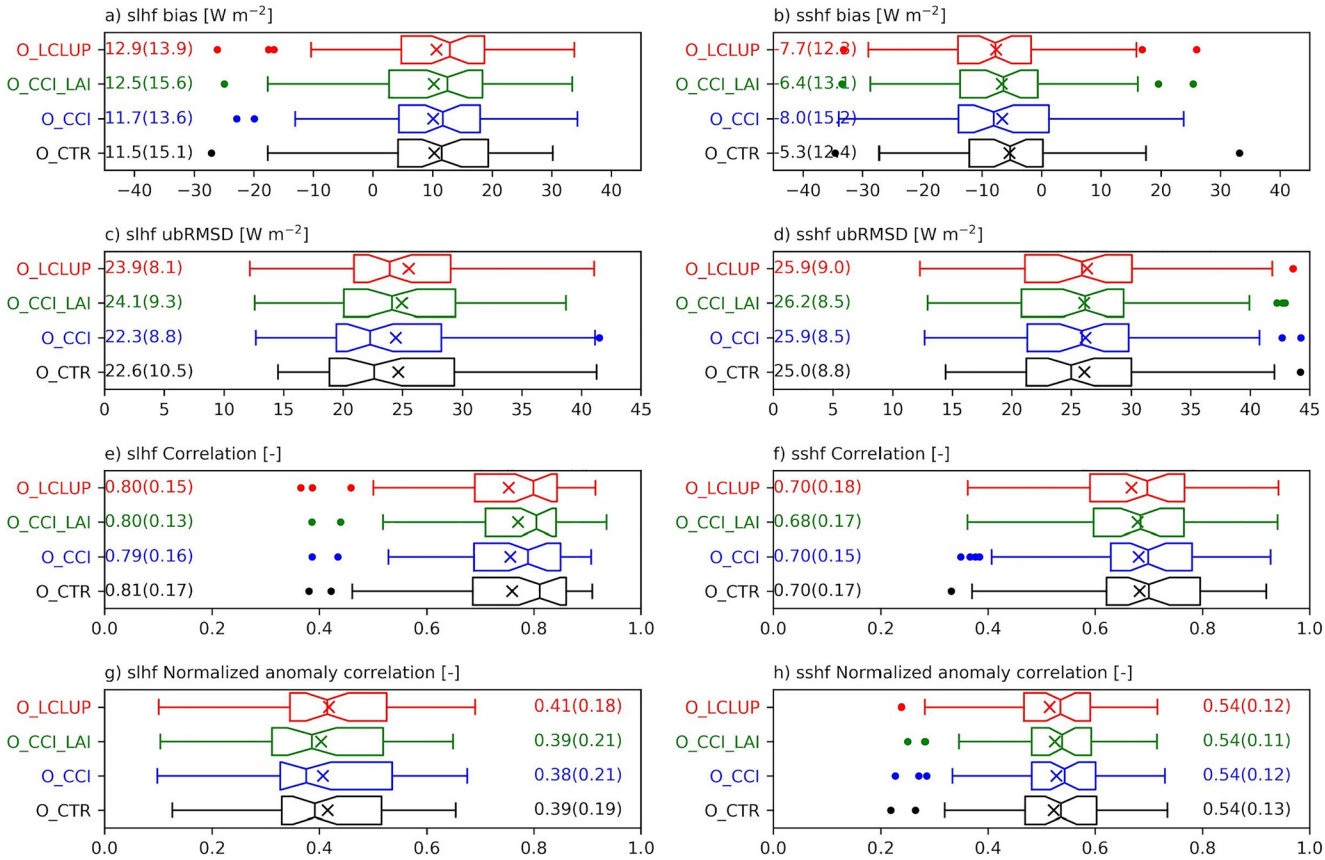

Figure 5. Evaluation of surface latent heat flux (left) and sensible heat flux (right) for the different simulations (vertical axis) against the measurements from 51 FLUXNET sites. The boxplots extend from the minimum to maximum with the boxes showing the 25th and 75th percentiles, with the mean and median represented by the cross and vertical line, respectively. The outliers (dots) denote values beyond 1.5 times the interquartile range from the closest 25th or 75th percentile. The statistical error metrics include the bias (a and b), unbiased root mean square error (c and d), the Pearson's correlation (d and f), and the Pearson's correlation of the standardized anomalies (g and h). Next to each boxplot, the values indicate the median and interquartile range.

low vegetation clumping (Figure 6). However, this influence showed large regional and seasonal variability. Specifically, the incorporation of ESA-CCI land cover in N_CCI resulted in a larger than +1°C increase in daily maximum LST relative to N_CTR over 8% of global land during DJF (Figure 6a), 9% during MAM (Figure 6d), 18% during JJA (Figure 6g), and 10% during SON (Figure 6j). In contrast, N_CCI showed decreases in daily maximum LST of more than −1°C compared to N_CTR in 15% of global land during DJF, 21% during MAM, 10% during JJA, and 9% during SON.

Most noticeably, N_CCI showed widespread cooling of northern Eurasia and North America during DJF and MAM (generally below −2°C but with values up to −7°C emerging during MAM over parts of northern Asia and eastern North America, Figures 6a and 6b), and a cooling of western Australia, southern Asia, and south and western North America in all seasons (mostly below −2°C, except for a specific region in western North America where the cooling reached −9°C during JJA). N_CCI results also showed widespread warming of eastern and northwestern North America and northern Eurasia during JJA (mostly below +3°C and +4°C, respectively, in northern Eurasia and North America, but with values up to +8°C in certain areas, Figure 6g), warming over much central and southern Africa in all seasons (up to +7°C), and warming of eastern Brazil in all seasons (lower than +2°C during DJF and MAM, but reaching up to +5° during JJA and SON).

Moreover, we noticed that similar changes to high and low vegetation fractions introduced by CCI-based update resulted in different LST responses over different regions. For example, strong increases to low vegetation fraction and respective decreases to high vegetation fraction emerged over northern Eurasian and southern Africa (Figures 1a and 1b), but their LST responses clearly differ (Figures 6a, 6d, 6g and 6j).

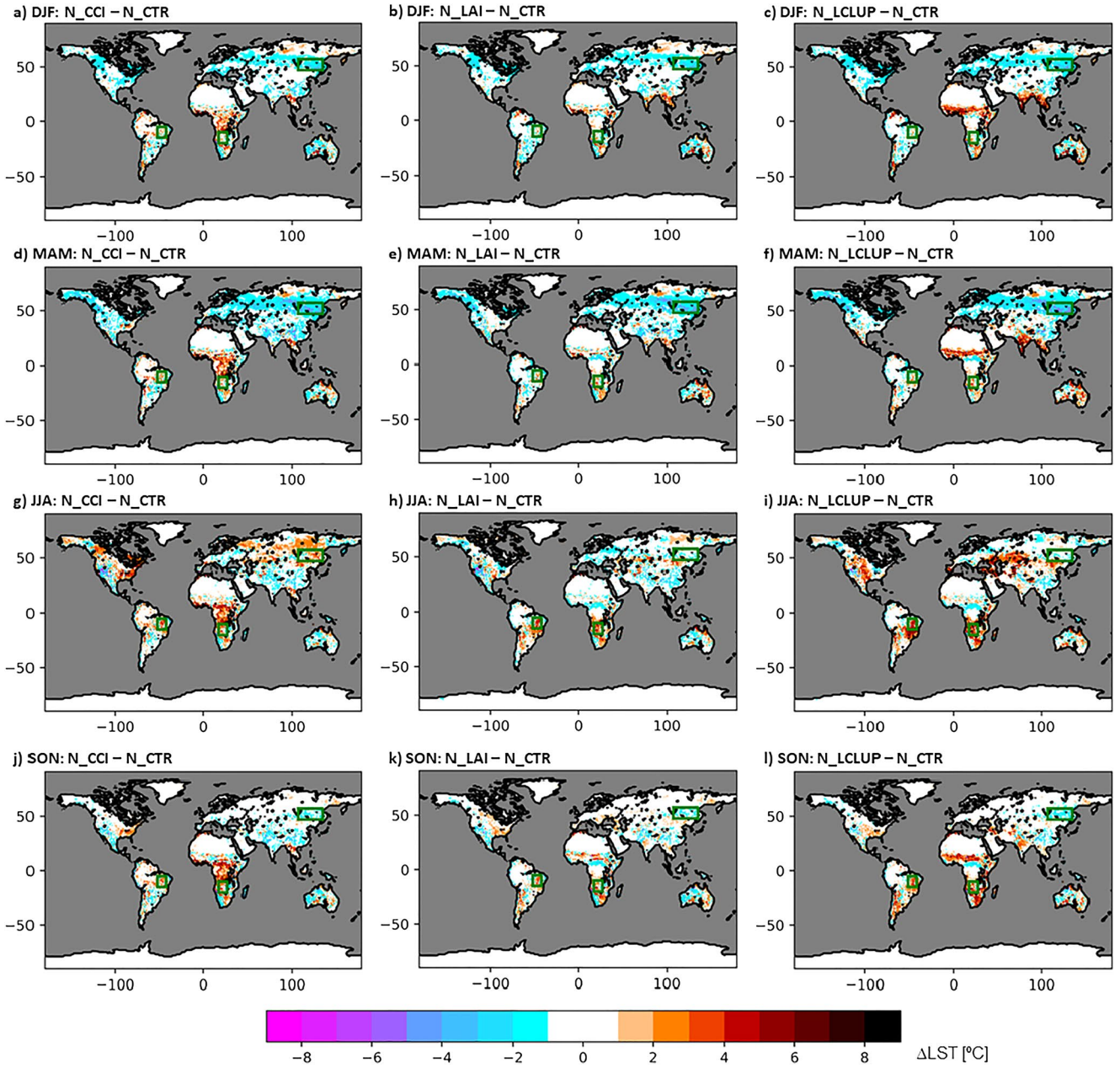

Figure 6. Differences in seasonal averaged daily maximum LST between the experiments with updated vegetation and the control using the coupled simulations. From top to bottom, the different panels represent different seasons (respectively DJF, MAM, JJA, and SON). From left to right the different panels represent differences between different experiments (respectively N_CCI–N_CTR, N_LAI–N_CTR, and N_LCLUP–N_CTR). The green boxes identify the eBrz and eRus sub-domains (see Section 5).

The CGLS-based LAI resulted in an increase of daily maximum LST larger than +1°C compared to N_CTR in 8% of global land during DJF (Figure 6b), 8% during MAM (Figure 6e), 11% during JJA (Figure 6h), and 10% during SON (Figure 6k). Conversely, N_CCI resulted in decreases to daily maximum LST of more than −1°C compared to N_CTR in 14% of global land during DJF, 19% during MAM, 11% during JJA, and 7% during SON.

Overall, Figures 6b, 6e, 6h and 6k showed that the spatial patterns of the changes to daily maximum LST for N_LAI–N_CTR were similar to those found for N_CCI–N_CTR. The main exceptions included eastern United States and northern Asia during JJA, and over central Africa in all seasons. In the cases of central

Africa and northern Asia, the warmer LST values when comparing N_CCI to N_CTR mostly disappeared for N_LAI (or even switched its sign, e.g., over eastern United States during JJA and a narrow band in central Africa during MAM and JJA, where cooling magnitudes up to −2°C emerged). We noted that the LAI seasonal amplitude increased over most of these three regions after the CGLS LAI-based update (with differences exceeding +2.0°C compared to N_CTR, Figure 1d). Additionally, the yearly averaged LAI (Figure 1c) mostly decreased over central Africa and northern Asia (between −0.5 and −2.0) and eastern United States (below −1.0). Interestingly, the decrease in yearly averaged LAI (below −1.0, Figure 1c) and increase in LAI seasonality (between +0.5 and +2.0, Figure 1d) over eastern Brazil and India, which were similar to the changes over eastern United States, induced an opposite sign daily maximum LST change, with a weak warming (+1°C on average) over eastern Brazil during JJA (Figure 6h) and over India during DJF (Figure 6b).

Implementation of low vegetation clumping also resulted in an increase of daily maximum LST larger than +1°C compared to N_CTR in 10% of global land during DJF (Figure 6c), 11% during MAM (Figure 6f), 16% during JJA (Figure 6l), and 12% during SON (Figure 6l). Conversely, N_LCLUP resulted in a decrease of daily maximum LST of more than −1°C compared to N_CTR in 16% of global land during DJF, 18% during MAM, 9% during JJA, and 6% during SON. The main differences between N_LCLUP and N_LAI included a further amplified the daily maximum LST over southern Africa and Brazil during JJA (on average +2°C, Figure 6i) and SON (by about +1°C, Figure 6l), over southern South America during SON (on average +1°C, Figure 6l), and southern Asia in DJF (+1°C, Figure 6c) and MAM (between +1°C to +2°C, Figure 6f). Additionally, Figure 6 also showed strong warming (up to +7°C) in N_LCLUP compared to N_LAI over a narrow band over tropical Africa during DJF (Figure 6c), MAM (Figure 6f) and SON (Figure 6j), and over Iberia, southeastern Europe, western Asia and central United States during JJA (Figure 6i).

3.5. Global Response of T_{2m} to CHTESSEL Vegetation Updates

Figure 7 showed that the changes to CHTESSEL vegetation coverage and the associated changes to daily maximum LST affected the simulated daily maximum near-surface air temperature in coupled simulations. On the one hand, N_CCI, N_LAI, and N_LCLUP resulted in increases to daily maximum T_{2m} larger than +1°C compared to N_CTR, respectively, over 2%, 1%,1% of global land during DJF, 4%, 3%, 3% during MAM, 9%, 5%, 4% during JJA, and 3%, 1%, 1% during SON, respectively. On the other hand, the percentage of pixels with decreases to daily maximum T_{2m} of more than −1°C for N_CCI, N_LAI, and N_LCLUP was, respectively, 5%, 4%, and 8% during DJF, 9%, 8%, and 13% during MAM, 0%, 1%, and 2% during JJA, and 0%, 0%, 1 during SON.

We found widespread reduction of T_{2m} for N_CCI, N_LAI, and N_LCLUP during DJF (Figures 7a–7c) and MAM (Figures 7d–7f) over the northern Eurasia and northern North America regions (mostly below −1°C, but with values up to −4°C over localized regions) concurrent to the surface cooling showed in Figure 6. During JJA, the surface warming over northern Eurasia found for N_CCI (cf. Figure 6g) was also reflected in daily maximum T_{2m} increases (ranging between small values below +1°C and values up to +3°C, Figure 7g), but with T_{2m} warming covering a larger area (particularly in northeastern Asia). This areal expansion was larger for N_LAI (Figure 7h) and N_LCLUP (Figure 7i), where the T_{2m} warming over northern Siberia during JJA is clearly stronger compared to the respective LST change (cf. Figures 6h and 6i). Updating LAI reduced the extent of T_{2m} warming over northern Eurasia and northern North America (Figure 7h).

The surface warming over central Africa for N_CCI during all seasons (Figures 6a, 6d, 6g and 6j) was also reflected in daily maximum T_{2m} (up to +5°C, Figures 7a, 7d, 7g and 7j). In contrast, the strong daily maximum LST warming signal over southern Africa during JJA, particularly in N_CLUP (Figure 6i), was largely reduced (or even changed signal) for daily maximum T_{2m} (Figure 6i). Similar to what was found for LST, the T_{2m} warming over central Africa that originated after incorporating the CCI land cover (Figures 7a–7j) was entirely offset after updating LAI, or even changed in sign (Figures 7b, 7e, 7h and 7k). The same was true over eastern United States during JJA (Figures 7g–7i). Additionally, updating LAI reduced (increased) the JJA and SON warming over northern (southern) South America with differences mostly below 1°C, and caused warming in southern Asia during DJF (up to +2°C, Figure 7b). These results highlighted the sensitivity of near-surface air temperature to the seasonal the cycle of LAI.

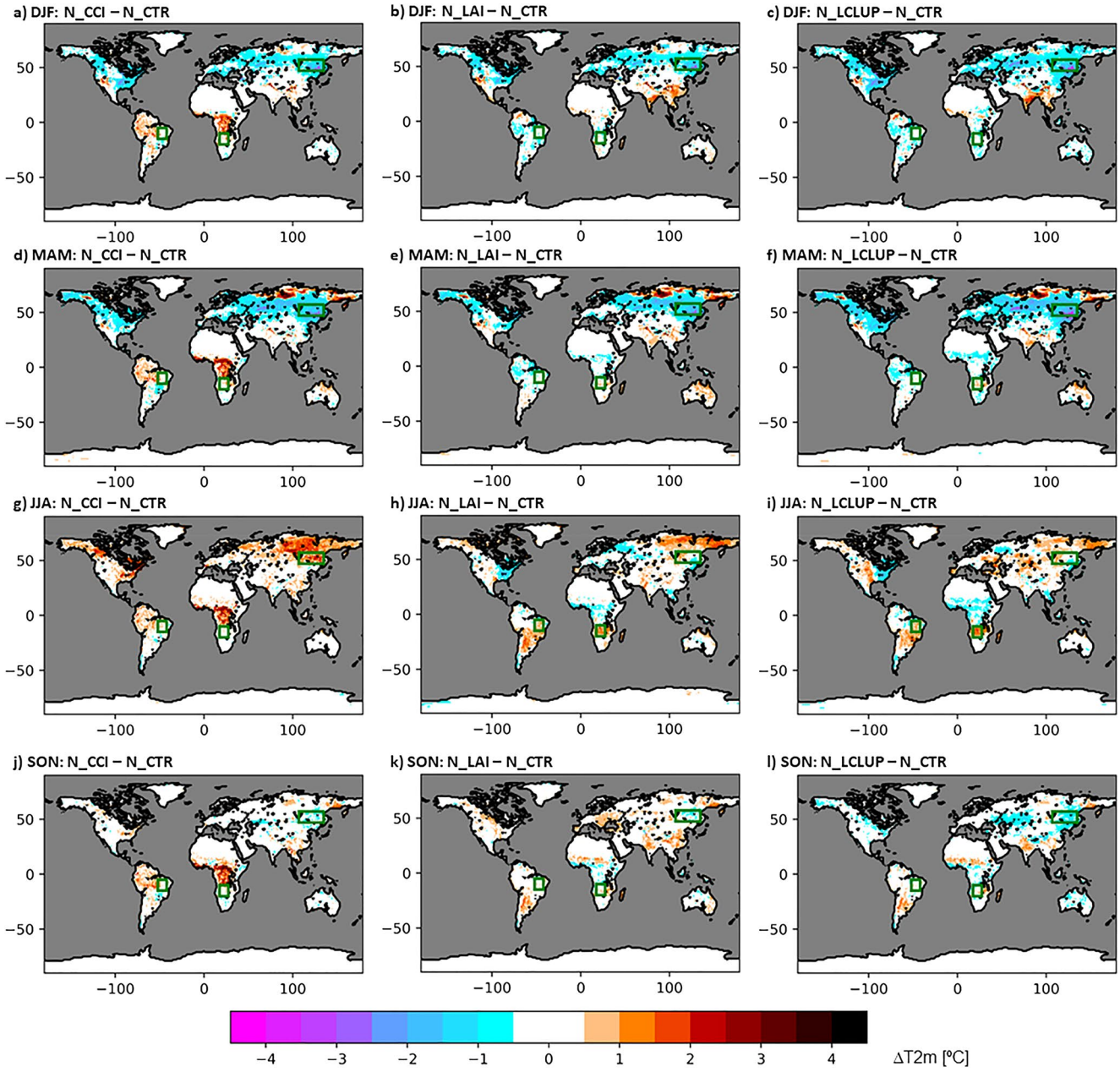

Figure 7. Same as Figure 5 but for 2-m air temperature.

Finally, the sensitivity of simulated T_{2m} to the implementation of low vegetation clumping in the coupled experiments was shown in Figures 7c, 7f, 7i and 7l. N_LCLUP resulted in slightly increased cooling magnitude and areal extent over northern Eurasia and northern North America during DJF and MAM, and slightly enhanced warming over southern Asia during DJF, as well as over western United States, western Asia and Eastern Europe, eastern Brazil, and part of southern Africa during JJA.

3.6. Global Atmospheric Response to CHTESSEL Vegetation Updates

We proceeded to assess the overall performance of the coupled IFS experiments. This was achieved by computing the RMSE at different pressure levels in three regions (NH, TROP, and SH) using ERA5 as reference. The main goal was to evaluate the impact of the large-scale temperature change patterns (e.g., the

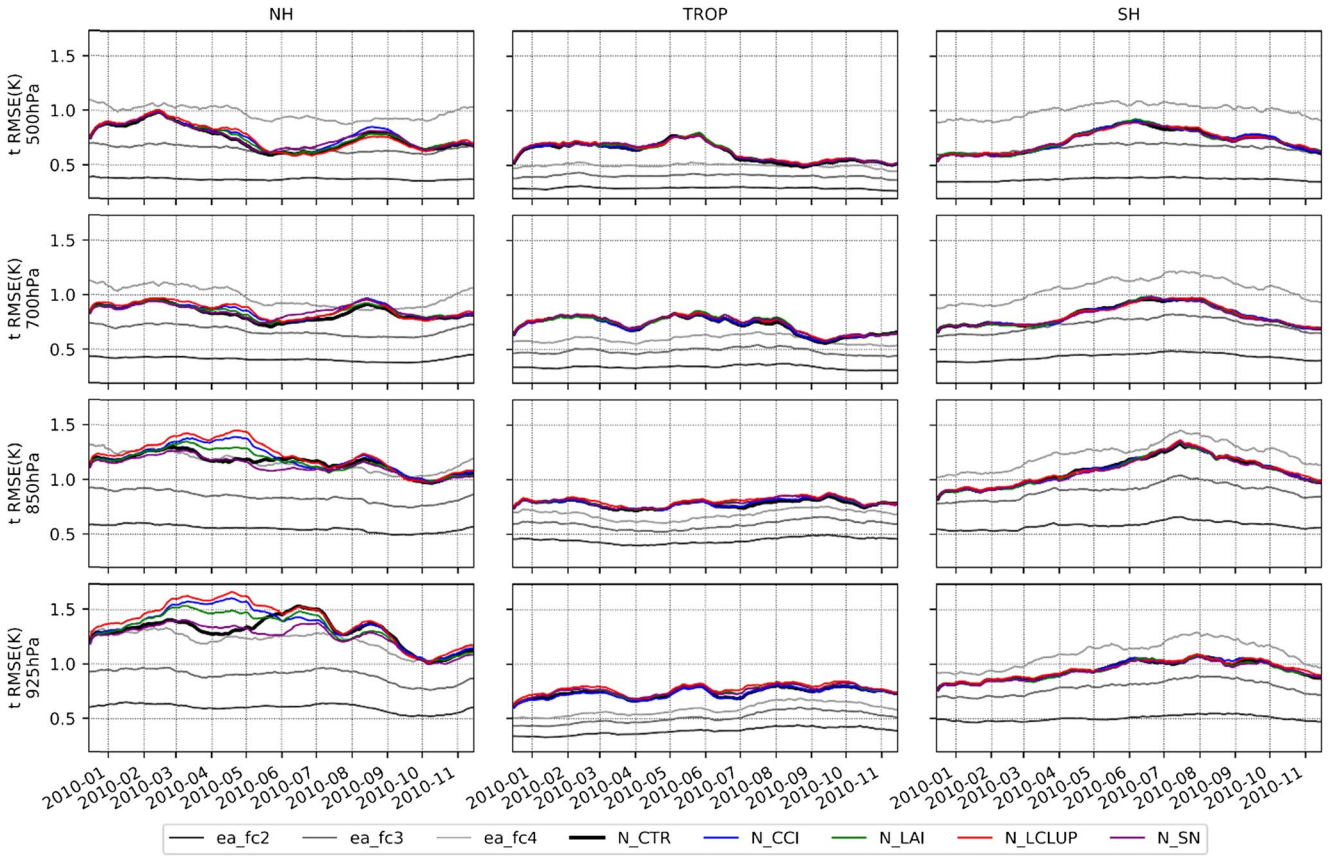

Figure 8. Mean root mean square difference in respect to ERA5 of temperature at different pressure levels (from top to bottom: 500, 700, 850, and 925 hPa), and different regions (from left to right: extratropical Northern Hemisphere region [NH], Tropical region [TROP], and extratropical Southern Hemisphere [SH]). The daily RMSE curves are smoothed with a 30-day running window. The gray curves correspond to ERA5 day 2, 3, and 4 forecasts (ea_fc2, ea_fc3, and ea_fc4) and the experiments N_CTR (black), N_CCI (blue), N_LAI (green), N_LCLUP (red), and N_SN (purple).

DJF and MAM cooling and JJA warming over the northern latitudes) on the upper levels of the simulated atmosphere.

The N_CTR simulations had errors comparable with day 4 forecasts for temperature (Figure 8) and humidity (Figure S11), while geopotential height showed errors equivalent to the day 2–3 forecast range for SH and NH region, and higher than day 4 for TROP (Figure S12). For wind speed, the errors were similar to a day 2 forecast range at 500 hPa increasing to day 3 near the surface (Figure S13). This was expected due to the nudging of vorticity from 800 hPa upwards that keeps the large-scale circulation close to ERA5. Despite the errors in temperature and humidity, consequence of free evolving state in the simulations, a performance comparable to a day 4 short-range weather forecast was considered reasonable to assess the large-scale impact of the model changes.

In TROP and SH, there was no difference among the different experiments for the temperature RMSE (Figure 8). For wind-speed and humidity, all experiments were similar in the three regions. Notice, however, that this did not guarantee that the model changes were neutral in terms of large-scale circulation since this was constrained by the nudging.

The main feature of the atmospheric impact associated with the proposed CHTESSEL vegetation updates occurred in the NH region, corresponding to an increase of the temperature errors in spring (MAM) at 925 and 850 hPa (Figure 8). These increased errors emerged in N_CCI, with a slight reduction in N_LAI, and increase with N_LCLUP. These temperature errors were also reflected in the geopotential height errors (Figure S12) at 500 hPa, despite the vorticity nudging at those levels.

A closer focus near the surface and considering only land points (cf. T_{2m} and T_{925hPa} RMSE over the NH region in Figure S14) showed that N_CCI started to deviate from N_CTR by mid-January returning to errors close to N_CTR by the end of May. These errors were introduced by the widespread cooling over northern Eurasia and North America resulting from the implemented vegetation updates (cf. Sections 3.3 and 3.4). Indeed, experiment N_SN, where the vegetation shadowing effect over snow was turned off resulted in an overall reduction of the NH atmospheric temperature errors during spring (Figure 8, discussed in more detail in the following section).

3.7. Issues Raised by the CHTESSEL Vegetation Updates

This section presents a detailed regional analysis on three sub-regions (represented by the green boxes in Figures 6 and 7) deemed to be representative of the improvements and deteriorations introduced by the implemented updates to CHTESSEL vegetation coverage. The first sub-region, located in eastern Brazil (eBrz, spanning between 15°S and 5°S, and 52°W and 42°W), experienced excessive surface warming during JJA (cf. Figure 3). The second sub-region, located in eastern Russia (eRus, spanning between 47°N and 57°N, and 107°E and 135°E) was deemed to be representative of the widespread cooling in surface and near-surface air temperatures over northern Eurasia and northern North America during winter and spring (cf. Figures 6 and 7). The third sub-region, located over southern Africa (sAfr, spanning between 20°S and 10°S, and 17.5°E and 27.5°E), experienced a strong reduction in LST errors during JJA.

Figure 9a showed that the JJA excessive warming in daily maximum LST over eBrz was induced in N_CCI (causing an increase up to +1.4°C compared to N_CTR) and amplified after updating the LAI (causing an increase up to +2.1°C compared to N_CTR). Furthermore, introducing the low vegetation clumping resulted in further amplification of the warming effect over eBrz (causing an increase up to +3.5°C compared to N_CTR).

The temperature amplification in N_LAI and N_LCLUP was associated with decreased surface latent heat fluxes up to −13 Wm^{-2} (Figure 9m), and increased surface sensible heat fluxes by +6 Wm^{-2} (Figure 9p). N_CCI, N_LAI, and N_LCLUP also displayed increases to the average 10-m wind speed over June–September up to +0.5 m/s (Figure 9j).

Notice that, in N_CTR, the eBrz region had, respectively, 31% and 67% of low and high vegetation cover. The high vegetation was dominated by the interrupted forest type (see Figures 1a–1c). With the ESA-CCI land cover update the region changed to, respectively, 70% and 20% of low and high vegetation, mainly dominated by "crops" (32%) and "evergreen shrubs" (28%). The change from high to low vegetation explained the increased near-surface wind-speed due to a reduction in roughness length (see Table S1). The decrease in latent heat flux likely resulted from both the change from high to low vegetation, in particular the increase of shrubs coverage, as well as a decrease of LAI from experiment N_CTR to experiment N_LAI. Specifically, the "Shrubs" vegetation type had a higher minimum stomatal resistance (225 s/m) than the other low vegetation types (mostly 100 s/m) and, also, than the original dominant interrupted forest (175 s/m). To test the impact of the association of model parameters with vegetation type, in the experiment N_TG we replaced all the parameters for the "Shrubs" vegetation type by those of "Tall Grass" (cf. Table S1). This change of the parameters had the same effect as modifying the cross-walking table so that shrublands in ESA-CCI were classified as "Tall Grass." Such change was justified by the uncertainty related in the cross-walking table used to convert the ESA-CCI vegetation types into the Biosphere Atmosphere Transfer Scheme vegetation classes used in CHTESSEL. The results showed that the daily maximum LST over eBrz in N_TG was within ±1°C of N_CTR throughout the full diurnal cycle (Figure 9a), thus reducing the warm bias introduced by the vegetation types and fractions and LAI updates. This bias reduction was accompanied by a reduction in 10-m wind speed when compared with N_LAI due to the increased roughness lengths (from 0.1 m in "Shrubs" to 0.47 m in 'Tall Grass", Figure 9j). Additionally, N_TG showed increased evaporation (Figure 9m) compared to N_LAI due to decreased rsmin (from 225 s/m in "Shrubs" to 100 s/m in "Tall Grass"). Figure 9d showed a relatively small impact of the warm bias induced by the vegetation updates to the T_{2m}, with an overall maximum value of +0.5°C. This impact was even smaller for the temperature at 850 hPa (Figure 9g, maximum of +0.3°C), with the changes induced by the vegetation updates being smaller than the difference between ERA5 and the control coupled simulation.

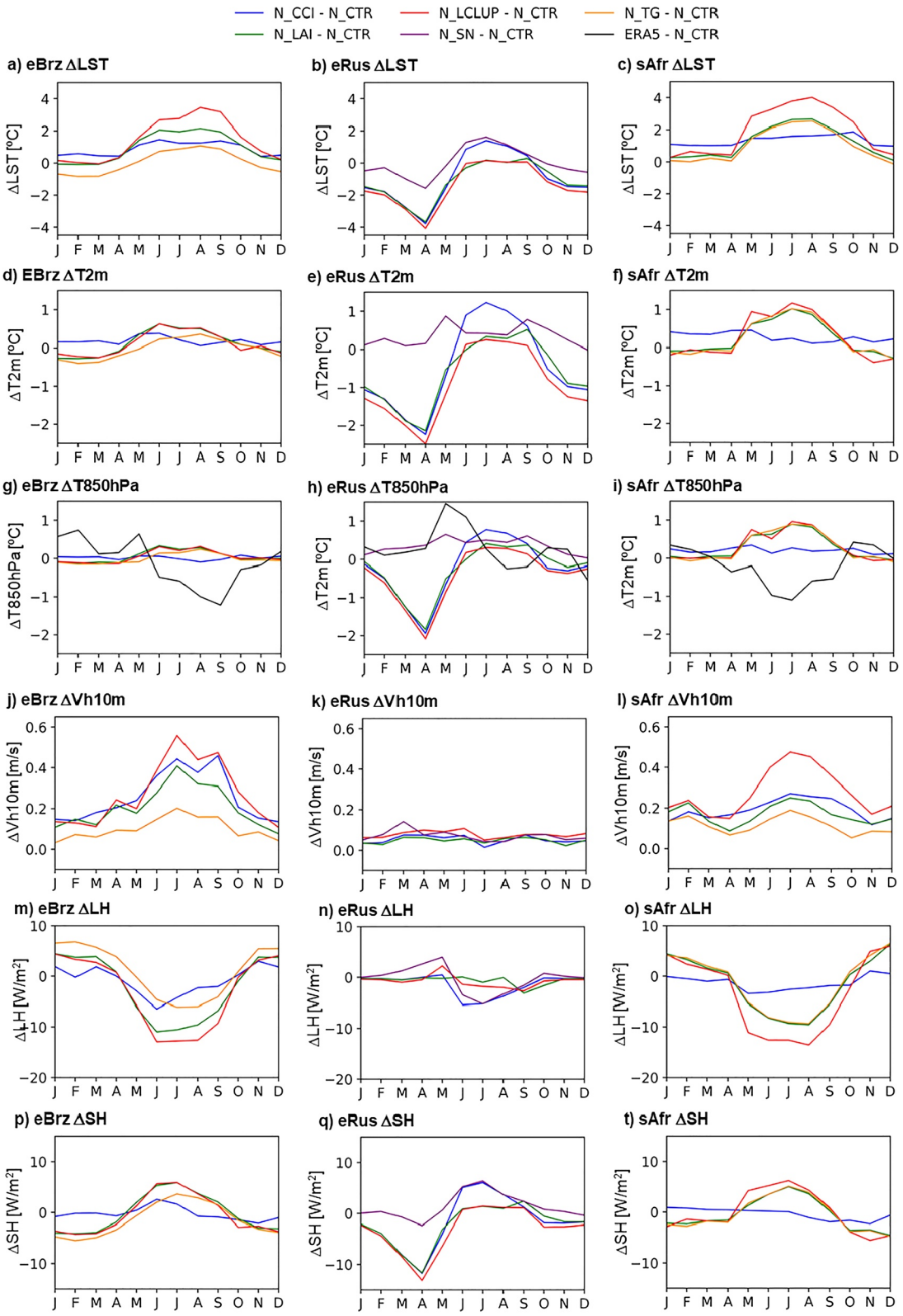

Figure 9. Average seasonal cycle over eBrz (left column) and eRus (right column) sub-domains between coupled simulations with modified vegetation and N_CTR. From top to bottom, the panels represent the seasonal cycles of daily maximum LST (a and b); daily maximum T_{2m} (c and d); daily mean T850 hPa (e and f); daily mean vh10m (g and h); daily mean latent heat flux (i and j); and daily mean sensible heat flux (k and l). The blue lines represent N_CCI–N_CTR, green is N_LAI–N_CTR, red is N_LCLUP–N_CTR, orange is N_TG–N_CTR, purple is N_SN–N_CTR, and black is ERA5–N_CTR.

The large systematic cooling of daily maximum LST over eRus in all simulations (compared to CTR) occurred mostly during January–April (JFMA), peaking in April with a value of −4.1°C (Figure 9b), while a warming effect was found during JJA (maximum of +1.5°C). The JFMA cooling was clearly reflected into daily maximum T_{2m} (Figure 9e, with differences no N_CTR up to −2.5°C) and T_{850hPa} (Figure 9h, with differences no N_CTR up to −2.1°C). Figure 9a showed that this cooling was introduced in N_CCI, while updating LAI and implementing low vegetation clumping had negligible additional impacts during JFMA (but reduced the warming in JJA). Figure 9q showed that the JFMA cooling was tightly linked with a large reduction of the sensible heat fluxes (up to −13 Wm^{-2}). In contrast, only small impacts of the implemented vegetation updates over eRus were found for near-surface wind speed (Figure 9k, below 0.1 m/s) and for evaporation/sublimation (Figure 9n, below 5 Wm^{-2}).

The experiment N_SN, where the vegetation shadowing effect over snow was turned off demonstrated that the widespread cooling of the northern latitudes was due to the model representation of the shading effect of vegetation over snow-covered land, which modulated the surface albedo. This was because the decreased high-vegetation cover after the ESA-CCI-based updated vegetation fractions and types (Figure 1) reduced the snow shading effect, that is, it exposed more snow. In turn, this resulted in an increased surface albedo and decreased surface sensible heat flux (Figure 9q, differences between N_SN and N_CTR below 3 Wm^{-2}), as well as daily maximum LST (Figure 9b, differences between N_SN and N_CTR below 1.4°C), T_{2m} (Figure 9e, differences between N_SN and N_CTR below 0.9°C) and T_{850hPa} (Figure 9h, differences between N_SN and N_CTR below 0.7°C). In fact, the change in T850 hPa between N_SN and N_CTR was of the same order of magnitude as the difference between ERA5 and N_CTR (up to 1.5°C). This represented an improvement to the simulated atmospheric temperature under the assumption that ERA5 represented reality due to data assimilation. The large-scale evaluation of the temperature errors against ERA5 at different pressure levels supported these results, showing that the N_SN experiment (purple lines in Figure 8) removed the increased temperature errors for March–May that were found for N_CCI and N_LAI when averaged over the entire NH region. Moreover, between May and June, N_SN had lower temperature errors than N_CTR at 925 hPa and 850 hPa (also in 2-m temperature, see Figure S14), showing the potential improvements of the land cover and LAI updates in the coupled model.

However, we point out that removing the vegetation snow shadowing effect greatly amplifies the warming over northern Eurasia and portions of northwestern America during warmer months, which was first introduced in N_CCI (Figure S15). This effect indicated that the surface vegetation parameters and interaction with the snowpack in this region requires further investigations. It is also important to assess the impact of these changes together with the recently developed multilayer snow scheme for the ECMWF model (Arduini et al., 2019). The reduced thermal inertia of the multilayer snow scheme is likely to interact with the handling of shaded snow albedo locally, but also regionally, considering the large extent of shaded snow regions in Eurasia.

The differences in the response of surface turbulent fluxes between eBrz and eRus reflected the well-known importance of these fluxes to the simulated atmospheric temperatures. Additionally, a global comparison of the sensitivity of daily maximum LST (Figure 6), and daily maximum T_{2m} (Figure 7) with the sensitivity patterns of latent and sensible heat fluxes (Figures S16 and S17, respectively) highlighted the key role of model vegetation to the modulation of heat and moisture exchanges between the surface and the atmosphere, with important impacts on the model climate.

The N_CCI simulation resulted in increased daily maximum LST compared to N_CTR over all months over sAfr (Figure 9c, differences ranging between +1.0°C and +1.9°C), which was responsible for a reduction of the observed strong cold bias in N_CTR over sAfr during MAM, JJA, and SON (cf. Figure S1, notice that during DJF there was not sufficient valid LSA-SAF data points to quantify the error). This improvement was related to large changes in low and high vegetation fractions introduced by the CCI-based updated: in N_CTR the sAfr region had, respectively, 43% and 56% of low and high vegetation cover, and with the ESA-CCI land cover update these changed to, respectively, 58% and 30% for low and high vegetation (Figures 1a and 1b). Additionally, there was also a change in the dominant low vegetation type, from tall grass (40%) in N_CTR to short grass (43%) in N_CCI.

Updating LAI resulted in increased seasonality of the warming introduced by the updated vegetation (Figure 9c), with no differences to N_CTR during DJF, and an increasing and decreasing cycle between MAM

and SON, peaking during JJA (with a maximum difference to N_CTR of +2.7°C). This seasonality was further enhanced in N_LCLUP (Figure 9c), with differences to N_CTR reaching +4.0°C during JJA. Notice that the increased warming during MAM, JJA, and SON corresponded to reductions of daily maximum LST systematic errors (cf. Figure S2).

The simulation N_TG had a negligible impact over the sAfr region as there was no shrubs coverage in the region (neither in CTR nor CCI). The increased seasonality and peak JJA warming for N_LAI and N_LCLUP are reflected in daily maximum T_{2m} (Figure 9f) and T_{850hPa} (Figure 9i) but with differences below +1°C compared to N_CTR (for T_{850hPa}, these differences are lower than the differences between ERA5 and N_CTR). Low vegetation clumping introduced an increase in Vh10 during JJA up to +0.5 m s^{-1} compared to N_CTR, which was nearly double the difference found for N_CCI and N_LAI. The increased seasonality and JJA warm peaks introduced in N_LAI (and enhanced in N_LCLUP) was associated with increased seasonality and in LH (Figure 9o, decreasing in JJA up to −14 Wm^{-2}) and in SH (Figure 9r, increasing in JJA up to +6 Wm^{-2}).

4. Discussion

4.1. Reconciling the LSA-SAF and FLUXNET Evaluations

The updates to CHTESSEL vegetation cover had an overall positive impact in the daily maximum LST estimated by coupled and offline simulations, with a reduction in biases over 17% and 16% of the grid points with valid observations, respectively. The largest bias reductions occurred over southern Europe and southern and central Africa. The daily maximum LST ubRMSE was also reduced over southern and central Africa, particularly during JJA and SON (with ubRMSE reductions up to −5°C).

However, the modifications showed mostly neutral impacts on the simulated surface latent and sensible heat fluxes when compared against 51 FLUXNET stations over 1996–2014. The reason for the lack of improvement in surface turbulent fluxes may be primarily attributed to the fact that FLUXNET towers were concentrated in central Europe and North America. Over central Europe, the changes to daily maximum LST errors associated with the vegetation updates were mostly within observational uncertainty (with magnitudes below 2°C, cf. Figure 3), which agrees with neutral impact of surface fluxes. This was corroborated by the coherent changes in daily maximum LST and the sensible and latent heat fluxes over sAfr region, where the updates had a positive impact. North America falls outside the MSG disk and, thus, a detailed analysis of the errors was not possible in the present work. Nonetheless, Figure S18 showed that the location of 51 towers is skewed toward energy-limited regimes, where the influence of land conditions was expected to be lower. Furthermore, it is worth noting the existence of the pixel-to-tower representativeness errors. This limited the interpretation of the results, as the land cover and LAI at the model grid-point scale (which also depended on the resolution) can be different from the actual local land and vegetation characteristics in the vicinity of the tower. Moreover, issues affecting the energy closure at eddy covariance sites may have induced systematic errors in the in situ measurements. In fact, latent heat flux is often underestimated by eddy covariance measurements (see e.g., Ershadi et al., 2014). These systematic errors in turbulent heat flux measurements can be accounted for by applying corrections based on assumptions regarding the source of this energy imbalance (Michel et al., 2016). However, those correction techniques can have potentially adverse effects on the actual fluxes being examined (Foken, 2008) and were not considered in this study. Nonetheless, we point out that the biases in latent heat flux and the relatively low anomaly correlations found in Section 3.3 indicated that there is still room to improve the evaporation formulation in CHTESSEL (e.g., canopy resistance formulation, soil moisture stress, etc.), despite recent reports of a marked improvement in ERA5 with respect to its predecessor, ERA Interim (Martens et al., 2020).

4.2. Open Regional Issues and Future Directions

Despite the overall positive impact, we found two regions where revised land cover and vegetation had a negative impact: the daily maximum LST bias increased up to +6°C in eastern Brazil during JJA, and a cooling signal (up to −7°C for daily maximum LST and to −4°C for daily maximum T_{2m}) emerged over wide portions of northern Eurasia during January–April.

On the one hand, we showed that the increased errors over eastern Brazil were primarily associated with the land cover revision and related uncertainty in vegetation classes cross-walking table. On the other hand, the errors over northern Eurasia were related to changes in surface albedo. These were induced by a reduction of high vegetation cover, exposing larger areas of snow that have a higher albedo than regions with snow under forests. We suggest several options to address these shortcomings, namely: (a) a detailed regional analysis, (b) a careful evaluation of the ESA-CCI/C3S data set and of the cross-walking table employed to convert between ESA-CCI/C3S and the model land cover classes, and (c) a recalibration of several model parameters closely related to land cover and vegetation (e.g., rsmin, $z0_h$, and $z0_m$), using a more extensive evaluation and multiple criteria. In fact, the calibration of model parameters is paramount, as the current model configuration has been tuned based on an outdated land cover data set. As an example, Stevens et al. (2020) investigated the close relationship between LAI and the minimum stomatal resistance and rooting depth in CHTESSEL showing some potential to improve the latent heat flux variability via an adjustment of these parameters. However, those results were based on offline simulations for 17 flux towers. Extending it globally, and within a coupled model perspective, represents a significant scientific and computational effort. It would require a large number of forecast and data assimilation experiments focusing on the impact on medium- to long-range forecasts and taking into consideration compensating errors related to the land-atmosphere interaction. Nonetheless, perspectives for representing interannual variability for both land-cover and vegetation states will be explored thanks to the ESA-CCI, C3S, and CGLS data sets availability over several decades. This will also require a more extensive evaluation using multiple assessment criteria, reference data sets and variables (e.g., Collier et al., 2018). Our results also showed that land cover classification and LAI cannot be treated independently from the model parameterization and parameters, as there are several assumptions in the model that are associated to the input land cover and LAI data provided.

Finally, we point out that the implemented changes to CHTESSEL vegetation had a negligible impact on daily maximum LST over wide areas of the MSG disk. These included the western Africa and Arabian Peninsula regions, which showed strong warm biases during JJA (up to +10°C, Figure 3). These systematic errors were likely related to aerosols affecting both the satellite retrievals and IFS radiation. Indeed, during JJA these regions have a significant aerosol load (dust) resulting in high aerosols optical depth (AOD) (Voss & Evan, 2020). The current algorithm used by LSA-SAF to derive LST does not consider explicitly AOD, which could lead to an underestimation of LST as the signal measured at the sensor is inherently influenced by AOD (e.g., Rosas et al., 2017). Therefore, this warm bias can result from an actual overestimation by the IFS but may also relate to an underestimation of the LSA-SAF LST. Further research is needed to understand and quantify the potential role of AOD on the LST satellite retrievals. But a more profound analysis of this issue is beyond the scope of the present work.

5. Conclusions

The present work demonstrated the existence of large magnitude biases (up to ~10°C, either positive or negative depending on region) in daily maximum LST simulated by the uncoupled CHTESSEL simulations and coupled IFS simulations. The comparison was performed over the MSG disk using the LSA-SAF LST product as the reference, which was limited to clear-sky conditions. Error magnitudes were particularly high over southern Europe, southern and central Africa, and Brazil.

Building upon recent work (JO19 and NO20) we updated the global vegetation characterization in CHTESSEL leveraging on state-of-the-art EO products. Namely, we updated the land cover, including vegetation types and fractions based on the ESA-CCI land cover data set, and the LAI based on the GCLS product. Moreover, a clumping parameterization was also tested, introducing seasonality in the effective low vegetation coverage which was currently neglected in the model. The implemented improvements to the vegetation cover resulted in an overall positive impact on the simulated daily maximum LST, with the area where the daily maximum LST bias decreased being more than double that where it increased (17% compared to 8% in coupled simulations). The ubRMSE also supported the overall improvement resulting from the implemented vegetation coverage updates (with a domain average reduction over all seasons ranging between −0.11°C and −0.06°C). Moreover, the proposed updates had a neutral impact on daily minimum LST.

The present results may have important implications to the surface and near-surface fields in ECMWF forecasts and reanalysis, which are based on IFS simulations using the CHTESSEL LSM. However, two issues emerged which will require further attention prior to the operational implementation of the proposed vegetation updates. The first was the degraded daily maximum LST over eastern Brazil, which was primarily associated with the land cover revision and related uncertainty in vegetation classes cross-walking table. The second was the widespread cooling of northern latitudes during winter and spring, at odds with ERA5 air temperatures (benefiting from data assimilation). Our analysis revealed that this issue was induced by a reduction of high vegetation cover exposing larger areas of snow that have a higher albedo than regions with snow under forests. We suggest several options to address these shortcomings, but an operational implementation of the proposed vegetation updates still requires further work, namely: (a) a detailed regional analysis, (b) a careful evaluation of the ESA-CCI/C3S data set and of the cross-walking table employed to convert between ESA-CCI/C3S and the model land cover classes, and c) a recalibration of several model parameters closely related to land cover and vegetation (e.g., rsmin, $z0_h$, and $z0_m$).

Finally, the analysis of the different experiments considered here, together with the detailed regional analysis over southern Africa, eastern Brazil, and eastern Russia, clearly illustrated the well-known importance of the accurate representation of vegetation for the accurate simulation of land-atmosphere interactions (Miralles et al., 2019; Seneviratne et al., 2010). At the same time, this study also highlighted the complexity of the land-atmosphere coupled system, with very distinct levels of atmospheric sensitivity to changes in vegetation coverage.

Data Availability Statement

ERA5 data can be obtained freely from the Copernicus Climate Change Service Information website (https://climate.copernicus.eu/). The CGLS LAI can be obtained freely from the CGLS website (https://land.copernicus.eu/). The ESA-CCI land cover can be obtained freely from their website (https://www.esa-landcover-cci.org/). The LSA-SAF LST can be obtained freely from their website (https://landsaf.ipma.pt/en/data/catalogue/). The FLUXNET 2015 data can be accessed from the FLUXNET data portal at https://fluxnet.fluxdata.org/data/fluxnet2015-dataset/ (FLUXNET, 2020; Pastorello et al., 2020). The simulations carried out in this study are available from ECMWF data archive (required login at: https://apps.ecmwf.int/mars-catalogue/?class=rd) with the following research experimental IDs: O_CTR is a04d, O_CCI is a04e, O_CCI is a04f, O_LCLUP is a04k, O_TG is a04s, O_SN is a04m, N_CTR is a046, N_CCI is a048, N_LAI is a04b, N_LCLUP is a04l, N_TG is a04t, and N_SN is a04n.

Acknowledgments
This research was funded by the Fundação para a Ciência e a Tecnologia (FCT) grant number PTDC/CTA-MET/28946/2017 (CONTROL). The authors would also like to acknowledge the financial support of FCT through project UIDB/50019/2020—IDL. The authors would like to thank Anton Beljaars and three anonymous reviewers for the interesting comments and discussion on our work. Diego G. Miralles acknowledges support from the European Research Council (ERC) (grant no. 715254, DRY–2–DRY) and the Belgian Science Policy Office (BELSPO) (grant no. SR/00/373, ALBERI). Acknowledgment is made for the use of ECMWF's computing and archive facilities in this research. This work used eddy-covariance data acquired and shared by the FLUXNET community, including the following networks: AmeriFlux, AfriFlux, AsiaFlux, CarboAfrica, CarboEuropeIP, CarboItaly, CarboMont, ChinaFlux, Fluxnet-Canada, GreenGrass, ICOS, KoFlux, LBA, NECC, OzFlux-TERN, TCOS-Siberia, and USCCC. The FLUXNET eddy-covariance data processing and harmonization were carried out by the ICOS Ecosystem Thematic Center, AmeriFlux Management Project and Fluxdata project of FLUXNET, with the support of CDIAC, and the OzFlux, ChinaFlux, and AsiaFlux offices.

References

Albergel, C., Dutra, E., Bonan, B., Zheng, Y., Munier, S., Balsamo, G., et al. (2019). Monitoring and forecasting the impact of the 2018 summer heatwave on vegetation. *Remote Sensing*, *11*, 520. https://doi.org/10.3390/rs11050520

Alessandri, A., Catalano, F., De Felice, M., Van Den Hurk, B., Doblas Reyes, F., Boussetta, S., et al. (2017). Multi-scale enhancement of climate prediction over land by increasing the model sensitivity to vegetation variability in EC-Earth. *Climate Dynamics*, *49*, 1215–1237. https://doi.org/10.1007/s00382-016-3372-4

Anderson, M. C., Norman, J. M., Kustas, W. P., Li, F., Prueger, J. H., & Mecikalski, J. R. (2005). Effects of vegetation clumping on two-source model estimates of surface energy fluxes from an agricultural landscape during SMACEX. *Journal of Hydrometeorology*, *6*, 892–909. https://doi.org/10.1175/JHM465.1

Arduini, G., Balsamo, G., Dutra, E., Day, J. J., Sandu, I., Boussetta, S., & Haiden, T. (2019). Impact of a multi-layer snow scheme on near-surface weather forecasts. *Journal of Advances in Modeling Earth Systems*, *11*(12), 4687–4710. https://doi.org/10.1029/2019MS001725

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*, 2038. https://doi.org/10.3390/rs10122038

Balsamo, G., Viterbo, P., Beljaars, A., van den Hurk, B., Hirschi, M., Betts, A. K., & Scipal, K. (2009). A revised hydrology for the ECMWF model: Verification from field site to terrestrial water storage and impact in the integrated forecast system. *Journal of Hydrometeorology*, *10*, 623–643. https://doi.org/10.1175/2008JHM1068.1

Beck, H. E., Pan, M., Roy, T., Weedon, G. P., Pappenberger, F., van Dijk, A. I. J. M., et al. (2019). Daily evaluation of 26 precipitation datasets using Stage-IV gauge-radar data for the CONUS. *Hydrology and Earth System Sciences*, *23*, 207–224. https://doi.org/10.5194/hess-23-207-2019

Beljaars, A. C. (2020). *Towards optimal parameters for the prediction of near surface temperature and dewpoint* (ECMWF Technical Memoranda 868). https://doi.org/10.21957/yt64x7rth

Beljaars, A. C., Viterbo, P., Miller, M. J., & Betts, A. K. (1996). The anomalous rainfall over the United States during July 1993: Sensitivity to land surface parameterization and soil moisture anomalies. *Monthly Weather Review*, *124*, 362–383. https://doi.org/10.1175/1520-0493(1996)124<0362:TAROTU>2.0.CO;2

Best, M. J., Abramowitz, G., Johnson, H. R., Pitman, A. J., Balsamo, G., Boone, A., et al. (2015). The plumbing of land surface models: Benchmarking model performance. *Journal of Hydrometeorology*, *16*(3), 1425–1442. https://doi.org/10.1175/JHM-D-14-0158.1

Bonan, G. B. (2008). Forests and climate change: Forcings, feedbacks, and the climate benefits of forests. *Science*, *320*(5882), 1444–1449. https://doi.org/10.1126/science.1155121

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., Beljaars, A., Panareda, A. A., Calvet, J. C., Jacobs, C., et al. (2013). Natural land carbon dioxide exchanges in the ECMWF Integrated Forecasting System: Implementation and offline validation. *Journal of Geophysical Research: Atmospheres*, *118*(12), 5923–5946. https://doi.org/10.1002/jgrd.50488

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*(15), 111–126. https://doi.org/10.1016/j.rse.2015.03.009

Collier, N., Hoffman, F. M., Lawrence, D. M., Keppel-Aleks, G., Koven, C. D., Riley, W. J., et al. (2018). The International Land Model Benchmarking (ILAMB) system: Design, theory, and implementation. *Journal of Advances in Modeling Earth Systems*, *10*, 2731–2754. https://doi.org/10.1029/2018MS001354

Defourny, P., Boettcher, M., Bontemps, S., Kirches, G., Krueger, O., Lamarche, C., et al. (2014). *Algorithm theoretical basis document for land cover climate change initiative* (Technical Report). European Space Agency.

Delworth, T., & Manaba, S. (1993). Climate variability and land-surface processes. *Advances in Water Resources*, *16*, 3–20. https://doi.org/10.1016/0309-1708(93)90026-C

de Noblet-Ducoudré, N., Boisier, J. P., Pitman, A., Bonan, G. B., Brovkin, V., Cruz, F., et al. (2012). Determining robust impacts of land-use-induced land cover changes on surface climate over North America and Eurasia: Results from the first set of LUCID experiments. *Journal of Climate*, *25*(9), 3261–3281. https://doi.org/10.1175/JCLI-D-11-00338.1

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. *The Quarterly Journal of the Royal Meteorological Society*, *139*, 1199–1213. https://doi.org/10.1002/qj.2023

Dirmeyer, P. A. (2003). The role of the land surface background state in climate predictability. *Journal of Hydrometeorology*, *4*, 599–610. https://doi.org/10.1175/1525-7541(2003)004<0599:TROTLS>2.0.CO;2

Dutra, E., Balsamo, G., Viterbo, P., Miranda, P. M., 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

ECMWF. (2019). IFS documentation CY46R1—Part IV: Physical processes. In *IFS documentation CY46R1* (IFS Documentation). https://doi.org/10.21957/xphfxep8c

Ermida, S. L., Trigo, I. F., DaCamara, C. C., Jiménez, C., & Prigent, C. (2019). Quantifying the clear-sky bias of satellite land surface temperature using microwave-based estimates. *Journal of Geophysical Research: Atmospheres*, *124*, 844–857. https://doi.org/10.1029/2018JD029354

Ershadi, A., Mccabe, M. F., Evans, J. P., Chaney, N. W., & Wood, E. F. (2014). Multi-site evaluation of terrestrial evaporation models using FLUXNET data. *Agricultural and Forest Meteorology, Agricultural and Forest Meteorology*, *187*, 46–61. https://doi.org/10.1016/j.agrformet.2013.11.008

Foken, T. (2008). The energy balance closure problem: An overview. *Ecological Applications*, *18*(6), 1351–1367. https://doi.org/10.1890/06-0922.1

Freitas, S. C., Trigo, I. F., Bioucas-Dias, J. M., & Goettsche, F.-M. (2010). Quantifying the uncertainty of land surface temperature retrievals from SEVIRI/Meteosat. *IEEE Transactions on Geoscience and Remote Sensing*, *48*, 523–534. https://doi.org/10.1109/TGRS.2009.2027697

Gevaert, A. I., Miralles, D. G., De Jeu, R. A. M., Schellekens, J., & Dolman, A. J. (2018). Soil moisture-temperature coupling in a set of land surface models. *Journal of Geophysical Research: Atmospheres*, *69*(2), 1481–1498. https://doi.org/10.1002/2017JD027346

Ghent, D., Kaduk, J., Remedios, J., Ardö, J., & Balzter, H. (2010). Assimilation of land surface temperature into the land surface model JULES with an ensemble Kalman filter. *Journal of Geophysical Research*, *115*, D19112. https://doi.org/10.1029/2010jd014392

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

Guo, Z., Dirmeyer, P. A., DelSole, T., & Koster, R. D. (2012). Rebound in atmospheric predictability and the role of the land surface. *Journal of Climate*, *25*, 4744–4749. https://doi.org/10.1175/JCLI-D-11-00651.1

Hauser, M., Orth, R., & Seneviratne, S. I. (2016). Role of soil moisture vs. recent climate change for heat waves in western Russia. *Geophysical Research Letters*, *43*, 2819–2826. https://doi.org/10.1002/2016GL068036

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

Hurk, B., Doblas-Reyes, F., Balsamo, G., Koster, R. D., Seneviratne, S. I., & Camargo, H. (2010). Soil moisture effects on seasonal temperature and precipitation forecast scores in Europe. *Climate Dynamics*, *38*(1–2), 349–362. https://doi.org/10.1007/s00382-010-0956-2

Jarvis, P. G. (1976). The interpretation of the variations in leaf water potential and stomatal conductance found in canopies in the field. *Philosophical Transactions of the Royal Society of London B Biological Sciences*, *273*, 593–610. https://doi.org/10.1098/rstb.1976.0035

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*, 2570. https://doi.org/10.3390/rs11212570

Koster, R. D., Dirmeyer, P. A., Guo, Z., Bonan, G., Chan, E., Cox, P., et al. (2004). Regions of strong coupling between soil moisture and precipitation. *Science*, *305*, 1138–1140. https://doi.org/10.1126/science.1100217

Kumar, S., Dirmeyer, P. A., Merwade, V., DelSole, T., Adams, J. M., & Niyogi, D. (2013). Land use/cover change impacts in CMIP5 climate simulations: A new methodology and 21st century challenges. *Journal of Geophysical Research: Atmospheres*, *118*, 6337–6353. https://doi.org/10.1002/jgrd.50463

Le Moigne, P. (2018). *SURFEX scientific documentation, V8.1*. Retrieved from http://www.umr-cnrm.fr/surfex/IMG/pdf/surfex_scidoc_v8.1.pdf

Li, Z.-L., Tang, B.-H., Wu, H., Ren, H., Yan, G., Wan, Z., et al. (2013). Satellite-derived land surface temperature: Current status and perspectives. *Remote Sensing of Environment*, *131*, 14–37. https://doi.org/10.1016/j.rse.2012.12.008

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*, 1303–1330. https://doi.org/10.1080/014311600210191

Martens, B., Miralles, D. G., Lievens, H., van der 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

Martens, B., Schumacher, D. L., Wouters, H., Muñoz-Sabater, J., Verhoest, N. E. C., & Miralles, D. G. (2020). Evaluating the land-surface energy partitioning in ERA5. *Geoscientific Model Development*, *13*(9), 4159–4181. https://doi.org/10.5194/gmd-13-4159-2020

Massari, C., Camici, S., Ciabatta, L., & Brocca, L. (2018). Exploiting satellite-based surface soil moisture for flood forecasting in the Mediterranean area: State update versus rainfall correction. *Remote Sensing*, *10*, 292. https://doi.org/10.3390/rs10020292

Michel, D., Jiménez, C., Miralles, D. G., Jung, M., Hirschi, M., Ershadi, A., et al. (2016). The WACMOS-ET project—Part 1: Tower-scale evaluation of four remote-sensing-based evapotranspiration algorithms. *Hydrology and Earth System Sciences*, *20*(2), 803–822. https://doi.org/10.5194/hess-20-803-2016

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*, 19–35. https://doi.org/10.1111/nyas.13912

Nogueira, M. (2020). Inter-comparison of ERA-5, ERA-Interim and GPCP rainfall over the last 40 years: Process-based analysis of systematic and random differences. *Journal of Hydrology (Wellington North)*, *583*, 124632. https://doi.org/10.1016/j.jhydrol.2020.124632

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*, 3975–3993. https://doi.org/10.5194/gmd-13-3975-2020

Nogueira, M., Lima, D. C., & Soares, P. M. (2020). An integrated approach to project the future urban climate response: Changes to Lisbon's urban heat island and temperature extremes. *Urban Climate*, *34*, 100683. https://doi.org/10.1016/j.uclim.2020.100683

Nogueira, M., & Soares, P. M. M. (2019). A surface modelling approach for attribution and disentanglement of the effects of global warming from urbanization in temperature extremes: Application to Lisbon. *Environmental Research Letters*, *14*, 114023. https://doi.org/10.1088/1748-9326/ab465f

Orth, R., Dutra, E., & Pappenberger, F. (2016). Improving weather predictability by including land-surface model parameter uncertainty. *Monthly Weather Review*, *144*, 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

Pastorello, G., Trotta, C., Canfora, E., Chu, H., Christianson, D., Cheah, Y. W., et al. (2020). The FLUXNET2015 dataset and the ONEFlux processing pipeline for eddy covariance data. *Scientific Data*, *7*(1), 1–27. https://doi.org/10.1038/s41597-020-0534-3

Pipunic, R. C., Walker, J. P., & Western, A. (2008). Assimilation of remotely sensed data for improved latent and sensible heat flux prediction: A comparative synthetic study. *Remote Sensing of Environment*, *112*(4), 1295–1305. https://doi.org/10.1016/j.rse.2007.02.038

Pitman, A. J., de Noblet-Ducoudré, N., Cruz, F. T., Davin, E. L., Bonan, G. B., Brovkin, V., et al. (2009). Uncertainties in climate responses to past land cover change: First results from the LUCID intercomparison study. *Geophysical Research Letters*, *36*(14), L14814. https://doi.org/10.1029/2009GL039076

Poulter, B., MacBean, N., Hartley, A., Khlystova, I., Arino, O., Betts, R., et al. (2015). Plant functional type classification for earth system models: Results from the European Space Agency's Land Cover Climate Change Initiative. *Geoscientific Model Development*, *8*, 2315–2328. https://doi.org/10.5194/gmd-8-2315-2015

Rasmijn, L. M., Van der Schrier, G., Bintanja, R., Barkmeijer, J., Sterl, A., & Hazeleger, W. (2018). Future equivalent of 2010 Russian heatwave intensified by weakening soil moisture constraints. *Nature Climate Change*, *8*(5), 381–385. https://doi.org/10.1038/s41558-018-0114-0

Rivas, M. R., & Stoffelen, A. (2019). Characterizing ERA-Interim and ERA5 surface wind biases using ASCAT. *Ocean Science*, *15*, 831–852. https://doi.org/10.5194/os-15-831-2019

Rosas, J., Houborg, R., & McCabe, M. F. (2017). Sensitivity of Landsat 8 surface temperature estimates to atmospheric profile data: A study using MODTRAN in dryland irrigated systems. *Remote Sensing*, *9*, 988. https://doi.org/10.3390/rs9100988

Schlosser, A., & Dirmeyer, P. (2001). Potential predictability of Eurasian snow cover. *Atmospheric Science Letters*, *2*(1–4), 1–8. https://doi.org/10.1006/asle.2001.0037

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*(3–4), 125–161. https://doi.org/10.1016/j.earscirev.2010.02.004

Shukla, J., & Mintz, Y. (1982). Influence of land-surface evapotranspiration on the Earth's climate. *Science*, *215*, 1498–1501. https://doi.org/10.1126/science.215.4539.1498

Stegehuis, A. I., Vautard, R., Ciais, P., Teuling, A. J., Miralles, D. G., & Wild, M. (2015). An observation-constrained multi-physics WRF ensemble for simulating European mega heat waves. *Geoscientific Model Development*, *8*(7), 2285–2298. https://doi.org/10.5194/gmd-8-2285-2015-supplement

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

Trigo, I. F., Boussetta, S., Viterbo, P., Balsamo, G., Beljaars, A., & Sandu, I. (2015). Comparison of model land skin temperature with remotely sensed estimates and assessment of surface-atmosphere coupling. *Journal of Geophysical Research: Atmospheres*, *120*, 12–96. https://doi.org/10.1002/2015JD023812

Trigo, I. F., Da Camara, C. C., Viterbo, P., Roujean, J. L., Olesen, F., Barroso, C., et al. (2011). The satellite application facility on land surface analysis. *International Journal of Remote Sensing*, *32*, 2725–2744. https://doi.org/10.1080/01431161003743199

Urraca, R., Huld, T., Gracia-Amillo, A., Martinez-de-Pison, F. J., Kaspar, F., & Sanz-Garcia, A. (2018). Evaluation of global horizontal irradiance estimates from ERA5 and COSMO-REA6 reanalyses using ground and satellite-based data. *Solar Energy*, *164*, 339–354. https://doi.org/10.1016/j.solener.2018.02.059

van den Hurk, B. J. J. M., Viterbo, P., Beljaars, A., & Betts, A. K. (2000). *Offline validation of the ERA-40 surface scheme* (ECMWF Technical Memoranda 295).

Verger, A., Baret, F., & Weiss, M. (2014). Near real-time vegetation monitoring at global scale. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *7*, 3473–3481. https://doi.org/10.1109/JSTARS.2014.2328632

Voss, K. K., & Evan, A. T. (2020). A new satellite-based global climatology of dust aerosol optical depth. *Journal of Applied Meteorology and Climatology*, *59*(1), 83–102. https://doi.org/10.1175/JAMC-D-19-0194.1

Wang, W., & Kumar, A. (1998). A GCM assessment of atmospheric seasonal predictability associated with soil moisture anomalies over North America. *Journal of Geophysical Research*, *103*, 28637–28646. https://doi.org/10.1029/1998JD200010