We would like to express our gratitude for the constructive reviews of our manuscript. We have the impression that we greatly improved the quality of our manuscript due to the inputs of the referees. Hereafter we provide an overview of the most important changes in the updated manuscript. After that, we list the comments of referees 1 and 2 in black and a description of the modifications of the manuscript in response to the comment in blue (note that the references made in the responses relate to the unmarked manuscript). After, we display the changes made to the manuscript. Parts that were removed from the manuscript are marked in red and stroke-through, whereas added parts are marked in blue.

There were three major changes to our manuscript:

1) The metrics used to assess the agreement between the model and the observations were calculated over finer climate zones to avoid too heterogeneous groups. The new climate zones are displayed in Fig. 1. The metrics over these climate zones are shown in Figs. 3, 4, and A7.

2) We now evaluate the variables in terms of the root mean squared deviation and the index of agreement (Duveiller et al., 2016), to assess the agreement between the model and the observations more objectively. These data are displayed in Figs. 4 and A7.

3) A Discussion section was added to the manuscript. In the first paragraph, the findings from the Results section are discussed in a broader context. In particular, we focus on how (remaining) biases can further alleviated. The second paragraph discusses why modeled ΔT2M is not in agreement with observations. In the third paragraph we list the limitations of the comparison made in this study. And finally, an outlook on what could be done in future studies of a similar kind is provided in the fourth paragraph.

Referee 1

From the stylistic point of view, in my opinion the paper would read better if it is moderately restructured. The description of the sensitivity experiment (section 3.2) should be located in the methodology section before the result. This would avoid the feeling of jumping back and forth from results to methods, and would help justify the presence of CLM-PLUS results in the earlier figures (before its description). I know this may feel awkward as the CLM – PLUS simulation is seen as a response to the problems (i.e. results) identified in section 3.1, but with some effort I am confident the restructuring can be done. I would introduce this idea (that a CLM – PLUS simulation is done as a response to the first results) in the last paragraph of the introduction, and then describe it in the last section of the methods, saying you are anticipating (in the text) the results that will be presented thereafter.

We agree that the experiment description in section 3.2 could be placed elsewhere. We decided to add the methodological details to the appendix dedicated to the sensitivity experiments including the detailed description of the implementation (P16-19). We now focus only on the results in section 3.2 and give a brief overview of the sensitivity experiments in the method section with a reference to the appendix for the more complete details (P4 L25 to P5 L19). As the referee

proposes we describe the overall goal/motivation of the sensitivity experiment in the last part of the introduction (P3 L18-20).

Another restructuring point I would strongly recommend is to try to separate Results from Discussion. The combined section currently works quite well for 'Results', as such a section should not be just a description of results but also an interpretation of them. But some parts can be moved to a more general 'discussion' section in which the whole approach is discussed in a broader sense, providing more insight of the caveats and advantages of the whole experiment, and how it relates to the broader picture in Earth System Science

This is a good suggestion. We added a separate Discussion section in the revised manuscript (P12-14).

Deforestation is more complex than a simple transition from forest to open land described in the Li et al 2015 MODIS dataset, as different types of forest (e.g. evergreen or deciduous) would have different effects (on snow masking and albedo for instance), and different kind of open lands will also behave differently (management would arguably have a strong influence). With the GETA data, the authors do explore this variability for ET to some extent. In my opinion a more thorough discussion is warranted, even if further analyses are not required within this study. Could anything be said on PFT specific differences for albedo and LST? Are there other field-based datasets such as GETA that could be used for these variables? Could other datasets from remote sensing that differentiate amongst forest types be used? If not, mentioning this need could justify and stimulate the development of such products in the future.

We touch on these topics in the newly added Discussion section (P14 L15-20). In particular, it is now mentioned that other land cover conversions appear to be relevant in recently published observational data sets.

There are some doubts on how comparable the deltas that are extracted from GLEAM are with respect to the Li MODIS dataset and to the CLM sub-grid simulations. If I understand correctly, GLEAM provides separate values for tall canopies and low vegetation over the same 0.25 dd pixels, and to obtain a change between 'forest' to 'openland', one makes the difference a pixel level between the value for tall canopies and for low vegetation. However, to understand better the possible repercussions this may have on the analysis, it would be necessary to have more information on how the distinction between tall canopies and low vegetation is made in GLEAM. What land cover maps are used (if any)? How do these match with the CLM distribution of PFTs?

This is an important issue that is now discussed at P13 L31-34.

The MODIS Li et al. 2015 dataset depends on setting a threshold on the percentage of forest/trees that there are in a pixel so as to consider it 'forest' or 'openland'. They also show in their supplementary material that the choice of the threshold does have

some effect of the results. How does this affect the comparison with CLM sub-grid results, for which the signal is fully 'un-mixed'? In my understanding this has the effect that the MODIS delta will often relate to a comparison from a 'not-so-full-forest' to a 'not-so-treeless-openland', while the simulations are from a 'full-closed forest' to a 'treeless openland'. How does this impact the results? Can something be done about it?

We now discuss this in the section where we present the limitations of our evaluation (P13 L34 to P14 L7). Further, newer observation-based data sets resolving this issue are mentioned in the Discussion (P14 L18-22).

Regarding the discussion on T2M vs LST in both the models and observations, an Important point that is not completely clear is whether T2M is considered as 2m above the canopy or above the soil (i.e. within the forest). Note that in studies like Alkama & Cescatti (2016), the techniques to obtain T2M from satellite LST require weather stations, which typically use WMO definitions by which temperature is measured above a standard grass canopy, even if it is surrounded by forest. This means that the T2M obtained is not that which is observed within the canopy (i.e. under the trees) nor the one above the trees. In the model, and hence in this analysis, what temperature are we speaking about and how can the comparability between observations and models be ensured?

T2M in CLM4.5 is defined as the temperature 2 m above the apparent sink for sensible heat (Oleson et al., 2013; Eq. 5.58) which lies within the canopy air space. In the manuscript we argue that T2M is not the right temperature diagnostic to compare to LST observations (this is why we recalculated a radiative temperature (TRAD) based on the outgoing longwave radiation). We nevertheless show a comparison with daily maximum T2M difference in the appendix (Fig. A10) to highlight the different sign of the response in T2M compared to TRAD in CLM4.5. This result is surprising and is worth noting since it implies that modelling studies looking at land use effects might be affected by the choice of temperature diagnostics, which is an issue that has been overlooked in our community. That said, evaluating this T2M temperature signal in CLM4.5 against observations is very challenging since, as the referee rightfully points out, the WMO T2M concept is by definition not applicable to forest and therefore "T2M" in forest is ill-defined. For instance, the measurements of Lee et al. (2011) report "T2M" above the canopy and "T2M" in Alkama and Cescatti (2016) is indeed a mixed concept derived empirically, which is yet another definition compared to the CLM4.5 definition above.
We added these clarifications to the Discussion section (P13 L22-28). Also, we updated Fig. 9 in the manuscript to make it clearer for the reader that we do not intend to compare the T2M signal in CLM with the MODIS LST observations but emphasize that the T2M and the TRAD signal in CLM look very different.

Page 9 Line 31, Could you speculate on why the model would have this behaviour?

We did not further investigate why the model exhibits this behavior. We think that we can exclude the fact that our simulations were made in offline mode as the cause, since online simulations using CLM exhibit similar daily maximum T2M signals (P11 L26-29).

Could you add some info on whether this shallower root distribution
is closer to what is observed, perhaps based on information from the references cited
(Fan2017 & Canadell1996)? Ideally it would be good to have a line in Fig 6 for the
observations over 'openland'.

We now visualize the rooting depths reported in Fan et al. (2017) in Fig. A5.

How do you calculate the confidence interval in MODIS? Do they come from
the original product of Li et al. 2015? If so, do explain a bit more how they are calculated
and how should the reader interpret it?

We replaced the confidence intervals with the interquartile range which is more suitable to
visualize the variability of not-normally-distributed data in Figs. 2, 5, and 8.

For all plots like that of figure 1, I am not too sure how much we gain in insight by
having the fine 1dd resolution. I would recommend using broader latitudinal bins (e.g.
2.5 or perhaps even 5 dd) so as to have larger boxes in which the points of the t-test
are larger and clearer.

At 0.5° resolution the points displaying the results of the t-test were indeed hard to see. We
therefore averaged the CLM and GLEAM data to latitudinal bands of 2.5° in Figs. 2, 6, 7, 8, 9, A4,
and A8.

Referee 2

Note that we regrouped the statements of referee 2 regarding the importance of VTR to one
single group in order to avoid repetition in our replies.

1) Complete the analysis of the sources of error. You test only things related to vegeta-
tion transpiration and not soil evaporation. Your data do not clearly indicate that vege-
tation is the main driver, and in fact show that soil evaporation could also be a dominant
source of error. Just because your modifications for transpiration show some improve-
ment does not mean that they are correct, because you could be over-compensating
or over-fitting these parameter values.
page 8 lines 21-22 and 31-32: this statement is not supported by your data or the rest
of this paragraph. while the visual pattern between the VTR and total is similar, the
soil evap effects are compensated for by the interception effects, thus leaving VTR to
dominate the pattern. but this doesn't mean that the soil evap is not a main contributor,
especially in the tropics. and you mention the biases in the non-forest that contribute
to this discrepancy as well. Figure 4 also indicates that the soil evaporation dominates
the total ET pattern in the higher latitudes, which is where your modifications show little
improvement.
page 12 lines 11-22: this indicates that your hypothesis regarding VTR as the main

driver of discrepancies may not be correct. while you get improvements, soil evap remains a problem, and you may even be overcompensating with the VTR related modifications

We agree that the soil and canopy evaporation are both important for the creation of the total ET signal. In the initial manuscript this was not clear enough and the importance of VTR was emphasized too much. Therefore, several statements were reformulated in P10 L3-20. Additionally, it is now mentioned in the Discussion that soil and interception evaporation might contribute considerably to the total ET difference (P13 L11-14).
We assessed the importance of the individual components more objectively by calculating the pearson correlation and the index of agreement (Duveiller et al., 2016) between the monthly difference of forest minus open land of these individual components and the monthly difference in total ET. The difference in vegetation transpiration in CLM4.5 exhibits a stronger correlation over a given latitudinal band with the difference in total ET (r of 0.72) than the other two components (vegetation evaporation r of 0.33 and soil evaporation r of 0.19). Similarly, the transpiration difference exhibits a much higher index of agreement with the ET difference than the other two components of ET in the model (0.61, 0.22, and 0.23 for vegetation transpiration, vegetation evaporation, and soil evaporation, respectively; description of index of agreement in Duveiller et al., 2016).

2) Please provide a metric for quantifying the effects of the modifications. Figure 7 (and the aggregate climate zones) is not adequate for demonstrating significant improvement of the results due to the modifications.
Also, while the pft level comparison with GETA looks good, the climate zone comparison is more difficult to evaluate. Aggregating to these climate zones smooths out a lot of spatial variability, and may be too coarse to adequately evaluate the modifications. can you calculate a metric to quantify the effects of the modifications? what do pixel-level correlations between the model and the obs look like? are these correlations improved by the modifications? would zonal grouping make more sense than the climate zones?

We completely agree that introducing a metric helps assessing the performance of different model configurations more objectively. We tested two additional metrics in response to the referee's concern: The root-mean-squared deviation (RMSD) and the index of agreement (IA, as described in Duveiller et al., 2016). For some of the variables with relatively poor agreement (e.g. daily minimum LST difference) the IA tends to be zero or very close to zero in all climate zones. We therefore added a figure displaying the RMSD over the Köppen-Geiger climate zones to the main section of the manuscript (Fig. 4) and the figure displaying the IA to the appendix (Fig. A7). A description of these metrics was added to the Model Evaluation section (P7 L19 to P8 L4).

We agree that aggregating over the climate zones smooths out some of the signal. We argue that zonal grouping will create more heterogeneous groups since there can be strong climatic variations at a given latitude as the referee mentions later in his review. We therefore calculated the metrics over the following refined climate zones: Equatorial humid (Ef and Em), equatorial seasonally dry (Es and Ew), arid, warm temperate fully humid (Cfa, Cfb, and Cfc), warm temperate summer dry (Csa, Csb, and Csc), warm temperate winter dry (Cwa, Cwb, and Cwc), snow warm summer (Dfa, Dfb, Dsa, Dsb, Dwa, and Dwb), and snow cold summer (Dfc, Dfd, Dsc, Dsd, Dwc, and

Dwd). These new climate zones are displayed in Fig. 1. The values of the different metrics over these climate zones are shown in Figs. 3, 4, and A7.

page 6 lines 7-11: I think CLM also outputs a surface radiative temperature. Why didn't you use this?

To our knowledge, CLM4.5 does not provide a radiative temperature output. Therefore, we added a runtime calculation of radiative temperature in the code which however is calculated according to the CLM4.5 Tech-note (Eq. 4.10 of Oleson et al., 2013) which is described in P7 L7-13.

page 7 line 21: confidence in which observations? the non-outliers I assume.

The statement we try to make is that the agreement across different independent data sources gives more confidence in the fact that ET is generally higher over forest. This section was reformulated (P9 L4-6).

page 10 lines 5-6: if comparing for lee et al, why reference alkama and cescatti for the amplification? you should include the delta LST per degree from lee et al for a consistent comparison, and to show that these observations also show this amplification

The structure and formulation of this section was indeed confusing. The statements regarding T2M were made clearer in the manuscript (P11 L23-29 and P13 L22-29).

page 13 line 18: is this because you used prescribed atmospheric forcing?

We indeed considered this hypothesis. But the fact that Lejeune et al. (2017) observed the same for coupled simulations with CLM tends to rule it out (P11 L26-29).

Generally, why show a CI for only the modis zonal average? What about the other data and the model outputs? And is CI the best metric to depict variability here? There are many reasons for variability around the globe at a given latitude (e.g., different weather patterns, continental vs maritime), and we should not expect a zonal mean to behave like a population mean estimate that supposedly characterizes a more homogenous group.

We agree the confidence interval is not the ideal metric to display variability here. We therefore now plot the mean and the interquartile range instead of the confidence interval in Figs. 2, 5, and 8.

Figure 2a: this does not appear to be the correct figure. it does not match with the averages in panel 2c, nor table 2

This was a graphical issue that has now been corrected for Fig. 5.

Evaluating and Improving the Community Land Model's Sensitivity to Land Cover

Ronny Meier[1], Edouard L. Davin[1], Quentin Lejeune[1,2], Mathias Hauser[1], Yan Li[3], Brecht Martens[4], Natalie M. Schultz[5], Shannon Sterling[6], and Wim Thiery[1,7]

[1]ETH Zuerich, Department of Environmental Systems Science, Universitaetstrasse 16, 8092 Zuerich, Switzerland
[2]Now at: Climate Analytics, Ritterstrasse 3, 10969 Berlin, Germany
[3]University of Illinois at Urbana-Champaign, Department of Natural Resource and Environmental Sciences, 1102 South Goodwin Avenue, Urbana IL 61801, USA
[4]Ghent University, Department of Forest and Water Management, Coupure links 653, 9000 Ghent, Belgium
[5]Yale University, School of Forestry and Environmental Studies, 195 Prospect Street, New Haven CT 06511, USA
[6]Dalhousie University, Department fo Earth Sciences, 1459 Oxford Street, Halifax NS B3H 4R2, Canada
[7]Vrije Universiteit Brussel, Department of Hydrology and Hydraulic Engineering, Pleinlaan 2, 1050 Brussels, Belgium

Correspondence to: Ronny Meier (ronny.meier@env.ethz.ch) or Edouard L. Davin (edouard.davin@env.ethz.ch)

Abstract. Modelling studies have shown the importance of biogeophysical effects of deforestation on local climate conditions, but have also highlighted the lack of agreement across different models. Recently, remote sensing observations have been used to assess the contrast in albedo, evapotranspiration (ET), and land surface temperature (LST) between forest and nearby open land on a global scale. These observations provide an unprecedented opportunity to evaluate the ability of land surface models to simulate the biogeophysical effects of forests. Here, we evaluate the representation of the difference of forest minus open land (i.e., grassland and cropland) in albedo, ET, and LST in the Community Land Model version 4.5 (CLM4.5) using various remote sensing and in-situ data sources. To extract the local sensitivity to land cover we analyze plant functional type level output from global CLM4.5 simulations, using a model configuration that attributes a separate soil column to each plant functional type. Using the separated soil column configuration, CLM4.5 is able to realistically reproduce the biogeophysical contrast between forest and open land in terms of albedo, daily mean LST, and daily maximum LST, while the effect on daily minimum LST is not well captured by the model. Furthermore, we identify that the ET contrast between forests and open land is underestimated in CLM4.5 compared to observation-based products and even reversed in sign for some regions, even when considering uncertainties in these products. We then show that these biases can be partly alleviated by modifying several model parameters, such as the root distribution, the formulation of plant water uptake, the light limitation of photosynthesis, and the maximum rate of carboxylation. Furthermore, the ET contrast between forest and open land needs to be better constrained by observations in order to foster convergence amongst different land surface models on the biogeophysical effects of forests. Overall, this study demonstrates the potential of comparing sub-grid model output to local observations to improve current land surface models' ability to simulate land cover change effects, which is a promising approach to reduce uncertainties in future assessments of land use impacts on climate.

1

1 Introduction

While the forested area has stabilized or is even increasing over Europe and North America, deforestation is still ongoing at a fast pace in some areas of South America, Africa, and Southeast Asia (Huang et al., 2009; Hansen et al., 2013; Margono et al., 2014; McGrath et al., 2015). In addition, carbon sequestration by re- or afforestation has been proposed as a strategy to mitigate anthropogenic climate change (Brown et al., 1996; Sonntag et al., 2016), making forest loss or gain likely an essential component of future climate change. Changes in forest coverage impact climate by altering both the carbon cycle (Ciais et al., 2013) and various biogeophysical properties of the land surface such as albedo, evaporative fraction and roughness length (Bonan, 2008; Pitman et al., 2009; Davin and de Noblet-Ducoudré, 2010; Li et al., 2015). However, there exist considerable discrepancies in the representation of biogeophysical effects amongst land surface models, thus generating a need for a thorough evaluation of the representation of these effects in individual models.

Model simulations indicate that the biogeophysical effects of historical deforestation have been rather small on a global scale (Davin et al., 2007; Findell et al., 2007; Davin and de Noblet-Ducoudré, 2010; De Noblet-Ducoudré et al., 2012; Malyshev et al., 2015). However, they have likely been significant on regional and local scales, especially over areas which experienced intense deforestation rates (Pongratz et al., 2010; De Noblet-Ducoudré et al., 2012; Kumar et al., 2013; Malyshev et al., 2015; Lejeune et al., 2017, 2018). Similarly, present-day observational data, either based on in-situ (Juang et al., 2007; Lee et al., 2011; Zhang et al., 2014; Bright et al., 2017) or remote-sensing measurements (Li et al., 2015; Alkama and Cescatti, 2016; Li et al., 2016; Duveiller et al., 2018) show that biogeophysical effects of forests can strongly influence local climate conditions. Among the different biophysical effects, the increased surface albedo (cooling effect), the alteration of the evaporative fraction (warming or cooling effect, depending on the region and season), and the lower surface roughness causing a reduction of the turbulent heat fluxes (warming effect) have been identified as the three main drivers of the climate impact of deforestation (Bonan, 2008; Pitman et al., 2009; Davin and de Noblet-Ducoudré, 2010; Li et al., 2015). However, some of these biogeophysical processes are not well represented in current land surface models. The model intercomparison projects LUCID (Land-Use and Climate, IDentification of robust impacts) and CMIP5 (Coupled Model Intercomparison Project Phase 5) exposed the lack of model agreement concerning the biogeophysical impacts of historical land use and land cover change (LULCC), especially regarding the impact on evapotranspiration (ET) and temperature during the warm season over the mid-latitudes of the northern hemisphere (De Noblet-Ducoudré et al., 2012; Kumar et al., 2013; Lejeune et al., 2017). In addition, distinct discrepancies between present-day temperature observations and the simulated historical effects of LULCC over North America were identified (Lejeune et al., 2017). This highlights the need for systematic evaluation and improvement of the representation of biogeophysical processes in land surface models.

Observing the local climatic impact of LULCC is not straightforward. When temporally comparing observational data over an area undergoing LULCC, it is difficult to disentangle the effect of the LULCC forcing from other climatic forcings (e.g.,

greenhouse gas forcing). To overcome this difficulty, observational studies often spatially compare nearby sites of differing land cover, assuming that they receive the same atmospheric forcing (e.g., von Randow et al., 2004; Lee et al., 2011). Hence, the sensitivity of land surface models to land cover can be evaluated best with observational data by spatially comparing different land cover types in models. Recently, Malyshev et al. (2015) employed a new approach to assess the local impacts of

5   LULCC in land surface models by comparing climate variables over tiles corresponding to different plant functional types (PFTs) located within the same grid cell. Since PFT tiles within the same grid cell experience exactly the same atmospheric forcing, the resulting sub-grid ~~deforestation~~ land cover signal extracted by this method achieves good comparability to local observations which contrast neighboring forest and open land sites (Lee et al., 2011; Li et al., 2015; Alkama and Cescatti, 2016; Li et al., 2016).

10

Here we aim to evaluate and improve the sensitivity of the Community Land Model 4.5 (CLM4.5) to land cover, using observational data of the local contrast between forest and open land (i.e., grassland and cropland). In Section 3.1 of this study, we systematically analyze the representation of the local difference of forest minus open land in albedo, ET and land surface temperature (LST) in CLM4.5 against the newly released observational remote-sensing-based products of Li et al. (2015).

15   The forest signal in CLM4.5 is extracted by comparing tiles corresponding to forest and open land, similar to Malyshev et al. (2015). Given the uncertainties in observation-based ET estimates, we further extend our evaluation by including data from the Global Land Evaporation Amsterdam Model (GLEAM) version 3.1a (Miralles et al., 2011; Martens et al., 2017) and the Global ET Assembly (GETA) 2.0 (Ambrose and Sterling, 2014), which are based on remote-sensing and in-situ observations, respectively. Finally, a sensitivity experiment is presented in Section 3.2, in order to explore the possibilities to better represent the

20   ET impact of forests in CLM4.5~~(Section 3.2)~~. This configuration of CLM4.5 incorporates modifications in root distribution, plant water uptake, light limitation of photosynthesis, and maximum rates of carboxylation.

2   Methods and Data

2.1   Model Description and Set Up

CLM is the land surface component of the Community Earth System Model (CESM), a state-of-the-art earth system model

25   widely applied in the climate science community (Hurrell et al., 2013). CLM represents the interaction of the terrestrial ecosystem with the atmosphere by simulating fluxes of energy, water and a number of chemical species at the interface between the land and the atmosphere. The represented biogeophysical processes include absorption and reflection of both diffuse and direct solar radiation by the vegetation and soil surface, emission and absorption of longwave radiation, latent and sensible heat fluxes from the soil and canopy, and heat transfer into the snow and soil. Sub-grid heterogeneity is taken into account in CLM by the

30   subdivision of each land grid cell in five land ~~cover~~ units (glacier, wetland, vegetated, lake and urban). The vegetated land unit is further divided into 16 tiles representing different PFTs (including bare soil).

We run ~~the latest~~ CLM version 4.5 at 0.5° resolution for the period 1997-2010. A five-year (1997-2001) spin-up period is

excluded from the analysis to minimize the impact of the model initialization. The analysis of CLM4.5 therefore covers the period of 2002 to 2010 which matches well with the observation period of 2002 to 2012 of Li et al. (2015). Assuming that the feedback of the land surface to the atmosphere is of minor importance for the sub-grid contrast between forest and open land tiles, simulations are performed in offline mode using atmospheric forcing ~~data~~ from the CRUNCEP v4 reanalysis product

5 (Vivoy, 2009; Harris et al., 2014). The land cover map and vegetation state data are prescribed based on MODIS observations (Lawrence and Chase, 2007, Fig. A1). The land cover map from the year 2000 is kept static during the entire simulation period, since no land cover change is required to retrieve a spatial contrast between forest and open land~~(Fig. A1). .~~. The optional carbon and nitrogen module of CLM4.5 as well as the crop- and irrigation modules are kept inactive in our simulations.

10 By default, all PFTs within a grid cell in CLM4.5 share a single soil column (Oleson et al., 2013), implying that all PFTs experience the same soil temperature and soil moisture (SM). Further, the surface energy balance at PFT level is closed using the ground heat flux (GHF; i.e., ~~it~~ GHF is calculated as the residual of the other energy fluxes). Hence, the soil warms in case of an energy excess at the land surface and vice versa. Warmer ~~/cooler~~(cooler) soil in turn will result in increased ~~/decreased~~(decreased) sensible and latent heat fluxes away from the ground and/or increased ~~/decreased~~(decreased) emit-

15 ted longwave radiation, thereby counteracting the initial energy imbalance. Consequently, this model architecture eventually results in near-zero daily mean GHF, once the soil temperature has adjusted to an equilibrium state with a near-zero energy imbalance. On shared soil columns (ShSCs), however, GHFs can reach unrealistically high values for individual PFTs (Fig. A2~~a~~ and c), because a common soil temperature is artificially maintained for all PFTs, which differs from their individual equilibrium states. This assumption leads to a ~~non-zero~~ net GHF into the soil over open land PFTs and out of the soil over forest

20 PFTs for the majority of the locations across the globe, implying a lateral subsurface heat transport from open land towards forests (Schultz et al., 2016). To resolve this issue, Schultz et al. (2016) proposed a modification of CLM4.5 which attributes a separate soil column (SeSC) to each PFT. This modification allows the soil of individual PFTs to equilibrate to a different temperature (Fig. A3) and suppresses these unrealistically high (lateral) GHFs (Fig. A2~~b~~ and d). Here, we present results from a simulation on SeSCs, called CLM - BASE, unless stated otherwise (Table ~~??~~A4). We also performed a simulation on ShSCs

25 named CLM - DFLT.

Further, we present a sensitivity experiment, named CLM - PLUS in Section 3.2, in which we try to alleviate ~~some of the detected biases by modifying~~ detected biases in ET. Besides the SeSCs, four aspects in the parameterization of vegetation transpiration ~~in~~ (VTR) are modified in this sensitivity experiment:

30 – Shallower root distribution for grass- and cropland PFTs. CLM4.5 ~~: (1) a shallower~~ accounts for SM stress on transpiration through a stress function β_t, which ranges from zero (when soil moisture limitation completely suppresses VTR) to one (corresponding to no soil moisture limitation of VTR). Forests for the most part experience higher SM stress than open land in CLM - DFLT except in the northern high-latitude winter (Fig. A4), partly caused by the similar root distribution for all PFTs but evergreen broadleaf trees (Fig. A5). In reality, observed maximum rooting

35 depths are considerably higher for forests than for grassland and cropland (Canadell et al., 1996; Fan et al., 2017).

Likewise, in-situ observations in the tropics show that grassland ET decreases during dry periods, because grasses have only limited access to water reservoirs located below a depth of 2 m (von Randow et al., 2004). Hence, we aim to increase SM stress of open land PFTs ~~, (2)a dynamic~~ and reduce their ability to extract water from the lower part of the soil, by introducing a shallower root distribution for these PFTs (Fig. A5). This root distribution was not fitted to a particular observed root distribution. However, the new root distribution agrees better with the average rooting depth of annual grass reported by Fan et al. (2017).

- Dynamic plant water uptake~~, in which plants only access~~. Tropical forests are often observed to exhibit increased ET during dry periods, due to increased incoming shortwave radiation (da Rocha et al., 2004; Huete et al., 2006; Saleska et al., 2007). That is, despite the upper soil being dry, tropical trees still have sufficient access to water from deeper soil layers (Jipp et al., 1998; von Randow et al., 2004). We aim to allow a similar behaviour in CLM4.5 by introducing a dynamic plant water uptake, where plants only extract water from the 10 % of the roots ~~where water is most easily available, (3)decreased and increased~~ with best access to SM (example in Fig. A6).

- Light limitation reduction for C_3 PFTs and enhancement for C_4 PFTs. In CLM‑BASE ET of boreal PFTs is underestimated compared to GETA 2.0 (Fig. 3f). Since VTR of these PFTs is only weakly affected by SM stress, light limitation for C_3 plants ~~and~~ is reduced. On the other hand, C_4 ~~plants, respectively, and (4) adapted values in the maximum rate of carboxylation.~~ grass shows a considerable positive bias in ET, which we try to alleviate by increasing the light limitation of this PFT.

- Modified maximum rates of carboxylation (V_{cmax}; Table A1). This PFT-specific parameter is suitable to tune VTR, since it is not well constrained from observations and VTR in models is highly sensitive to this parameter (Bonan et al., 2011). The new values were chosen with the aim to alleviate biases relative to GETA 2.0 (Fig. 3f) and still lie well within the range of observations collected in the TRY plant trait database (Boenisch and Kattge, 2017). Additionally, the minimum stomatal conductance of C_4 plants, which is by default four times larger than that of C_3 plants, is reduced.

A technical description of these modifications as well as a discussion of the effect on ET by each individual modifications is provided in Appendix A.

2.2 Observational Data

~~A MODIS-based dataset is~~ The data published in Li et al. (2015) ~~is~~ are used to evaluate the effects of ~~forest~~ forests on local climate variables in CLM4.5. This data set was created by applying a window searching algorithm to remote-sensing LST, albedo, and ET products from the MODerate resolution Imaging Spectroradioameter (MODIS), in order to systematically compare these variables over forest and open land on a global scale. The data of this study, hereafter referred to as MODIS,

cover the period of 2002 to 2012 and were aggregated from the initial window size of $0.45° \times 0.25°$ to $\cancel{1}0.5° \times 0.5°$ spatial resolution. Hence, the similar spatial scale of the MODIS data and the CLM4.5 simulations allows for good comparability between these two data sources.

5   We also use two additional observation-based datasets of ET to consider uncertainties in present-day ET estimates. A number of different global ET products are available which, however, exhibit substantial discrepancies (Mueller et al., 2011; Wang and Dickinson, 2012; Mueller et al., 2013; Michel et al., 2016; Miralles et al., 2016). In particular, the algorithm from Mu et al. (2011) used to retrieve the MODIS ET product was found to systematically underestimate ET compared to in-situ and catchment-scale observations (Michel et al., 2016; Miralles et al., 2016). In addition, algorithms used to infer ET from remote-

10   sensing observations make assumptions on how ~~LC~~ the land cover type influences ET, preventing an independent identification of the influence of LULCCs on ET. We therefore complement our evaluation of the ET impact of forest in CLM4.5 with two additional data sets: GLEAM version 3.1a and GETA 2.0.

GLEAM was introduced in 2011 (Miralles et al., 2011) and revised twice, resulting in the current version 3.1 (Martens et al.,

15   2017). It provides ~~ET estimates~~ estimates of potential ET for tall canopy, bare soil, and low vegetation after Priestly and Taylor (1972). Potential ET of vegetated land surfaces is converted into actual ET using vegetation-dependent parameterizations of evaporative stress. Canopy interception evaporation is calculated separately using the parameterization of Gash and Stewart (1979). GLEAM uses surface radiation, near-surface air temperature, surface SM, precipitation, snow water equivalent, and vegetation optical depth observations to estimate ET globally at $0.25°$ resolution. To maximize spatial and temporal overlap

20   with the MODIS observations, we choose GLEAM version 3.1 a (hereafter referred to as GLEAM), which incorporates re-analysis input besides satellite observations. We compare the ET estimates for tall canopy and low vegetation to model output for forests and open land, respectively. Since interception loss is only estimated for tall canopy, it was fully attributed to ET from forests.

25   GETA 2.0 (Ambrose and Sterling, 2014) is a suite of global-scale fields of actual ET for 16 separate land cover types (LCTs), derived from a collection of in-situ measurements between 1850 and 2010. Using a linear mixed effect model with air temperature, precipitation, and incoming shortwave radiation as predictors, ~~they obtained~~ yearly ET estimates for each of these 16 different LCTs have been obtained with a global coverage and $1°$ spatial resolution. We then use the same land cover map employed for the CLM4.5 simulations to weigh the different LCTs in this data set and retrieve an ET value for forest and open

30   land (see Section 2.3 for more details). Since our CLM4.5 simulations were conducted without irrigation, we did not include the GETA 2.0 irrigation layer. We refer to this data set as GETA in this study.

2.3 ~~Comparison Strategy~~ Model Evaluation

The forest signal in CLM4.5 is extracted by comparing the area-weighted mean of the variables of interest over all ~~tiles corresponding to forest PFTs to the area-weighted mean over the tiles corresponding to open land~~ forest tiles to its corresponding values over open land tiles (i.e., grassland and cropland)~~PFTs~~, similar to Malyshev et al. (2015). As such, it becomes possible to ~~deduce~~ infer a forest signal for every model grid cell containing any forest and any open land PFT, no matter how small the fraction of the grid cell covered by these PFTs. The different PFT tiles within a $0.5° \times 0.5°$ grid cell in our CLM4.5 simulations are subject to the exact same atmospheric forcing and are hence comparable to the almost local effect of forests retrieved at a resolution of $0.45° \times 0.25°$ in MODIS. It needs to be noted that the MODIS observations can only be retrieved under clear-sky conditions, thereby potentially impairing the comparability to our CLM4.5 data which are not filtered for clear-sky days. Nevertheless, it was decided to include cloudy days for the analysis of the CLM4.5 simulations, to preserve the comparability to studies which do not distinguish between cloudy and clear-sky days (e.g. GLEAM; GETA; da Rocha et al., 2004; von Randow et al., 2004; Liu et al., 2005).

~~12~~ Twelve of the 16 PFTs of CLM4.5 are attributed to either the forest or the open land class as described in Table A2. Consistent with Li et al. (2015), open land was considered the combination of grassland and cropland. Hence, bare soil as well as shrubland are excluded from our analysis. Forest and open land ET of GETA was aggregated similarly using the same LC map as in the CLM4.5 simulations, with the LCTs of GETA attributed to the different CLM4.5 PFTs as listed in Table A3. To ensure a consistent comparison with the LST data from MODIS, we derive a radiative temperature (T_{rad}) from the emitted longwave radiation output (LW_{up}) in CLM4.5 according to Stefan-Boltzmann's law (assuming that emissivity is 1 as in Eq. 4.10 of Oleson et al., 2013):

$$T_{rad} = \sqrt[4]{\frac{LW_{up}}{\sigma}} \tag{1}$$

with σ being the Stefan-Boltzmann constant ($5.567 \times 10^{-8} \, \mathrm{W \, m^{-2} \, K^{-4}}$). Hereafter T_{rad} will be referred to as LST. For the local difference of forest minus open land in albedo, ET, daily mean LST, daily maximum LST, and daily minimum LST we will use the symbols $\Delta\alpha$(f-o), ΔET(f-o), ΔLST$_{avg}$(f-o), ΔLST$_{max}$(f-o), and ΔLST$_{min}$(f-o), respectively.

In order to evaluate the different CLM4.5 simulations objectively, three different metrics are calculated over the following eight Köppen-Geiger climate zones (Kottek et al., 2006): Equatorial humid (E-h), equatorial seasonally dry (E-sd), arid (Ar), warm temperate winter dry (T-wd), warm temperate summer dry (T-sd), warm temperate fully humid (T-fh), snow warm summer (S-ws), and snow cold summer (S-cs) (Fig. 1). As a first metric, the area-weighted mean for a given variable over these climate zones ($\overline{\Delta x}$) is calculated as follows:

$$\overline{\Delta x} = \frac{\sum\limits_{i} A_i \Delta x_i}{\sum\limits_{i} A_i} \tag{2}$$

where Δx_i are the differences of forest minus open land in variable x of all the grid cells i belonging to the respective climate zone and A_i their areas. Secondly, the CLM4.5 simulations are compared in terms of the area-weighted root mean squared deviation (RMSD) to the observation-based data sources:

$$\text{RMSD}(\Delta x) = \sqrt{\frac{\sum\limits_i A_i (\Delta x_i^{sim} - \Delta x_i^{obs})^2}{\sum\limits_i A_i}} \tag{3}$$

where Δx_i^{sim} and Δx_i^{obs} are the simulated and observed differences of forest minus open land in variable x. RMSD for a Köppen-Geiger climate zone is calculated from a data pool collecting all monthly values with data in CLM4.5 and the given observational data which lie within the respective climate zone (except when comparing to GETA for which only long-term annual means are available).

Lastly, the index of agreement (IA; Duveiller et al., 2016) was calculated for the same data pools as RMSD. This dimensionless metric describes the agreement between two data sets, with 0 indicating no agreement and 1 indicating perfect agreement. By definition, this metric is set to 0 if the two compared data sets exhibit a negative pearson correlation. Since results of this metric generally support those of RMSD, they are shown in the Appendix (Fig. A7).

3   Results~~and Discussion~~

3.1   Evaluation of the Local Effect of Forests in CLM4.5

3.1.1   Albedo

The MODIS satellite observations and CLM - BASE agree on a generally negative $\Delta\alpha$(f-o) (Fig. 2). ~~This difference is amplified towards the poles and in wintertime due to the snow masking effect. $\Delta\alpha$(f-o) tends to be more negative in CLM - BASE than in the satellite observations in all Köppen-Geiger climate zones (Table ??).~~ Effectively, MODIS observations show ~~a~~ slightly positive $\Delta\alpha$(f-o) for some latitude-month combinations concentrated in the tropics and sub-tropics (Fig. 2); however, these differences are mostly insignificant and ~~are likely related to~~ must be considered in the light of uncertainties in the MODIS observations, which are more sparse over these regions due to frequent cloud coverage (Li et al., 2015). ~~MODIS albedo retrievals tend to underestimate albedo over grass- and cropland, especially under the presence of snow, and overestimate it over forests~~ The negative albedo difference is amplified towards the poles and in wintertime due to the ~~heterogeneity of land cover within pixels . Therefore, it is likely that the magnitude of $\Delta\alpha$(f-o)is underestimated in MODIS. Consistently with this hypothesis, in-situ observations of paired forest and open land sites support the higher~~ snow masking effect (Harding et al., 2001). Among the non-snow climate zones, the albedo contrast between forest and open land is strongest in the Ar and the T-sd climate zones (Fig. 3a). This could be related to the occurrence of dry periods in these climate zones during which open land dries out more easily than forests due to their shallower root profiles (Canadell et al., 1996; Fan et al., 2017). As green leaves have lower albedo than dry leaves and the soil, the albedo contrast between the still-green forest and the dried-out open land

8

would be intensified in such a scenario (Dorman and Sellers, 1989). $\Delta\alpha$(f-o) ~~found~~ tends to be more negative in CLM-BASE ~~. In agreement with the observations described above, earth system models concordantly simulate an albedo increase due to deforestation, which is enhanced during winter at higher latitudes . Nevertheless, considerable discrepancies amongst different land surface models remain regarding the order of magnitude of this increase. Hence, we encourage an extension of our analysis of CLM4.5 to the land component of other earth system models and possibly using additional observational constraints.~~ than in the satellite observations in all Köppen-Geiger climate zones, especially in the snow climate zones. RMSD values over the climate zones exhibit similar tendencies as the magnitudes of mean $\Delta\alpha$(f-o) and have roughly 50 % the magnitude of mean $\Delta\alpha$(f-o) ~~is especially uncertain at northern high-latitudes . Since this variable has the highest magnitude at the very same geographic location, a focus on these areas could be highly relevant in future studies.~~(Fig. 4a). The exception to this are the tropical climate zones where the magnitude of RMSD is roughly the same as the mean values of mean $\Delta\alpha$(f-o). This is likely related to the fact that MODIS observes only a weak albedo signal of forests in these climate zones.

3.1.2 Evapotranspiration

All of the considered observation-based ET products indicate that annual mean ΔET(f-o) is positive in ~~all~~ every climate zones, despite considerable variations in the magnitude of this difference (~~Table~~Fig. ~~??~~3e). GLEAM suggests a near zero ΔET(f-o) in the arid climate zone most likely because it uses surface SM data as an input to estimate ET. Also, GLEAM exhibits positive ΔET(f-o) ~~over forests~~ throughout the year in the mid-latitudes, unlike MODIS which ~~proposes~~ has a negative ΔET(f-o) during winter ~~.~~(Fig. 6). Paired-site FLUXNET studies offer ~~and~~ an additional opportunity to compare ET over forest and over open land on a point scale. Overall, they report higher ET for tropical forests (Jipp et al., 1998; von Randow et al., 2004; Wolf et al., 2011). In the mid- and high-latitudes a number of FLUXNET studies observe a positive ΔET(f-o) during summer, and a near-zero negative ΔET(f-o) during winter, similar to MODIS (Fig. 6; Liu et al., 2005; Stoy et al., 2006; Juang et al., 2007; Baldocchi and Ma, 2013; Vanden Broucke et al., 2015; Chen et al., 2018). On the other hand, ~~there are also FLUXNET observations indicating a~~ negative ΔET(f-o) have been observed at some paired FLUXNET sites in the tropics (Van der Molen et al., 2006) and in the mid-latitudes during summer (Teuling et al., 2010). ~~Although these contradicting results highlight some uncertainties, overall the fact that the two remote sensing-based data sets considered in this study, GETA, which is based on in-situ ET measurements, as well as most paired-site FLUXNET studies consistently exhibit positive ΔET (f-o)~~ The considered global ET data sets however consistently exhibit higher ET over forests in most regions ~~,~~(Fig. 5). This agreement across the different independent global data sources gives some confidence in ~~these observations~~the fact that ET is generally higher over forests. Nevertheless, it needs to be noted that ΔET(f-o) ~~in GETA looks fundamentally different when~~ GETA shows fundamentally different results when considering the data over irrigated crops instead of data over rainfed ~~are considered~~ (resulting in negative ΔET(f-o) at many locations). Therefore, distinguishing irrigated from rainfed crops in future evaluations would be essential, but remains beyond the scope of this study.

CLM‑BASE exhibits considerable discrepancies in ΔET(f-o) to the observation-based data sets both for the annual mean values (Fig. 5) and the seasonal cycle (Fig. 6). ΔET(f-o) in CLM‑BASE is near zero ~~at almost all latitudes~~ in all climate zones (Fig. ~~5~~3e), and even negative in the ~~equatorial and arid climate zones~~E-sd climate zone, unlike the global ET datasets which clearly suggest positive values~~for this variable (except for GLEAM in the arid region).~~. The large bias of ΔET(f-o) in CLM‑BASE is also apparent in the RMSD values, which are similar in magnitude to the observed mean signal (compare Figs. 3e and 4e). A comparison of the absolute ET values of each PFT in CLM‑BASE versus the GETA data reveals that CLM‑BASE generally exhibits similar ET averages for needleleaf PFTs, lower ET averages for broadleaf deciduous trees as well as crops, and higher ET averages for non-arctic grasses and broadleaf evergreen trees (Fig. 3f). Notably, evergreen and deciduous tropical broadleaf trees as well as C_4 grass have a bias larger than $0.2\,\mathrm{mm\,day^{-1}}$ relative to GETA. The biases of these PFTs can have a large effect on the overall ΔET(f-o) as they cover a large proportion of the land surface (9.5 %, 8.0 %, and 8.0 %, respectively). Similarly, CLM‑BASE overestimates ET compared to in-situ measurements conducted over a pasture site in the Amazon by von Randow et al. (2004) and underestimates ET compared to the two forest sites in Alaska reported in the study of Liu et al. (2005) (Table 1).

Interestingly, deciduous trees are mostly responsible for this discrepancy in ΔET(f-o) at latitudes below 30° (Fig. A8). In the mid-latitudes, on the other hand, both deciduous and evergreen trees show lower ET than open land during summer and higher ET during winter, which is inconsistent with GLEAM and, even more so, inconsistent with the seasonally-varying ΔET(f-o) in MODIS. Another noteworthy result is that the SeSC configuration (i.e., CLM‑BASE) appears to impair the agreement on ΔET(f-o) between CLM4.5 and the ~~considered~~ observations (Fig. 6). CLM‑DFLT exhibits a positive ΔET(f-o) throughout the year except for the tropical dry season which is caused by deciduous broadleaf trees exhibiting lower ET than open land (Fig. ~~??~~A8a, b, and c). There are two ~~likely~~ potential reasons for the negative bias in ΔET(f-o) introduced by SeSCs. First, the implicit lateral ~~GH flux~~ GHF from open land towards forest which occurs in CLM‑DFLT (Fig. A2) provides additional energy over forests for turbulent heat fluxes. This energy source ~~/sink~~~~for forests/open land~~(sink) for forests (open land) is disabled by SeSCs. Second, the lower soil temperature of forests in CLM‑BASE (Fig. A3) reduces the specific humidity gradient between the soil surface and the atmosphere and hence also the absolute soil evaporation. It needs to be noted that the weaker agreement with observational data of ΔET(f-o) in CLM‑BASE than in CLM‑DFLT does not necessarily imply a worse representation of the evaporative processes in CLM‑BASE, but could also originate from the fact that CLM4.5 was tuned to retrieve realistic ET values on ShSCs.

To shed light on the origin of the ΔET(f-o) bias in CLM4.5, we separately analyze the three components of ET in CLM4.5: soil evaporation (including sublimation/evaporation from the snow- and water-covered fraction of the soil), canopy interception evaporation, and vegetation transpiration (VTR). As ~~can be~~ seen in Fig. 7d, there is a distinct band around the equator where soil evaporation is considerably lower in forests than in open land. Interestingly, both the study of Chen et al. (2018) and ours show that the lower soil evaporation signal only arises for the configuration with SeSCs (data of CLM‑DFLT are not

presented here). Thus, lower soil evaporation around the equator in CLM - BASE is likely related to the diminution of the soil temperature and of the available energy mentioned earlier in this section. It appears reasonable that, in comparison with open land, forests have lower soil evaporation since (1) the forest soil surface receives less incoming solar radiation, (2) more of the incoming precipitation is intercepted by the canopy, and (3) the water vapour concentrations within the canopy are higher.

Yet soil evaporation and canopy interception evaporation ~~make up a much larger proportion of the~~ contribute a larger proportion to total ET in CLM4.5 (31 % and 19 %) compared to GLEAM ~~(14 % and 10 %;)~~ (14 % and 10 %; Martens and Miralles, 2017). It is thus possible that the strength of this effect is too large in CLM4.5. However, most ET measurement techniques cannot distinguish among the different components of ET, making it difficult to assess which partitioning is more realistic. ~~Anyhow~~Overall, negative ΔET(f-o) values in ~~CLM4.5 are mainly driven by the lower VTR of forests in most regions~~CLM - BASE typically coincide with negative differences in VTR, in particular during the wet season in the tropics and sub-tropics and during summer at higher latitudes (Fig. 7 c and f)~~. A comparison of the absolute ET values of each PFT in CLM~~, whereas negative values in the soil evaporation difference are partly compensated by positive values in interception evaporation (Fig. ~~BASE versus the GETA data reveals that CLM - BASE generally exhibits similar ET averages for needleleaf PFTs, lower ET averages for broadleaf deciduous trees as well as crops, and higher ET averages for non-arctic grasses and broadleaf evergreen trees (Table ??). Particularly, evergreen and deciduous tropical broadleaf trees as well as C₄ grass have a bias larger than 0.2 relative to GETA. The biases of these PFTs can have a large effect on the overall~~ 7d and e). It is therefore likely, that VTR is the main driver behind the ΔET(f-o) ~~as they cover a large proportion of the land surface. Similarly, CLM~~bias even though the contribution of the individual ET components to the total signal cannot be evaluated. For this reason, the modifications in the CLM - ~~BASE considerably overestimates ET compared to in-situ measurements conducted over a pasture site in the Amazon by and underestimates ET compared to the two forest sites in Alaska reported in the study of (Table 1)~~PLUS sensitivity experiment are targeted at altering vegetation transpiration.

In summary, ΔET(f-o) in CLM4.5 exhibits considerable discrepancies to the considered global ET datasets and in-situ observations. The SeSC configuration amplifies these discrepancies, which are ~~mainly~~ typically driven by the difference in VTR of forest minus open land.

3.1.3 Land Surface Temperature

~~CLM~~The overall local temperature impact of forests is the result of several biogeophysical properties acting simultaneously. They include lower albedo of forests (warming effect), higher surface roughness (cooling effect if land surface is warmer than boundary layer), and alteration of the evaporative fraction (Bonan, 2008; Pitman et al., 2009; Davin and de Noblet-Ducoudré, 2010; Li et al., 2015). For daily mean LST forests exhibit a cooling effect in MODIS except for the winter months at latitudes exceeding 30° (Fig. ~~BASE generally captures the sign of~~ 8a). This implies that the cooling effects of higher surface roughness and generally higher evaporative fraction over forests are stronger than the warming effect due to their lower albedo. ΔLST$_{avg}$(f-o) and ΔLST$_{max}$(f-o) ~~, but shows considerable discrepancies for~~ are positive only under the presence of snow, as $\Delta\alpha$(f-o) is amplified due to the snow masking effect (moreover sensible heat fluxes

are often directed towards the land surface during winter at high latitudes, resulting in warmer forests due to their higher surface roughness inducing stronger turbulent heat fluxes; Liu et al., 2005). The observed magnitude of ΔLST_{max}(f-o) tends to be larger than that of ΔLST_{avg}(f-o) likely due to the fact that the observed daytime effect is partly compensated by an opposing nighttime effect (Fig. 3b, c, and d). MODIS exhibits an overall cooling effect of forests on daily mean LST in all climate zones, including the snow climate zone where the sign of the difference changes seasonally (Fig. 8d). Further, this dataset shows a slightly negative ΔLST_{min}(f-o) in tropical and sub-tropical regions and even a positive ΔLST_{min}(f-o) in the mid-latitudes ~~, compared to MODIS~~ (Fig. 8~~)~~g). This nighttime signal in the mid-latitudes is observed in several observational studies but its source is not yet fully determined (Lee et al., 2011; Vanden Broucke et al., 2015; Li et al., 2015).

CLM‑BASE ~~and MODIS both exhibit an overall negative~~ generally captures the sign and magnitude of ΔLST_{avg}(f-o) ~~, except for winter at latitudes exceeding 30 (~~and ΔLST_{max}(f-o) compared to MODIS (Fig. 8). The SeSCs used in CLM ~~a, b, and c)~~ BASE allow for larger LST differences between forest and open land than the default version of CLM4.5 (CLM‑DFLT) on ShSCs, resulting in a better agreement with the observed magnitudes. This is due to the fact the GHF on ShSCs counteracts the soil temperature difference and thereby also the LST difference between forest and open land. Nevertheless, there are still some discrepancies between the LST signal in CLM‑BASE and the MODIS observations. It appears that ΔLST_{avg}(f-o) in CLM‑BASE has a positive bias in the equatorial, the arid, and the snow climate ~~zone~~ zones and a negative bias in the ~~warm temperate zone (Table ??). The results for ΔLST_{max}~~T-wd and T-fh climate zones ~~(f-o) appear similar to those for ΔLST_{avg}(f-o) overall (Fig. 8). However, the observed magnitude of ΔLST_{max}(f-o) tends to be larger (Table ??). For this variable MODIS exhibits an overall cooling effect of forests in all climate zones , including the snow climate zone where the sign of the difference changes seasonally (~~Fig. 8~~ d~~3b). ΔLST_{max}(f-o) in CLM‑BASE appears qualitatively similar to the MODIS observations (Fig. 8~~d, e, and f) but is biased positively in all climate zones (~~Table~~Fig. ~~??~~3c). In contrast, daily minimum LST shows much larger discrepancies between CLM‑BASE and MODIS (Fig. 8g, h, and i). ~~The MODIS data indicate that forests have a weak and mostly insignificant nighttime cooling effect in the equatorial region and a significant nighttime warming in the mid-latitudes throughout the year. On the other hand~~In CLM‑BASE, ΔLST_{min}(f-o) ~~in CLM‑BASE~~ is similar to ΔLST_{avg}(f-o) and ΔLST_{max}(f-o)~~(Fig. 8)~~, i.e. ~~,~~ forests have an overall nighttime cooling effect in all climate zones except for the neutral signal in the snow climate ~~zone (Table~~zones, whereas MODIS exhibits an only weak nighttime cooling effect in the tropical climate zones and a clear nighttime warming effect in all other climate zones (Fig. ~~??). While this comparison suggests considerable biases of CLM, there are three factors that can impair a full quantitative comparison of the MODIS LST~~difference between forest and open land with our CLM4.5 simulations. (1) As for the MODIS albedo product, the LST measurements of MODIS might often retrieve a mixed signal of various land cover types, which likely dampens~~ 3d). The weak performance of CLM‑BASE in terms of ΔLST_{min}(f-o) is also visible in the RMSD values which are considerably larger than the mean ΔLST_{min}(f-o) ~~. (2) MODIS LST data are retrieved under clear-sky conditions only, whereas we do not mask out cloudy days in the evaluation of the CLM4.5 simulations.Whilst albedo is likely unaffected by cloud coverage, it might well be relevant for LST, as it affects the radiation budget of the land surface. (3) The overpass times~~

of the MODIS satellite system are at 1:30 ~~signal (compare Figs.~~ ~~am and 1:30 pm, hence not necessarily coinciding with the~~ ~~daily maximum and minimum LST in CLM4.5.~~ signal (compare Figs. 3d and 4d)

Interestingly, and in contrast to LST, CLM4.5 simulates a small year-round warming effect of forests on daily maximum

5 2 m air temperature (T2M, Fig. 9). This contradicts a number of observational studies which show that the T2M difference of forest minus open land (ΔT2M(f-o)) has the same sign, but is attenuated compared to ΔLST(f-o) (Li et al., 2015; Vanden Broucke et al., 2015; Alkama and Cescatti, 2016; Li et al., 2016). The fact ~~that we ran offline simulations~~ we use offline simulations in our experiments might explain this behaviour, because some land-atmosphere feedbacks ~~could be missed~~ are not represented. However, ~~several studies~~ Lejeune et al. (2017) report similar discrepancies of ΔT2M(f-o) in CLM with observa-

10 tional data for ~~online CLM simulations~~ coupled simulations, suggesting that the behaviour of ΔT2M(f-o) in our simulations may not be related to the ~~offline mode.~~ lack of atmospheric feedbacks.

~~Besides , a number of observational studies investigated the effect of forests on LST and/or T2M.~~ ~~report similar ΔLST(f-o)~~ ~~as , which is not surprising, as both utilize data from the MODIS satellite system. Further, there are a number of studies~~

15 ~~that used paired-site in-situ measurements to infer the local climate impact of forests on LST and/or T2M . CLM - BASE~~ ~~exhibits a weaker latitudinal dependence of ΔT2M$_{avg}$(f-o) than the observational studies of and (e.g., slope of 0.028 in~~ ~~CLM - BASE compared to 0.070 in). In contrast, the latitudinal dependence ΔLST$_{avg}$(f-o) in CLM - BASE (0.10) is stronger~~ ~~than the latitudinal dependence of ΔT2M(f-o) in , consistent with the observation that ΔLST(f-o) is amplified compared~~ ~~with ΔT2M(f-o) . For both variables the shift from a cooling to a warming effect of forests is located further north in~~

20 ~~CLM - BASE (43 N) than the shift documented in the studies using paired-site in-situ measurements (35 N).~~ ~~In the past, a~~ ~~number of modeling studies investigated the temperature effects of deforestation, using different methodologies and land~~ ~~surface models. An assessment of the two model intercomparison projects CMIP5 and LUCID revealed the inability of most~~ ~~land surface models to reproduce the observed effect of forests on daily maximum and minimum temperatures in North~~ ~~America . For CLM4.5, this issue can be resolved to a large extent for ΔLST$_{max}$(f-o) by introducing the SeSC modification~~

25 ~~proposed by (Fig. ??). On the other hand, CLM - BASE still is unable to represent the nighttime warming effect of forests in~~ ~~the mid-latitudes exhibited by observational data . Further, we highlight two modeling studies that can be directly compared to~~ ~~our study. applied a similar comparison strategy using coupled simulations conducted with the GFDL ESM2Mb model. Their~~ ~~results for the canopy temperature difference between cropland and natural vegetation look similar to ours for LST, with a~~ ~~typical shift from a cooling effect of forests at low latitudes to a warming effect north of roughly 45 N. Nevertheless, the region~~

30 ~~with a warming effect of forests extends more towards the equator in the study of . The second study we discuss here is the~~ ~~one of , investigating the contrast between forest and grassland in the MPI-ESM model. This model exhibits a similar daily~~ ~~mean local LST difference of forest minus grassland as CLM - BASE. Further the study of revealed that extensive changes in~~ ~~forest fraction can potentially have substantial non-local impacts not only on LST but also on precipitation, hence emphasizing~~ ~~that contrasting forests and open land locally is not sufficient to fully capture the biogeophysical impact of large-scale de- or~~

35 ~~afforestation.~~ ~~Inadequate representation or omission of several processes in CLM4.5 could be the source of the discrepancies of~~

5   potentially resolve the discrepancies of CLM4.5 with MODIS for ΔLST$_{max}$(f-o) and ΔLST$_{min}$(f-o). The positive bias in ΔLST$_{max}$(f-o) of CLM4.5 could also be related to the negative bias in ΔET(f-o), as increasing/decreasing ET over forest/open land would result in locally lower/higher temperatures.

3.2   Sensitivity Experiment to Alleviate ET Biases in CLM4.5

In the previous section, ~~the most~~ striking discrepancies between the effect of forests in CLM - BASE and observation-based

10   data were found for ΔET(f-o). ~~The main~~ An important driver responsible for these differences was identified to be VTR (Fig. 7). In addition, it became apparent that the SeSC configuration impairs the ΔET(f-o) compared to the ShSC configuration (~~Figs~~Fig. ~~A8 and ??~~6), despite improving ΔLST$_{avg}$(f-o) and ΔLST$_{max}$(f-o) (Fig. ~~??~~8). Hence, in this section we aim to improve the comparability of modeled ΔET(f-o) ~~in CLM - BASE compared~~ to observation-based results by testing a ~~new~~ modified parameterization of VTR in a sensitivity experiment called CLM - PLUS. This ~~configuration of CLM4.5 incorporates~~

15   ~~four additional modifications besides the SeSCs (Table ??) : Shallower root distribution for grass- and cropland PFTs. CLM4.5~~ ~~accounts for SM stress on transpiration through a stress function β_t, which ranges from zero (when soil moisture limitation~~ ~~completely suppresses VTR) to one (corresponding to no soil moisture limitation of VTR). Forests for the most part experience~~ ~~higher SM stress than open land in CLM - DFLT except in the northern high-latitude winter (Fig. A4), partly caused by the~~ ~~similar~~ model configuration comprises (1) a shallower root distribution for ~~all PFTs but evergreen broadleaf trees (Fig. A5).~~

20   ~~In reality, considerably higher maximum rooting depths of forest than grassland and cropland are observed . Likewise, in-situ~~ ~~observations in the tropics show that grassland ET decreases during dry periods, because grasses have only limited access to~~ ~~water reservoirs located below a depth of 2 . Hence, we aim to increase SM stress of~~ open land PFTs~~and reduce their ability to~~ ~~extract water from the lower part of the soil, by introducing a shallower root distribution for these PFTs (Fig. A5) . Dynamic~~ ~~plant water uptake. Tropical forests are often observed to exhibit increased ET during dry periods, due to increased incoming~~

25   ~~shortwave radiation . That is, despite the upper soil being dry, tropical trees still have sufficient access to water from deeper~~ ~~soil layers . We aim to allow a similar behaviour in CLM4.5 by introducing a dynamic~~ , (2) a modified plant water uptake ~~,~~ ~~where~~ scheme whereby plants only extract water from the 10 % of the roots with best access to SM~~(example in Fig. A6) .~~ ~~Light limitation reduction for all~~ , (3) altered light limitation of photosynthesis (decreased for C$_3$ ~~PFTs and enhancement~~ plants and increased for C$_4$ ~~PFTs. In CLM - BASE ET of boreal PFTs is underestimated compared to GETA (Table ??).~~

30   ~~Since VTR of these PFTs is only weakly affected by SM stress, light limitation for C$_3$ plants is reduced. On the other hand,~~ ~~C$_4$ grass shows a considerable positive bias in ET, which we try to alleviate by increasing the light limitation of this PFT.~~ ~~Modified maximum rates of carboxylation (~~plants), and (4) altered V$_{cmax}$ ~~; Table A1). This PFT-specific parameter is suitable~~ ~~to tune VTR, since it is not well constrained from observations and VTR in models is highly sensitive to this parameter .~~ ~~The new values were chosen with the aim to alleviate biases relative to GETA (Table ??) and still lie well within the range of~~

~~observations collected in the TRY plant trait database . Additionally, the minimum stomatal conductance of C₄ plants, which is~~ ~~by default four times larger than that of C₃ plants, was reduced from 40000~~ values to ~~20000 (see Eq. 8.1 in Oleson et al. 2013).~~ ~~A technical description of these modifications as well as a discussion of the effect on ET by each individual modifications is~~ ~~provided in Appendix A.~~alleviate ET biases at PFT level compared to the GETA data.

$\Delta\alpha$(f-o) is only marginally affected by the modifications of CLM-PLUS compared to CLM-BASE (Fig. 3~~a~~a). This is expected since the modifications are targeted at modifying VTR which is not linked directly to albedo. ΔET(f-o) in CLM-PLUS ~~is in better agreement with observations than that~~ becomes more positive than in CLM-BASE ~~. This is mainly due to an~~ ~~increase of ET over forests in tropical regions, thereby alleviating the observation-contradicting sign of the forest signal~~ in all climate zones, thereby better matching the observation-based estimates (Fig. 3e). The improvement is also apparent in the RMSD values which are reduced in CLM-~~BASE~~PLUS for all data sets and climate zones, except for GETA in the E-h climate zone (Fig. ~~6~~4e). The bias in average ET compared to GETA is smaller in CLM-PLUS than in CLM-BASE for all PFTs except for boreal deciduous needleleaf trees and crops (Fig. 3~~f~~f). Some discrepancies with observation-based ET products nevertheless remain. ΔET(f-o) in CLM-PLUS is still mostly less positive compared to remote sensing-based observations and GETA, and remains of opposite sign during the warm season in the temperate regions and in a narrow band around the Equator (~~Fig~~Figs. 6 and 3e). This band originates from a negative ΔET(f-o) around the western part of the equator in Africa and over Indonesia (Fig. ~~3 e). Similarly, the biases of~~ 5). GLEAM and GETA observations cover these areas which explains the only moderate reduction of RMSD of CLM-PLUS against GLEAM and the increase in RMSD against GETA. On the other hand, the RMSD against MODIS is reduced considerably, since MODIS observations are sparse over Africa and Indonesia (Fig. 4e). Also, relative to the in-situ observations of ~~are reduced , but are not completely eliminated (Table 1).~~ ~~Hence, our results call for stronger modifications of the PFT-specific representation of ET in CLM4.5. While testing new model~~ ~~configurations, care should be taken that the implemented modifications do not impair other features of the model, related not~~ ~~only to the water but also the energy and carbon budgets. Reassuringly, we find that global ET values are only weakly affected~~ ~~in the sensitivity experiment, with 70223 in~~ von Randow et al. (2004), biases in CLM-~~BASE compared to 69023~~PLUS are reduced, yet not completely eliminated (Table ~~in CLM-PLUS . These values correspond to averages of 1.43 and 1.41,~~ ~~respectively, which lie within the range of estimates from surface water budgets in the order of 1.2 to 1.5 . Nevertheless, it~~ ~~would be desirable in future studies to evaluate the biogeochemical effects of forests in the different model configurations~~ ~~investigated here alongside the biogeophysical effects of forests.~~1). As a consequence of the improved ΔET(f-o), we find that CLM-PLUS partly alleviates the positive bias in ΔLST$_{max}$(f-o) compared to the MODIS data, especially in the equatorial climate zone ~~(Fig~~which also reduces the RMSD in all but the Arid climate zone (Figs. 3~~b and~~c and 4c). This hints that a realistic representation of ΔET(f-o) is crucial for resolving the underestimated cooling effect of forests on daily maximum LST. ~~The fact that the bias in~~ Similarly, RMSD of ΔLST$_{\overline{max}\,avg}$(f-o) ~~is not completely removed in CLM~~decreases in the Equatorial and Arid climate zones, whereas it increases in the temperate and snow climate zones (Fig. ~~-PLUS is another~~ ~~indication that~~ 4b). At the same time, the RMSD of Δ~~ET(~~LST$_{min}$(f-o~~) is still underestimated in the model.~~We acknowledge ~~that parameter tuning might not be sufficient to remove the ET biases completely. Most values in the parameterization of~~

photosynthesis are shared by all C$_3$ PFTs, which hinders the tuning of the ET difference amongst different PFTs. For example, the introduction of PFT-specific values for some parameters in the calculation of the leaf stomatal resistance would allow for a more effective tuning of ET at PFT level. Additionally, the modifications added in CLM-PLUS only weakly affect ET at high-latitudes (e.g.,) is only marginally increased in all climate zones (Fig. 6). Modifying the temperature dependence of photosynthesis, a parameter which was not taken into consideration in this study, could be beneficial in those regions. 4d).

4 Discussion

The combination of SeSCs and the further modifications introduced in CLM-PLUS led to substantial improvements in CLM4.5's capability to represent forest/open land contrast. Nevertheless, some biases still persist. In particular, CLM4.5 is still unable to represent the nighttime warming effect of forests in the mid-latitudes exhibited by observational data (Lee et al., 2011; Zhang et al., 2014; Vanden Broucke et al., 2015; Li et al., 2015; Alkama and Cescatti, 2016; Li et al., 2016). Additionally, there is a remaining positive bias of $\Delta \mathrm{LST}_{max}$(f-o) compared with MODIS even though this bias is alleviated to some extent due to the more positive $\Delta \mathrm{ET}$(f-o). Inadequate representation or omission of several processes in CLM4.5 could be the source of these discrepancies with MODIS. The biases in both $\Delta \mathrm{LST}_{max}$(f-o) and $\Delta \mathrm{LST}_{min}$(f-o) could be alleviated by accounting for vegetation heat storage, a process which is currently disregarded in CLM4.5. Observed diurnal vegetation heat storage fluxes reach an amplitude of 10–20 $\mathrm{W\,m}^{-2}$ in the mid- and high-latitudes (McCaughey and Saxton, 1988; Lindroth et al., 2010; Kilinc et al., 2012) and 20–70 $\mathrm{W\,m}^{-2}$ in the tropics (Moore and Fisch, 1986; Meesters and Vugts, 1996; dos Santos Michiles and Gielow, 2008). Fluxes of this magnitude are sufficient to considerably alter the diurnal temperature cycle in forests and hence potentially resolve the discrepancies in $\Delta \mathrm{LST}_{max}$(f-o) and $\Delta \mathrm{LST}_{min}$(f-o) of CLM4.5 with MODIS. While $\Delta \mathrm{ET}$(f-o) in CLM-PLUS is improved against all the considered ET data sets in almost every climate zones, some biases persist especially concerning the seasonality in the mid- and high-latitudes as well as annual mean values around the equator. In CLM-PLUS the focus was on VTR, thereby neglecting the contribution from soil and interception evaporation. However, soil evaporation is considerably lower over forests around the equator in CLM-PLUS which might explain the remaining negative $\Delta \mathrm{ET}$(f-o) in this region. We therefore encourage additional sensitivity experiments which also focus on the other components of ET. When testing new model configurations, care should be taken that the implemented modifications do not impair other features of the model, related not only to the water but also the energy and carbon budgets. Reassuringly, we find that global ET averages are only weakly affected in the sensitivity experiment, with an average of 1.43 $\mathrm{mm\,day}^{-1}$ in CLM-BASE compared to 1.41 $\mathrm{mm\,day}^{-1}$ in CLM-PLUS. These values lie within the range of 1.2 $\mathrm{mm\,day}^{-1}$ to 1.5 $\mathrm{mm\,day}^{-1}$ estimated from surface water budgets (Wang and Dickinson, 2012). Nevertheless, it would be desirable in future studies to evaluate the biogeochemical effects of the different model configurations investigated here alongside the biogeophysical effects.

For comparison with LST data we used the radiative temperature in CLM4.5 rather than the more common T2M diagnostic which exhibits an observation-contradicting sign in CLM4.5 (compare Figs. 8e and 9). Such T2M-specific discrepancies with observations could be related to a differing definition of T2M over forests in the model and observations. For example, the differing sign of ΔT2M$_{max}$(f-o) in climate models using CLM and the observations of Lee et al. (2011) found in Lejeune

5 et al. (2017) might be related to the fact that T2M observations were made 2 to 15 m above the forest canopy whereas T2M of CLM4.5 lies within the forest canopy (Oleson et al., 2013). Therefore, T2M in CLM4.5 should be used with care when comparing to observations.

There are several factors which may affect the comparability of the signal extracted from our CLM4.5 simulations and

10 the considered observational data sets. (1) the different data sources use differing land cover information. For example, GLEAM uses the MOD44B product which provides the fraction of each grid cell covered by trees, non-tree vegetation, and non-vegetated land surfaces, whereas MODIS uses MCD12C1 product which provides the dominant IGBP land cover type (Li et al., 2015; Martens et al., 2017). Further, the definition of forest and open land in the Li et al. 2015 dataset can be a source of model-data discrepancy. The methodology applied by Li et al. (2015) relies on the definition of a threshold

15 (80 %) in the coverage of forest (open land) for a pixel to be classified as forest (open land). There are therefore some mixing effects between the forest and open land categories in this dataset, whereas our evaluation method isolates pure signals over forest and open land in CLM4.5. In fact, MODIS albedo retrievals were found to underestimate albedo over grass- and cropland, especially under the presence of snow, and overestimate it over forests due to the heterogeneity of land cover within pixels (Cescatti et al., 2012; Wang et al., 2014). Therefore, it is possible that the magnitude of $\Delta\alpha$(f-o)

20 is underestimated in MODIS rather than overestimated in CLM4.5. Consistently, in-situ observations of paired forest and open land sites support the higher $\Delta\alpha$(f-o) found in CLM-BASE (von Randow et al., 2004; Liu et al., 2005). (3) MODIS LST data are retrieved under clear-sky conditions only, whereas we do not mask out cloudy days in the evaluation of the CLM4.5 simulations. (4) the overpass times of the MODIS satellite system are at 1:30 am and 1:30 pm, hence not necessarily coinciding with the daily maximum and minimum LST in CLM4.5. And finally (5), the meteorological conditions

25 within one search window of MODIS may vary between the different pixels, whereas the different PFT tiles in our CLM4.5 simulations where subject to the exact same atmospheric forcing. However, Li et al. (2015) partly accounted for this effect by applying an elevation adjustment. Moreover, they found little sensitivity of the forest minus open land signal to the size of the chosen window.

30 In this study, we focused on the contrast between forest and open land. However, we acknowledge that future studies should consider other types of land conversions or land management changes, as an increasing number of studies have demonstrated that other LULCCs than de- or reforestation also have remarkable biogeophysical effects (e.g., Davin et al., 2014; Malyshev et al., 2015; Naudts et al., 2016; Thiery et al., 2017; Chen et al., 2018). The two new observation based data sets of Bright et al. (2017) and Duveiller et al. (2018) assess the biogeophysical consequences of a series of different

35 LULCCs globally, thereby enabling the evaluations of the sensitivity to additional types of land cover in future studies. An

17

additional advantage of these two studies is that they both provide a signal for a complete conversion from one land cover type to another (i.e. they do not rely on a coverage threshold as MODIS). In our evaluation approach we focus on the local climatic impact of forests, thereby neglecting feedback mechanisms between the atmosphere and the land surface. While they appear to be relevant in many climate models (Winckler et al., 2017; Devaraju et al., 2018), their evaluation is prevented by the lack of observations at the moment.

5 Conclusions

In this study we evaluate the representation of the local biogeophysical effects of forests in the Community Land Model 4.5 (CLM4.5), using recently published MODIS-based observations of the albedo, evapotranspiration (ET), and land surface temperature (LST) difference between forest and nearby open land. Given the uncertainties in observation-based ET estimates, we further extend our evaluation for this variable by including data from GLEAM v3.1a and GETA 2.0. In our model evaluation we extract a local signal of forests by analyzing PFT-level model output, allowing for good comparability with the high-resolution satellite observations. Further, we use a modified version of CLM4.5 which attributes a separated soil column to each plant functional type (PFT), resulting in a more realistic sub-grid contrast between forest and open land.

Overall, the lower albedo over forests in CLM4.5 is in line with the MODIS observations. However, the albedo contrast between forest and open land is somewhat more pronounced in the model. Ground observations support the stronger albedo contrast in CLM4.5, ~~indicating~~ suggesting that MODIS albedo observations should be used carefully when contrasting different land cover types, as satellite observations tend to retrieve a mixed signal of various land cover types due to their limited spatial resolution. By suppressing lateral ground heat fluxes, the soil column separation considerably improved the representation of the impact of deforestation on daily mean and maximum LST, resulting in a good agreement with the MODIS observations. Both exhibit an overall cooling effect of forests on these variables, except for winter at latitudes exceeding 30° ~~(Fig. 8).~~. Nevertheless, it appeared that the LST difference of forest minus open land in CLM4.5 tends to have a positive bias compared to observational studies. Also, it emerged that caution is required when comparing 2 m air temperature in CLM4.5 to observational data. This variable is only diagnostic in CLM4.5 and might not be conform with measurements, despite realistic LST values. The nighttime warming effect of forests in the mid-latitudes which emerged in a number of recent observational studies, is not represented by CLM4.5. ~~This issue has been observed in other modeling studies using CLM.~~ The biases in the daily maximum and minimum LST signal of forests might be at least partly alleviated by accounting for heat storage in the ~~forest~~ vegetation biomass. We therefore encourage a modification of CLM which enables the representation of canopy heat storage.

Observation-based ET estimates generally agree on higher ET over forests than open land throughout the year at low lati-

tudes and during summer at mid- and high latitudes. This was however not represented by the CLM4.5 configuration using separated soil columns. In fact, the soil column separation impaired the ET signal of forests in CLM4.5, despite improving the LST signal of forests considerably. Hence, a complete evaluation and verification of this modification of CLM4.5 should be undertaken before including it in future versions of CLM. ~~The ET difference of forest minus open land is mainly driven~~

5 ~~by vegetation transpiration. Therefore a revision of the parameterization of transpiration appears necessary to achieve better~~ ~~comparability with observations on the ET difference of forest minus open land.~~ We succeeded in attenuating the biases in ET and also daily ~~mean and~~ maximum LST in a sensitivity experiment which incorporated modifications on four aspects of the parameterization of vegetation transpiration: The root distribution, a dynamic plant water uptake instead of the current static one, the light limitation, and the maximum rate of carboxylation.

10

Historically the most important ~~LULCC~~ Land Use and Land Cover Change (LULCC) process, deforestation is still ongoing ~~mainly in~~ in large parts South America, Africa, and ~~south-east~~ Southeast Asia. A realistic representation of the biogeophysical effects of ~~Land Use and Land Cover Change (LULCC)~~ LULCC in climate models is needed as a number of observational studies revealed that they can have a considerable impact on the local climate. An appropriate representation of the effects of

15 LULCC is not only a feature land surface models need to have in order to understand the ~~past and~~ climate of the past and project future climate, but is also a chance to achieve a more realistic simulation of processes at the land surface. ~~As can be seen~~ ~~from~~ To this end, the analysis of ~~ET in this study,~~ model output at PFT level can ~~reveal~~ help revealing model deficiencies that otherwise would have been hidden below the veil of ~~aggregation and can thus facilitate a better understanding of the underlying~~ ~~processes~~grid-scale aggregation.

Appendix A: Sensitivity of CLM4.5 to Individual Modifications

5   Here we present a more detailed description and discussion of the individual modifications described in Section 3.2. In order to isolate the effect of the individual modifications three additional sensitivity experiments are presented: CLM‑ROOT, CLM‑10PER, and CLM‑LIGHT. Table A4 shows which modifications of CLM4.5 are incorporated in the different sensitivity experiments.

10

A0.1   Sensitivity to Root Distribution

In CLM4.5 ET is strongly and positively correlated to SM at most locations, indicating that SM limitation exerts a strong control on the magnitude of ET (not shown). In CLM‑DFLT, where SM is the same for all PFTs within a grid cell, forest mostly experiences higher SM stress except for the northern high-latitude winter (Fig. A4‑a). Once the SeSCs are introduced
15 in CLM‑BASE, the differences in the SM stress are also influenced by the differences in SM, which in term are affected by the various ET rates over forest and open land. In other terms, it is possible that forests experience less SM stress than open land but only because they evaporate less water and vice versa (Fig. A4‑b). We argue that the difference in the SM stress of forest minus open land in CLM‑DFLT is more representative, because it is unaffected by the ET rates of the individual PFTs in this model configuration. Under this assumption, forests are often more SM-limited than open land in CLM4.5. In contrast,
20 two observational studies comparing SM profiles of forest and nearby pasture sites in the Amazon reveal that forests have a considerably higher capacity to access water from the soil below a depth of $2\,\mathrm{m}$ (Jipp et al., 1998; von Randow et al., 2004). Further, there are a number of studies reporting increased forest ET during the dry season due to the higher amount of incoming shortwave radiation, whilst the response is the opposite over pasture (Jipp et al., 1998; da Rocha et al., 2004; von Randow et al., 2004; Huete et al., 2006; Saleska et al., 2007). Altogether these studies indicate that forest ET should be less SM-limited than
25 open land ET. It is thus possible that forests experience too high and/or open land too little SM stress in CLM4.5.

CLM4.5 accounts for SM stress on VTR through a stress function β_t, which ranges from zero (when soil moisture limitation completely suppresses VTR) to one (corresponding to no SM limitation on VTR). This function is calculated according to Eq. A1 as the sum of the root fraction in each soil layer (r_i) multiplied by a PFT-dependent wilting factor (w_i). The original
30 root distributions in CLM4.5 were adapted from Zeng (2001) and are rather similar for all PFTs, especially for needleleaf trees, broadleaf deciduous trees, and grassland in the lower part of the soil (Fig. A5). Therefore, there is no considerable difference in the default configuration of CLM4.5 regarding the ability to extract water from the lower part of the soil between forests and open land PFTs (except for broadleaf evergreen trees). Furthermore, all tree PFTs have a less negative soil matrix potential at which the stomata are fully closed and opened than the open land ones, i.e., tree PFTs have their permanent wilting point at a higher SM content than open land and hence use water more conservatively. In order to increase SM limitation for open land PFTs and thus reduce their ability to extract water from the lower part of the soil, we conduct a sensitivity experiment, called

CLM‑ROOT, with a much shallower root distribution for open land PFTs. The new values for the root distribution factors (r_a

5   and r_b) are shown in Table A1 and the resulting root distribution in Fig. A5.

$$\beta_t = \sum_i w_i r_i \tag{A1}$$

The modified root distributions strongly reduce the ET of non‑arctic open land PFTs, especially ET of C_4 grass (Table A5). Also, the ET of grassland at the location of the pasture site in the Amazon in the study of von Randow et al. (2004) is consider‑ably reduced during the dry period, even overcompensating the positive bias in CLM‑BASE (Table A6). On the other hand, it

10   does not affect ET during the wet season, when ET is not SM limited. Overall, this experiment reveals that modifying the root distribution has high potential to alleviate biases of CLM4.5 in ET, except for the arctic region where likely temperature and incoming shortwave radiation are the main factors limiting VTR.

15   **A0.2   Sensitivity to Dynamic Plant Water Uptake**

In the tropics forests often exhibit increased ET during dry periods, due to increased light availability (da Rocha et al., 2004; Huete et al., 2006; Saleska et al., 2007), even though the upper soil is dry, as they still have sufficient water supply from the lower part of the soil (Jipp et al., 1998; von Randow et al., 2004). We aim to allow a similar behaviour in CLM4.5 by introduc‑ing a dynamic plant water uptake, where plants only extract water from the 10 % of the roots with the highest wilting factor

20   (i.e., best access to SM) for the calculation of the β_t‑factor and the extraction of soil water (example in Fig. A6). The resulting model simulation, called CLM‑10PER, was conducted by adding this modification to the configuration from the CLM‑ROOT experiment.

This modification generally reduces SM stress for plants and hence increases ET for all non‑arctic PFTs (Table A5). Its impact

25   is limited for arctic PFTs where temperature and shortwave radiation are more important limiting factors of VTR than water availability. A notable improvement can be observed for tropical deciduous broadleaf trees for which average ET is increased by $0.11\,\mathrm{mm\,day^{-1}}$, thereby alleviating the negative bias compared to GETA. Furthermore, it improves the seasonal dynamics of forest ET in the tropics. With the 10 % modification forests show increased ET during the dry period at the forest site of da Rocha et al. (2004). This is the case as trees are now less SM‑limited during the dry period than in CLM‑BASE, since they

30   have a significant fraction of their roots in the ~~still moist~~ still‑moist lower part of the soil, allowing them to exploit the increase in incoming shortwave radiation. On the other hand, ET at the pasture site of von Randow et al. (2004) remains largely unaf‑fected, as grassland has only limited access to SM from the lower part of the soil due to the shallow root distribution introduced in CLM‑ROOT. It hence appears that a dynamic plant water uptake could be crucial for the representation of the seasonal dynamics of ET (and possibly photosynthetic activity in general) in the tropics.

21

A0.3 Sensitivity to Light Limitation

As arctic PFTs are only weakly affected by the previously introduced modifications of SM stress as well as the maximum rate of carboxylation described in the next section, we performed a sensitivity experiment with altered light limitation, which is called CLM‑LIGHT. Since ET values are strongly negatively biased for boreal deciduous broadleaf trees and C_3 arctic grass (Table A5), the light limitation of photosynthesis for C_3 plants was lessened by increasing the factor 0.5 in Eq. 8.7 of Oleson et al. (2013) to 0.6. Because ET of C_4 grass exhibits a strong positive bias, their quantum efficiency was reduced from 0.05 to 0.025 $\mathrm{mol\,CO_2\,mol^{-1}\,photon}$, thereby increasing their light limitation.

Altering the light limitation of photosynthesis impacts ET in all climate zones (Table A5). Its impact is strongest in the tropics and remains small in boreal regions. Of the C_3 PFTs tropical evergreen broadleaf trees are impacted strongest. The implemented modification alleviates the negative ET bias for evergreen broadleaf trees during the dry season but slightly increases the positive bias during the wet season, overall still leading to a further improvement of the difference between the two seasons (Table A6). Additionally, the increased light limitation reduces ET of C_4 grass during the wet season similar to the observations over the grassland site in von Randow et al. (2004). This is likely responsible for the increased ET during the dry season as well, since the reduced SM consumption during the wet season is carried over to the following dry season, therefore reducing the SM stress.

A0.4 Sensitivity to the Maximum Rate of Carboxylation

V_{cmax} appears to be a suitable parameter to tune VTR values, since it is not well constrained from observations and VTR in models is highly sensitive to this parameter (Bonan et al., 2011). In CLM4.5 the values reported by Kattge et al. (2009) are used except for tropical evergreen broadleaf trees, for which a higher value was chosen to alleviate model biases (Bonan et al., 2012; Oleson et al., 2013). In order to test the sensitivity of the PFT-specific ET values to V_{cmax}, we conduct a final sensitivity experiment with new values of this parameter ~~(Table A1)~~ in addition to the other modifications presented beforehand, with the aim to alleviate the biases to GETA (Table A1). Additionally, the minimum stomatal conductance of C_4 plants, which is by default four times larger than that of C_3 plants, was reduced from 40000 $\mathrm{\mu mol\,m^{-2}\,s^{-1}}$ to 20000 $\mathrm{\mu mol\,m^{-2}\,s^{-1}}$ (see Eq. 8.1 in Oleson et al. 2013) in this sensitivity experiment, which we call CLM‑PLUS.

As already shown by Bonan et al. (2011), photosynthetic activity of C_3 PFTs is strongly influenced by the choice of V_{cmax}, except for the boreal ones where light or temperature are more important limiting factors of photosynthesis~~(Tables A5 and A6)~~. The CLM‑PLUS simulation alleviates biases in ET averaged for the individual PFTs compared to GETA, in particular by reducing ET over temperate evergreen needleleaf trees, both temperate and tropical evergreen broadleaf trees, and C_4 grass, as well as by increasing ET of tropical deciduous broadleaf trees (Table A5). The mismatch between results of CLM4.5 and the

in-situ measurements of von Randow et al. (2004) and da Rocha et al. (2004) in the Amazon region are reduced in this new

5 configuration during the wet season, but enhanced during the dry one (Tables A6). As in the CLM‑LIGHT experiment the reduction of C_4 grass ET during the wet season at the pasture site of von Randow et al. (2004) is partly compensated by an ET increase during the dry period. Overall, ET of C_4 grass compares well with the mean value of GETA. The in-situ observations of von Randow et al. (2004) on the other hand support a stronger tuning for this particular PFT in order to further reduce its ET.

References

5   Alkama, R. and Cescatti, A.: Biophysical climate impacts of recent changes in global forest cover, Science, 351, 600–604, https://doi.org/10.1126/science.aac8083, 2016.

Ambrose, S. M. and Sterling, S. M.: Global patterns of annual actual evapotranspiration with land-cover type: Knowledge gained from a new observation-based database, Hydrol. Earth Syst. Sc. Discuss., 11, 12 103–12 135, https://doi.org/10.5194/hessd-11-12103-2014, 2014.

Baldocchi, D. and Ma, S.: How will land use affect air temperature in the surface boundary layer? Lessons learned from a
10   comparative study on the energy balance of an oak savanna and annual grassland in California, USA, Tellus B, 65, 19 994, https://doi.org/10.3402/tellusb.v65i0.19994, 2013.

Boenisch, G. and Kattge, J.: TRY Plant Trait Database (https://www.try-db.org/TryWeb/Home.php), 2017.

Bonan, G.: Forests and Climate Change: Forcings, Feedbacks, and the Climate Benefits of Forests, Science, 320, 1444–1449, https://doi.org/10.2307/20054256, 2008.

15   Bonan, G., Lawrence, P. J., Oleson, K. W., Levis, S., Jung, M., Reichstein, M., Lawrence, D. M., and Swenson, S. C.: Improving canopy processes in the Community Land Model version 4 (CLM4) using global flux fields empirically inferred from FLUXNET data, J. Geophys. Res.-Biogeo., 116, https://doi.org/10.1029/2010JG001593, 2011.

Bonan, G., Oleson, K. W., Fisher, R. A., Lasslop, G., and Reichstein, M.: Reconciling leaf physiological traits and canopy flux data: Use of the TRY and FLUXNET databases in the Community Land Model version 4, J. Geophys. Res.-Biogeo., 117,
20   https://doi.org/10.1029/2011JG001913, 2012.

Bright, R. M., Davin, E. L., O'Halloran, T., Pongratz, J., Zhao, K., and Cescatti, A.: Local temperature response to land cover and management change driven by non-radiative processes, Nat. Clim. Change, 7, 296–302, https://doi.org/10.1038/nclimate3250, 2017.

Brown, S., Sathaye, J., Cannell, M., and Kauppi, P. E.: Mitigation of carbon emissions to the atmosphere by forest management, Commonw. Forest. Rev., 75, 80–91, 1996.

25   Canadell, J., Jackson, R. B., Ehleringer, J. B., Mooney, H. A., Sala, O. E., and Schulze, E.-D.: Maximum rooting depth of vegetation types at the global scale, Oecologia, 108, 583–595, https://doi.org/10.1007/BF00329030, 1996.

Cescatti, A., Marcolla, B., Vannan, S. K. S., Pan, J. Y., Román, M. O., Yang, X., Ciais, P., Cook, R. B., Law, B. E., Matteucci, G., Migliavacca, M., Moors, E., Richardson, A. D., Seufert, G., and Schaaf, C. B.: Intercomparison of MODIS albedo retrievals and in situ measurements across the global FLUXNET network, Remote Sens. Environ., 121, 323–334, https://doi.org/10.1016/j.rse.2012.02.019, 2012.

30   Chen, L., Dirmeyer, P. A., Guo, Z., and Schultz, N. M.: Pairing FLUXNET sites to validate model representations of land-use/land-cover change, Hydrol. Earth Syst. Sc. Discuss., 22, 111–125, https://doi.org/10.5194/hess-22-111-2018, 2018.

Ciais, P., Sabine, C., Bala, G., Bopp, L., Brovkin, V., Canadell, J., A., C., DeFries, R., Galloway, J., Heimann, M., Jones, C., Le Quéré, C., Myneni, R. B., Piao, S., and Thornton, P.: Carbon and Other Biogeochemical Cycles. In: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panle on Climate Change [Stocker, T.F.,
35   D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley (eds.)], Cambridge University Press, 2013.

da Rocha, H. R., Goulden, M. L., Miller, S. D., Menton, M. C., Pinto, L. D. V. O., de Freitas, H. C., and e Silva Figueira, A. M.: Seasonality of Water and Heat Fluxes Over a Tropical Forest in Eastern Amazonia, Ecol. Appl., 14, 22–32, https://doi.org/10.1890/02-6001, 2004.

Davin, E. L. and de Noblet-Ducoudré, N.: Climatic impact of global-scale deforestation: Radiative versus nonradiative processes, J. Climate, 23, 97–112, https://doi.org/10.1175/2009JCLI3102.1, 2010.

Davin, E. L., de Noblet-Ducoudré, N., and Friedlingstein, P.: Impact of land cover change on surface climate: Relevance of the radiative forcing concept, Geophys. Res. Lett., 34, https://doi.org/10.1029/2007GL029678, 2007.

Davin, E. L., Seneviratne, S. I., Ciais, P., Olioso, A., and Wang, T.: Preferential cooling of hot extremes from cropland albedo management, P. Natl. Acad. Sci. USA, 111, 9757–9761, https://doi.org/10.1073/pnas.1317323111, 2014.

De Noblet-Ducoudré, N., Boisier, J.-P., Pitman, A., Bonan, G. B., Brovkin, V., Cruz, F., Delire, C., Gayler, V., van den Hurk, B. J. J. M., Lawrence, P. J., van der Molen, M. K., Müller, C., Reick, C. H., Strengers, B. J., and Voldoire, A.: 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, J. Climate, 25, 3261–3181, https://doi.org/10.1175/JCLI-D-11-00338.1, 2012.

Devaraju, N., de Noblet-Ducoudré, N., Quesada, B., and Bala, G.: Quantifying the Relative Importance of Direct and Indirect Biophysical Effects of Deforestation on Surface Temperature and Teleconnections, J. Climate, 31, 3811 – 3829, https://doi.org/10.1175/JCLI-D-17-0563.1, 2018.

Dorman, J. L. and Sellers, P. J.: A Global Climatology of Albedo, Roughness Length and Stomatal Resistance for Atmospheric General Circulation Models as Represented by the Simple Biosphere Model (SiB), J. Appl. Meteorol., 28, 833–855, https://doi.org/10.1175/1520-0450(1989)028<0833:AGCOAR>2.0.CO;2, 1989.

dos Santos Michiles, A. A. and Gielow, R.: Above-ground thermal energy storage rates, trunk heat fluxes and surface energy balance in a central Amazonian rainforest, Agr. Forest Meteorol., 148, 917–930, https://doi.org/10.1016/j.agrformet.2008.01.001, 2008.

Duveiller, G., Fasbender, D., and Meroni, M.: Revisiting the concept of a symmetric index of agreement for continuous datasets, Sci. Rep.-UK, 6, 19 401, https://doi.org/10.1038/srep19401, 2016.

Duveiller, G., Hooker, J., and Cescatti, A.: The mark of vegetation change on Earth's surface energy balance, Nat. Commun., 9, https://doi.org/10.5194/essd-2018-24, 2018.

Fan, Y., Miguez-Macho, G., Jobbágy, E. G., Jackson, R. B., and Otero-Casal, C.: Hydrologic regulation of plant rooting depth, P. Natl. Acad. Sci. USA, 114, 10 572–10 577, https://doi.org/10.1073/pnas.1712381114, 2017.

Findell, K. L., Shevliakova, E., Milly, P. C. D., and Stouffer, R.-J.: Modeled impact of anthropogenic land cover change on climate, J. Climate, 20, 3621–3634, https://doi.org/10.1175/JCLI4185.1, 2007.

Gash, J. H. C. and Stewart, J. B.: An analytical model of rainfall interception by forests, Q. J. Roy. Meteor. Soc., 105, 43–55, 1979.

Hansen, M. C., Potapov, P. V., Moore, R., Hancher, M., Turubanova, S. A., Tyukavina, A., Thau, D., Stehman, S. V., Goetz, S. J., Loveland, T. R., Kommareddy, A., Egorov, A., Chini, L., Justice, C. O., and Townshend, J. R. G.: High-Resolution Global Maps of 21st-Century Forest Cover Change, Science, 342, 850–853, https://doi.org/10.1126/science.1244693, 2013.

Harding, R. J., Gryning, S.-E., Halldin, S., and Lloyd, C. R.: Progress in understanding of land surface/atmosphere exchanges at high latitudes, Theor. Appl. Climatol., 70, 5–18, https://doi.org/10.1007/s007040170002, 2001.

Harris, I., Jones, P. D., Osborn, T. J., and Lister, D. H.: Updated high-resolution grids of monthly climatic observations – the CRU TS3.10 Dataset, Int. J. Climatol., 34, 623–642, https://doi.org/10.1002/joc.3711, 2014.

Huang, C., Kim, S., Song, K., Townshend, J. R., Davis, P., Altstatt, A., Rodas, O., Yanosky, A., Clay, R., Tucker, C. J., and Musinsky, J.: Assessment of Paraguay's forest cover change using Landsat observations, Global Planet. Change, 67, 1–12, https://doi.org/10.1016/j.gloplacha.2008.12.009, 2009.

Huete, A. R., Didan, K., Shimabukuro, Y. E., Ratana, P., Saleska, S. R., Hutyra, L. R., Yang, W., Nemani, R. R., and Myneni, R.: Amazon rainforests green-up with sunlight in dry season, Geophys. Res. Lett., 33, https://doi.org/10.1029/2005GL025583, 2006.

Hurrell, J. W., Holland, M. M., Gent, P. R., Ghan, S., Kay, J. E., Kushner, P. J., Lamarque, J.-F., Large, W. G., Lawrence, D., Lindsay, K., Lipscomb, W. H., Long, M. C., Mahowald, N., Marsh, D. R., Neale, R. B., Rasch, P., Vavrus, S., Vertenstein, M., Bader, D., Collins, W. D., Hack, J. J., Kiehl, J., and Marshall, S.: The Community Earth System Model: A Framework for Collaborative Research, B. Am. Meteorol. Soc., 94, 1339–1360, https://doi.org/10.1175/BAMS-D-12-00121.1, 2013.

Jipp, P. H., Nepstad, D. C., Cassel, D. K., and de Carvalho, C. R.: Deep Soil Moisture Storage and Transpiration in Forests and Pastures of Seasonally-dry Amazonia, Climate Change, 39, 395–412, https://doi.org/10.1023/A:1005308930871, 1998.

Juang, J.-Y., Katul, G., Siqueira, M., Stoy, P., and Novick, K.: Separating the effects of albedo from eco-physiological changes on surface temperature along a successional chronosequence in the southeastern United States, Geophys. Res. Lett., 34, https://doi.org/10.1029/2007GL031296, 2007.

Kattge, J., Knorr, W., Raddatz, T., and Wirth, C.: Quantifying photosynthetic capacity and its relationship to leaf nitrogen content for global-scale terrestrial biosphere models, Glob. Change Biol., 15, 976–991, https://doi.org/10.1111/j.1365-2486.2008.01744.x, 2009.

Kilinc, M., Beringer, J., Hutley, L. B., Haverd, V., and Tapper, N.: An analysis of the surface energy budget above the world's tallest angiosperm forest, Agr. Forest Meteorol., 166, 23–31, https://doi.org/10.1016/j.agrformet.2012.05.014, 2012.

Kottek, M., Grieser, J., Beck, C., Rudolf, B., and Rubel, F.: World Map of the Köppen-Geiger climate classification updated, Meteorol. Z., 15, 259–263, https://doi.org/10.1127/0941-2948/2006/0130, 2006.

Kumar, S., Dirmeyer, P. A., Merwade, V., DelSole, T., Adams, J. M., and Niyogi, D.: Land use/cover change impacts in CMIP5 climate simulations: A new methodology and 21st century challenges, J. Geophys. Res., 118, 6337–6353, https://doi.org/10.1002/jgrd.50, 2013.

Lawrence, P. J. and Chase, T. N.: Representing a new MODIS consistent land surface in the Community Land Model (CLM 3.0), J. Geophys. Res.-Biogeo., 112, https://doi.org/10.1029/2006JG000168, 2007.

Lee, X., Goulden, M. L., Hollinger, D. Y., Barr, A., Black, T. A., Bohrer, G., Bracho, R., Drake, B., Goldstein, A., Gu, L., Katul, G., Kolb, T., Law, B. E., Margolis, L. H., Meyers, T., Monson, R., Munger, W., Oren, R., Paw U, K. T., Richardson, A. D., Schmid, H. P. Staebler, R., Wofsy, S., and Zhao, L.: Observed increase in local cooling effect of deforestation at higher latitude, Nature, 479, 384–387, https://doi.org/10.1038/nature10588, 2011.

Lejeune, Q., Seneviratne, S. I., and Davin, E. L.: Historical Land-Cover Change Impacts on Climate: Comparative Assessment of LUCID and CMIP5 Multimodel Experiments, J. Climate, pp. 1439–1459, https://doi.org/10.1175/JCLI-D-16-0213.1, 2017.

Lejeune, Q., Davin, E. L., Gudmundsson, L., Winckler, J., and Seneviratne, S. I.: Historical deforestation locally increased the intensity of hot days in northern mid-latitudes, Nat. Clim. Change, 8, 386–390, https://doi.org/10.1038/s41558-018-0131-z; 10.1038/s41558-018-0131-z, 2018.

Li, Y., Zhao, M., Motesharrei, S., Mu, Q., Kalnay, E., and Li, S.: Local cooling and warming effects of forests based on satellite observations, Nat. Commun., 6, https://doi.org/10.1038/ncomms7603, 2015.

Li, Y., Zhao, M., Mildrexler, D. J., Motesharrei, S., Mu, Q., Kalnay, E., Zhao, F., Li, S., and Wang, K.: Potential and Actual impacts of deforestation and afforestation on land surface temperature, J. Geophys. Res., 121, 14 372–14 386, https://doi.org/10.1002/2016JD024969, 2016.

Lindroth, A., Mölder, M., and Lagergren, F.: Heat storage in forest biomass improves energy balance closure, Biogeoscience, 7, https://doi.org/10.5194/bg-7-301-2010, 2010.

Liu, H., Randerson, J. T., Lindfors, J., and Chapin III, F. S.: Changes in the surface energy budget after fire in boreal ecosystems of interior Alaska: An annual perspective, J. Geophys. Res., 110, https://doi.org/10.1029/2004JD005158, 2005.

Malyshev, S., Shevliakova, E., Stouffer, R. J., and Pacala, S. W.: Contrasting Local versus Regional Effects of Land-Use-Change-Induced Heterogeneity on Historical Climate: Analysis with the GFDL Earth System Model, J. Climate, 28, 5448–5469, https://doi.org/10.1175/JCLI-D-14-00586.1, 2015.

Margono, B. A., Potapov, P. V., Turubanova, S. A., Stolle, F., and Hansen, M. C.: Primary forest cover loss in Indonesia over 2000–2012, Nat. Clim. Change, 4, 730–735, https://doi.org/10.1038/nclimate2277, 2014.

Martens, B. and Miralles, D. G.: Global Land Evaporation Amsterdam Model, http://www.gleam.eu/, 2017.

Martens, B., Miralles, D. G., Lievens, H., van der Schalie, R., de Jeu, R. A. M., Fernández-Prieto, D., Beck, H. E., Dorigo, W. A., and Verhoest, N. E. C.: GLEAM v3: satellite-based land evaporation and root-zone soil moisture, Geosci. Model Dev., 10, https://doi.org/10.5194/gmd-2016-162, 2017.

McCaughey, J. and Saxton, W.: Energy balance storage terms in a mixed forest, Agr. Forest Meteorol., 44, 1–18, https://doi.org/10.1016/0168-1923(88)90029-9, 1988.

McGrath, M. J., Luyssaert, S., Meyfroidt, P., Kaplan, J. O., Bürgi, M., Chen, Y., Erb, K., Gimmi, U., McInerney, D., Naudts, K., Otto, J., Pasztor, F., Ryder, J., Schelhaas, M.-J., and Valade, A.: Reconstructing European forest management from 1600 to 2010, Biogeosciences, 12, 4291–4316, https://doi.org/10.5194/bg-12-4291-2015, 2015.

Meesters, A. and Vugts, H.: Calculation of heat storage in stems, Agr. Forest Meteorol., 78, 181–202, https://doi.org/10.1016/0168-1923(95)02251-1, 1996.

Michel, D., Jiménez, C., Miralles, D. G., Jung, M., Hirschi, M., Ershadi, A., Martens, B., McCabe, M. F., Fisher, J. B., Mu, Q., Seneviratne, S. I., Wood, E. F., and Fernández-Prieto, D.: The WACMOS-ET project - Part 1: Tower-scale evaluation of four remote-sensing-based evapotranspiration algorithms, Hydrol. Earth Syst. Sc., 20, 803–822, https://doi.org/10.5194/hess-20-803-2016, 2016.

Miralles, D. G., Holmes, T. R. H., De Jeu, R. A. M., Gash, J. H., Meesters, A. G. C. A., and Dolman, A. J.: Global land-surface evaporation estimated from satellite-based observations, Hydrol. Earth Syst. Sc., 15, 454–469, https://doi.org/10.5194/hess-15-453-2011, 2011.

Miralles, D. G., Jiménez, C., Jung, M., Michel, D., Ershadi, A., McCabe, M. F., Hirschi, M., Martens, B., Dolman, A. J., Fisher, J. B., Mu, Q., Seneviratne, S. I., Wood, E. F., and Fernández-Prieto, D.: The WACMOS-ET project - Part 2: Evaluation of global terrestrial evaporation data sets, Hydrol. Earth Syst. Sc., 20, 823–842, https://doi.org/10.5194/hess-20-823-2016, 2016.

Moore, C. and Fisch, G.: Estimating heat storage in Amazonian tropical forest, Agr. Forest Meteorol., 38, 147–168, https://doi.org/10.1016/0168-1923(86)90055-9, 1986.

Mu, Q. Z., Zhao, M. S., and Running, S. W.: Improvements to a MODIS global terrestrial evapotranspiration algorithm, Remote Sens. Environ., 115, 1781–1800, https://doi.org/10.1016/j.rse.2011.02.019, 2011.

Mueller, B., Seneviratne, S. I., Jimenez, C., Corti, T., Hirschi, M., Balsamo, G., Ciais, P., Dirmeyer, P., Fisher, J. B., Guo, Z., Jung, M., Maignan, F., McCabe, M. F., Reichle, R., Reichstein, M., Rodell, M., Sheffield, J., Teuling, A. J., Wang, K., Wood, E. F., and Zhang, Y.: Evaluation of global boservations-based evapotranspiration datasets and IPCC AR4 simulations, Geophys. Res. Lett., 38, https://doi.org/10.1029/2010GL046230, 2011.

Mueller, B., Hirschi, M., Jimenez, C., Ciais, P., Dirmeyer, P. A., Dolman, A. J., Fisher, J. B., Jung, M., Ludwig, F., Maignan, F., Miralles, D. G., McCabe, M. F., Reichstein, M., Sheffield, J., Wang, K., Wood, E. F., Zhang, Y., and Seneviratne, S. I.: Benchmark products for land evapotranspiration: LandFlux-EVAL multi-data set synthesis, Hydrol. Earth Syst. Sc., 17, 3707–3720, https://doi.org/10.5194/hess-17-3707-2013, 2013.

Naudts, K., Chen, Y., McGrath, M. J., Ryder, J., Valade, A., Otto, J., and Luyssaert, S.: Europe's forest management did not mitigate climate warming, Science, 351, 597–600, https://doi.org/10.1126/science.aad7270, 2016.

Oleson, K. W., Lawrence, D., B. Bonan, G., Drewniak, B., Huang, M., Koven, C., Levis, S., Li, F., Riley, W., M. Subin, Z., C. Swenson, S., E. Thornton, P., Bozbiyik, A., Fisher, R., Heald, C., Kluzek, E., Lamarque, J.-F., Lawrence, P., Leung, L., and Yang, Z.-L.: Technical Description of version 4.5 fo the Community Land Model (CLM), https://doi.org/10.5065/D6RR1W7M, 2013.

Pitman, A. J., de Noblet-Ducoudré, N., Cruz, F. T., Davin, E. L., Bonan, G. B., Brovkin, V., Claussen, M., Delire, C., Ganzeveld, L., Gayler, V., van den Hurk, B. J. J. M., Lawrence, P. J., van der Molen, M. K., Müller, C., Reick, C. H., Seneviratne, S. I., Strengers, B. J., and Voldoire, A.: Uncertainties in climate responses to past land cover change: First results from the LUCID intercomparison study, Geophys. Res. Lett., 36, https://doi.org/10.1029/2009GL039076, 2009.

Pongratz, J., Reick, C. H., Raddatz, T., and Claussen, M.: Biogeophysical versus biogeochemical climate response to historical anthropogenic land cover change, Geophys.Res. Lett., 37, https://doi.org/10.1029/2010GL043010, 2010.

Priestly, C. H. B. and Taylor, R. J.: On the assessment of surface heat flux and evaporation using large-scale parameters, Mon. Weather Rev., 100, 81–92, 1972.

Saleska, S. R., Didan, K., Huete, A. R., and da Rocha, H. R.: Amazon Forests Green-Up During 2005 Drought, Science, 318, 612, https://doi.org/10.1126/science.1146663, 2007.

Schultz, N. M., Lee, X., Lawrence, P. J., Lawrence, D. M., and Zhao, L.: Assessing the use of subgrid land model output to study impacts of land cover change, J. Geophys. Res., 121, 6133–6147, https://doi.org/10.1002/2016JD025094, 2016.

Sonntag, S., Pongratz, J., Reick, C. H., and Schmidt, H.: Reforestation in a high-CO2 world—Higher mitigation potential than expected, lower adaptation potential than hoped for, Geophys. Res. Lett., 43, 6546–6553, https://doi.org/10.1002/2016GL068824, 2016.

Stoy, P., Katul, G., Siqueira, M., Juang, J.-Y., Mccarthy, H., Oishi, A., M. Uebelherr, J., Kim, H.-S., and Oren, R.: Separating the Effects of Climate and Vegetation on Evapotranspiration along a Successional Chronosequence in the Southeastern US, Glob. Change Biol., 12, 2115–2135, https://doi.org/10.1111/j.1365-2486.2006.01244.x, 2006.

Teuling, A. T., Seneviratne, S. I., Stöckli, R., Reichstein, M., Moors, E., Ciais, P., Luyssaert, S., van den Hurk, B., Ammann, C., Bernhofer, C., Dellwik, E., Gianelle, D., Gielen, B., Grünwald, T., Klumpp, K., Montagnani, L., Moureaux, C., Scottocornola, M., and Wohlfahrt, G.: Contrasting response of European forest and grassland energy exchange ot heatwaves, Nat. Geosci., 3, 722–727, https://doi.org/10.1038/ngeo950, 2010.

Thiery, W., Davin, E. L., Lawrence, D. M., Hirsch, A. L., Hauser, M., and Seneviratne, S. I.: Present-day irrigation mitigates heat extremes, J. Geophys. Res., 122, 1403–1422, https://doi.org/10.1002/2016JD025740, 2017.

Van der Molen, M., Dolman, A., Waterloo, M., and Bruijnzeel, L.: Climate is affected more by maritime than by continental land use change: A multiple scale analysis, Global Planet. Change, 54, 128–149, https://doi.org/10.1016/j.gloplacha.2006.05.005, 2006.

Vanden Broucke, S., Luyssaert, S., Davin, E. L., Janssens, I., and van Lipzig, N.: New insights in the capability of climate models to simulate the impact of LUC based on temperature decomposition of paired site observations, J. Geophys. Res.-Atmos., 120, 5417–5436, https://doi.org/10.1002/2015JD023095, 2015.

Vivoy, N.: CRUNCEP dataset, ftp://nacp.ornl.gov/synthesis/2009/frescati/temp/land_use_change/original/readme.htm, 2009.

von Randow, C., Manzi, A. O., Kruijt, B., de Oliveira, P. J., Zanchi, F. B., Silva, R. L., Hodnett, M. G., Gash, J. H. C., Elbers, J. A., Waterloo, M. J., Cardoso, F. L., and Kabat, P.: Comparative measurements and seasonal variations in energy and carbon exchange over forest and pasture in South West Amazonia, Theor. Appl. Climatol., 78, 5–26, https://doi.org/10.1007/s00704-004-0041-z, 2004.

Wang, K. and Dickinson, R. E.: A Review of Global Terrrestrial Evapotranspiration: Observation, Modeling, Climatology, and Climatic Variablility, Rev. Geophys., 50, https://doi.org/10.1029/2011RG000373, 2012.

Wang, Z., Schaaf, C. B., Strahler, A. H., Chopping, M. J., Román, M. O., Shuai, Y., Woodcock, C. E., Hollinger, D. Y., and Fitzjarrald, D. R.:

Evaluation of MODIS albedo product (MCD43A) over grassland, agriculture and forest surface types during dormant and snow-covered periods, Remote Sens. Environ., 140, 60–77, https://doi.org/10.1016/j.rse.2013.08.025, 2014.

Winckler, J., Reick, C. H., and Pongratz, J.: Robust Identification of Local Biogeophysical Effects of Land-Cover Change in a Global Climate Model, J. Climate, 30, 1159–1176, https://doi.org/10.1175/JCLI-D-16-0067.1, https://doi.org/10.1175/JCLI-D-16-0067.1, 2017.

Wolf, S., Eugster, W., Majorek, S., and Buchmann, N.: Afforestation of Tropical Pasture Only Marginally Affects Ecosystem-Scale Evapotranspiration, Ecosystems, 14, 1264–1275, https://doi.org/10.1007/s10021-011-9478-y, 2011.

Zeng, X.: Global Vegetation Root Distribution for Land Modeling, J. Climate, https://doi.org/10.1175/1525-7541(2001)002<0525:GVRDFL>2.0.CO;2, 2001.

Zhang, M., Lee, X., Yu, G., Han, S., Wang, H., Yan, J., Zhang, Y., Li, Y., Ohta, T., Hirano, T., Kim, J., Yoshifuji, N., and Wang, W.: Response of surface air temperature to small-scale land clearing across latitudes, Environ. Reas. Lett., 9, https://doi.org/10.1088/1748-9326/9/3/034002, 2014.
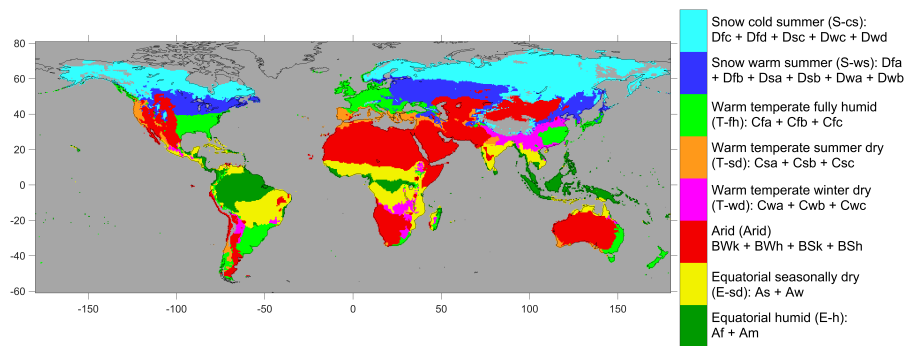
Figure 1. The Köppen-Geiger climate zones (Kottek et al., 2006) used for the analysis.
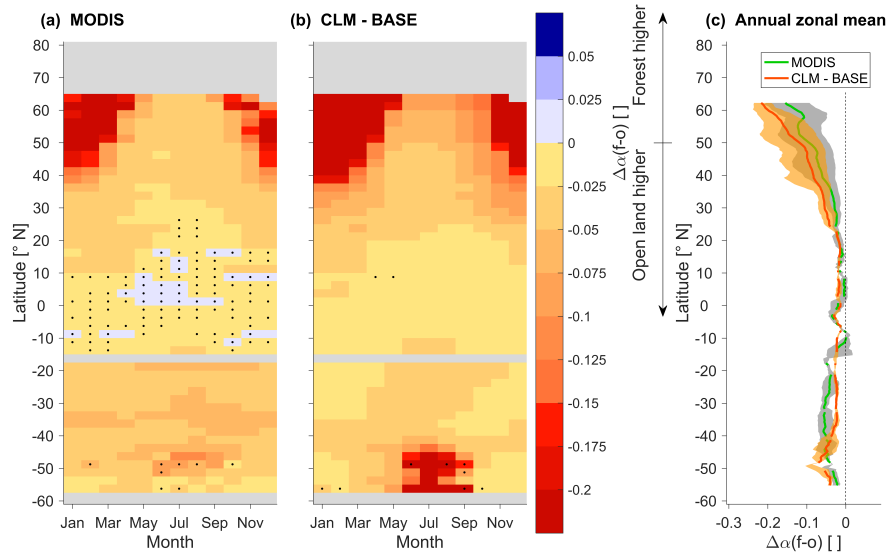
Figure 2. Seasonal and latitudinal variations of $\Delta\alpha$(f-o) in (a) the MODIS observations and (b) CLM‑BASE. Points with a mean which is insignificantly different from zero in a two-sided t-test at 95 % confidence level are marked with a black dot. All data from the 2002-2010 analysis period corresponding to a given latitude and a given month are pooled to derive the sample set for the test. Panel (c) shows the zonal annual mean of both MODIS (in green along with its ~~confidence interval~~ interquartile range in grey) and CLM‑BASE (in red, ~~confidence interval~~ interquartile range in orange). Note that on this subfigure results have been smoothed with a 4° latitudinally-running mean. Only grid cells containing valid data in the MODIS observations were considered for the analysis of CLM‑BASE.
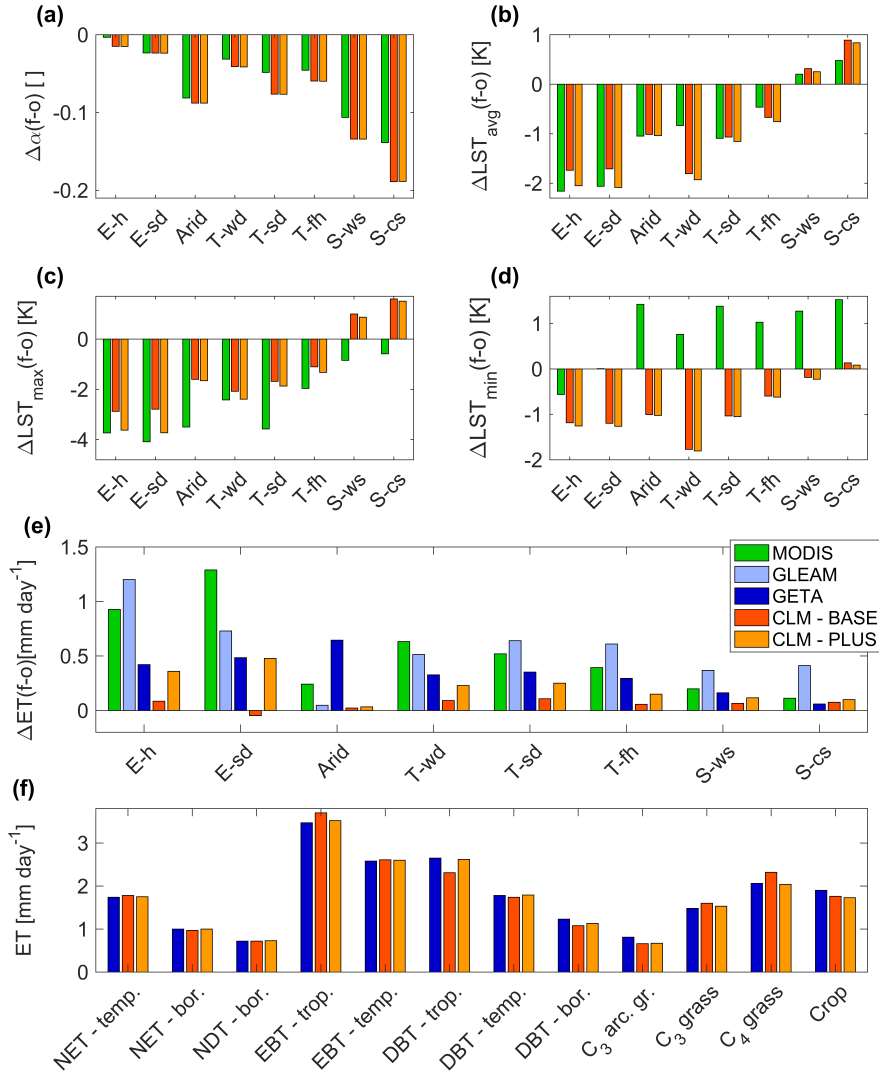
31

Figure 3. ~~Improvements in CLM‑PLUS compared to CLM‑BASE.~~ Area-weighted annual mean over Köppen-Geiger climate zones (Kottek et al., 2006, ; Fig. 1) of (a) $\Delta\alpha$(f-o), (b) $\Delta\mathrm{LST}_{avg}$(f-o), (c) $\Delta\mathrm{LST}_{max}$(f-o), and (d) $\Delta\mathrm{LST}_{min}$(f-o) in MODIS (green), CLM‑BASE (red), and CLM‑PLUS (orange). Only grid cells containing valid data in the MODIS observations were considered for analysis of CLM4.5. Panel (e) shows the ~~area weighted~~ area-weighted mean over the Köppen-Geiger climate zone of $\Delta\mathrm{ET}$(f-o) in MODIS (green), GLEAM (light blue), GETA (dark blue), CLM‑BASE (red), and CLM‑PLUS (orange) and panel (f) the ~~area weighted~~ area-weighted mean ET for each PFT analyzed in this study according to the GETA (dark blue), CLM‑BASE (red), and CLM‑PLUS (orange). The acronyms of the PFTs are defined in Table ~~??~~A2.
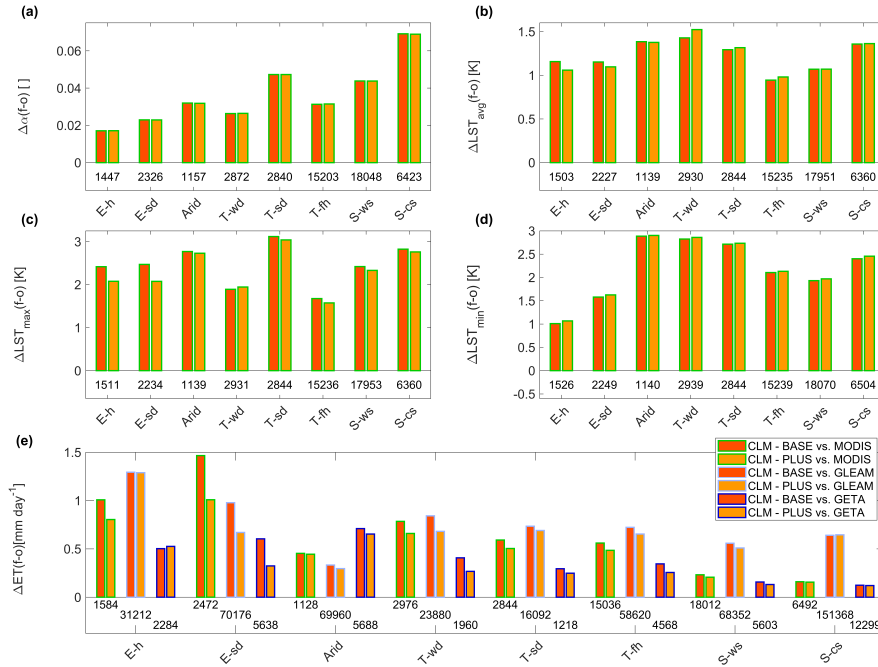
Figure 4. RMSD of CLM‑BASE (red), and CLM‑PLUS (orange) against MODIS observations over Köppen-Geiger climate zones (Kottek et al., 2006, ; Fig. 1) of monthly (a) $\Delta\alpha$(f-o), (b) ΔLST_{avg}(f-o), (c) ΔLST_{max}(f-o), and (d) ΔLST_{min}(f-o). Panel (e) shows the RMSD over the Köppen-Geiger climate zone of ΔET(f-o) of CLM‑BASE (red), and CLM‑PLUS (orange) against MODIS (green edge), GLEAM (light blue edge), GETA (dark blue edge). The numbers indicate the number of data samples used for the calculation of RMSD.
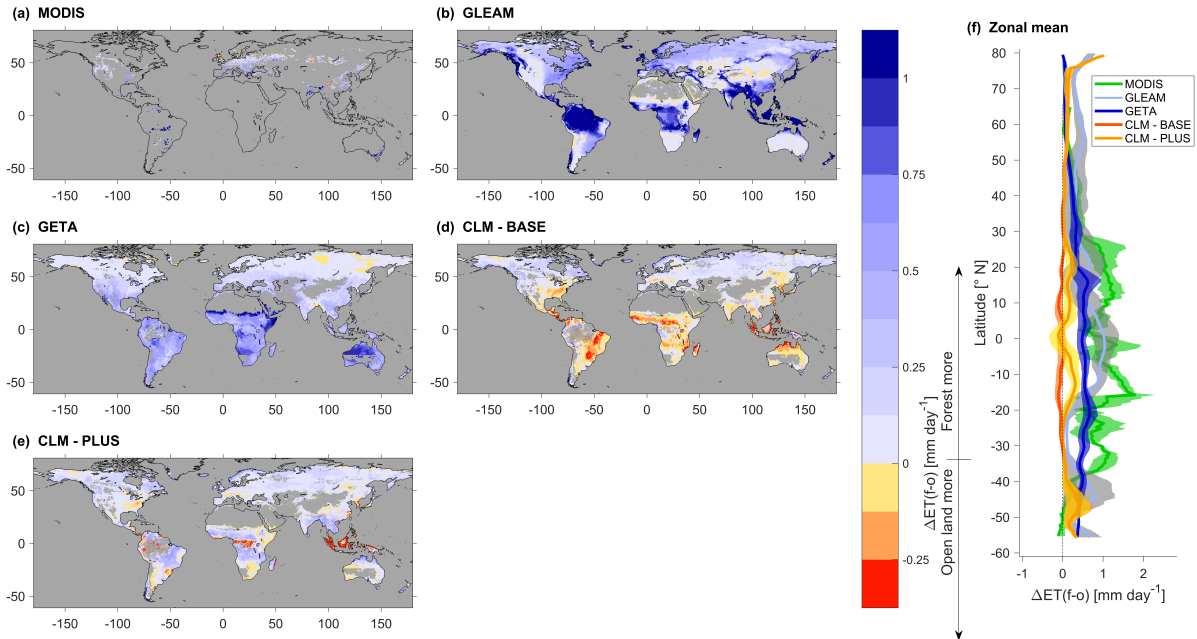
33

Figure 5. Annual mean ΔET(f-o) in (a) MODIS, (b) GLEAM, (c) GETA, ~~and~~ (d) CLM‑BASE, and (e) CLM‑PLUS. Panel (~~e~~f) shows the zonal mean (thick line) and interquartiel range (shading) of MODIS (~~in~~ green~~along with its confidence interval in grey~~), GLEAM (light blue, grey shading), GETA (~~orange~~dark blue), ~~and~~ CLM‑BASE (red), and CLM‑PLUS (orange). Note that on this subfigure results have been smoothed with a 4° ~~latitdudinally‑running~~ latitudinally‑running mean.
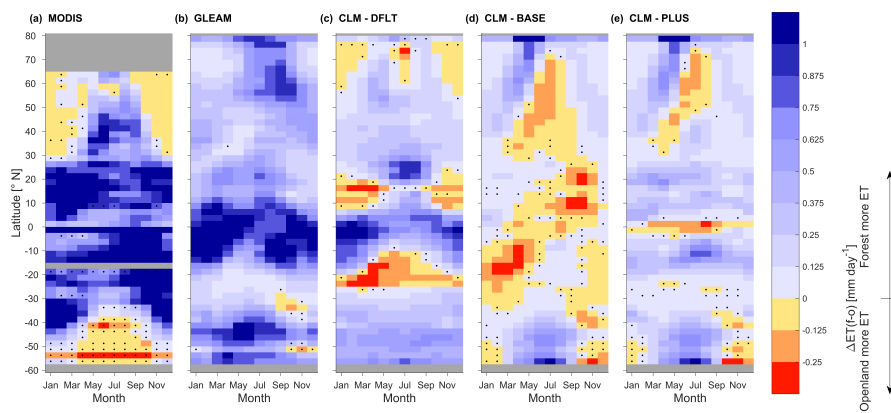
Figure 6. Seasonal and latitudinal variations of ΔET(f-o) in (a) the MODIS and (b) GLEAM observations, (c) CLM - DFLT, (d) CLM - BASE, and (e) CLM - PLUS. Points with a mean which is insignificantly different from zero in a two-sided t-test at 95 % confidence level are marked with a black dot. All data from the 2002-2010 analysis period corresponding to a given latitude and a given month are pooled to derive the sample set for the test.
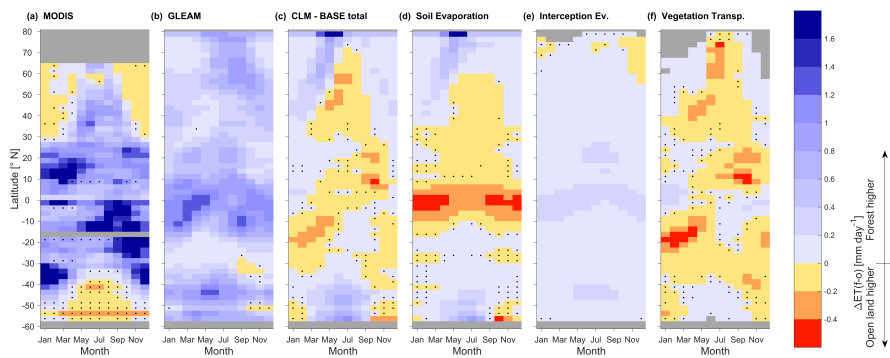
Figure 7. Seasonal and latitudinal variations of $\Delta ET(f\text{-}o)$ in (a) MODIS, (b) GLEAM, and difference of forest minus open land in (c) total ET, (d) soil evaporation, (e) canopy interception evaporation, and (f) vegetation transpiration in CLM - BASE. Points with a mean which is insignificantly different from zero in a two-sided t-test at 95 % confidence level are marked with a black dot.

Figure 8. Seasonal and latitudinal variations of ΔLST_{avg}(f-o) in (a) the MODIS observations~~and~~, (b) CLM‑DFLT, and (c) CLM‑BASE. Points with a mean which is insignificantly different from zero in a two-sided t-test at 95 % confidence level are marked with a black dot. All data from the 2002-2010 analysis period corresponding to a given latitude and a given mon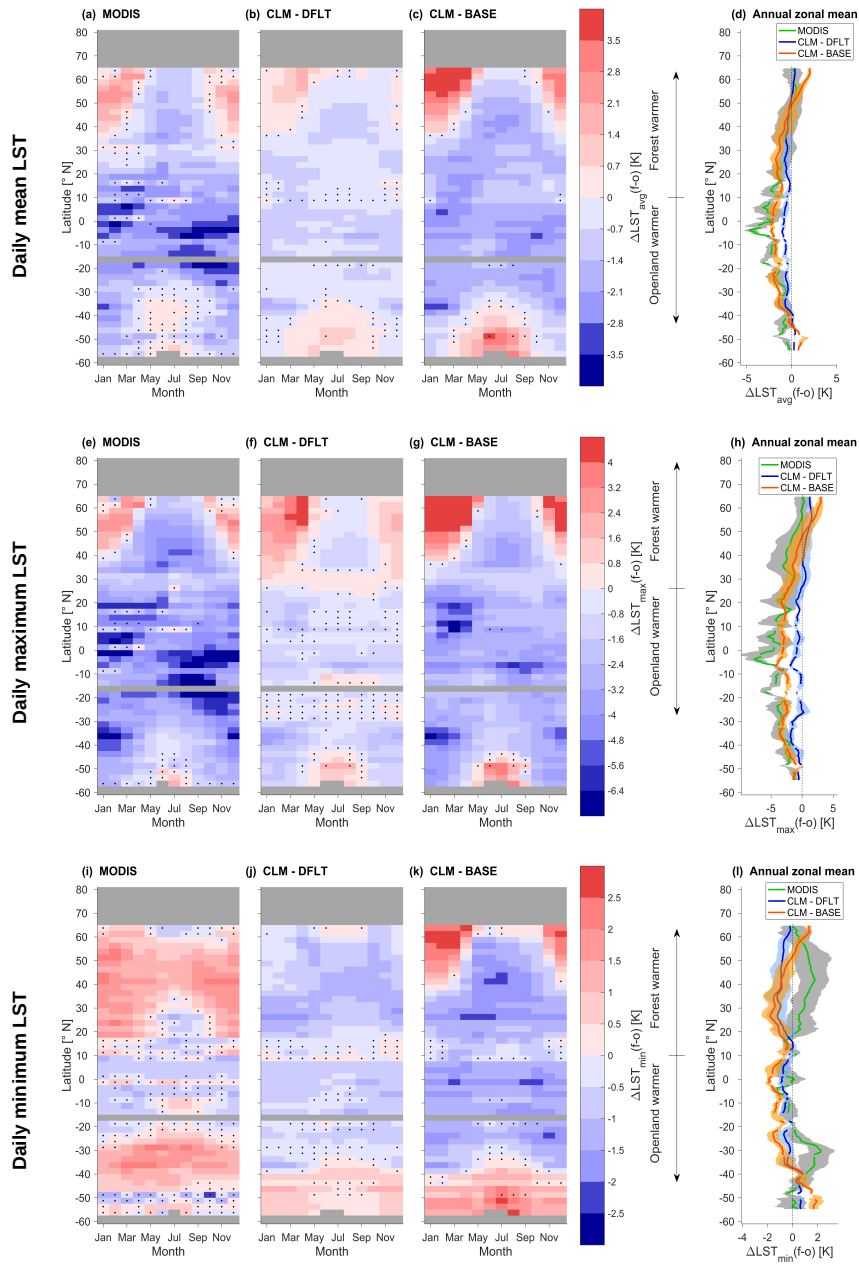th are pooled to derive the sample set for the test. Panel (~~e~~d) shows the zonal annual mean of ~~both~~ MODIS (~~in~~ green~~along with its confidence interval~~, interquartile range in grey), CLM‑DFLT (blue, interquartile range in blue), and CLM‑BASE (~~in~~ red, ~~confidence interval~~ interquartile range in orange). Note that on this subfigure results have been smoothed with a $4°$ latitudinally-running mean. Only grid cells containing valid data in the MODIS observations were considered for the analysis of CLM‑DFLT and CLM‑BASE. The same for ΔLST_{max}(f-o) in panels (~~d~~e), (~~e~~f), (~~f~~g), (h) and for ΔLST_{min}(f-o) in panels (~~g~~i), (~~h~~j), (~~i~~k), (l).

37
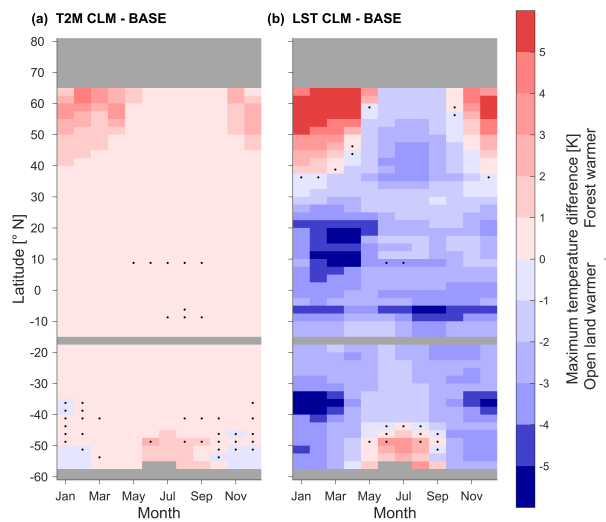
Figure 9. Seasonal and latitudinal variations of (a) daily maximum T2M difference of forest minus open land and (b) ΔLST_{max}(f-o) in CLM - BASE. Points with a mean which is insignificantly different from zero in a two-sided t-test at 95 % confidence level are marked with a black dot. All data from the 2002-2010 analysis period corresponding to a given latitude and a given month are pooled to derive the sample set for the test. Only grid cells containing valid data in the MODIS observations were considered for the analysis.

Overview of the different modifications of CLM4.5 incorporated in the simulations presented this study. Simulation SeSCs Shallow roots 10 % Light limitation V_{cmax} CLM - DFLT – – — – CLM - BASE – – – – CLM - PLUS

Area-weighted annual mean of $\Delta\alpha$(f-o), ΔLST_{avg}(f-o), ΔLST_{max}(f-o), and ΔLST_{min}(f-o) over the Köppen-Geiger climate zones in MODIS observations and CLM - BASE. For these variables only grid cells containing valid data in the MODIS observations were considered for the analysis of CLM - BASE. Area-weighted annual mean of ΔET(f-o) over the Köppen-Geiger climate zones in MODIS, GLEAM, GETA and CLM - BASE. Also shown is the fraction of these climate zones covered by the MODIS $\Delta\alpha$(f-o) observations (note that these fractions vary slightly for different variables).

~~Area-weighted annual mean ET for each PFT analyzed in this study according to the GETA data and in CLM - BASE as well as the fraction of the land surface covered by the different PFTs. Abbreviation Full name Frac. %GETA CLM - BASE NET - temperate Temperate evergreen needleleaf tree 3.2 1.74 1.78 NET - boreal Boreal evergreen needleleaf tree 6.9 1.00 0.97 NDT - boreal Boreal deciduous needleleaf tree 1.0 0.72 0.72 EBT - tropical Tropical evergreen broadleaf tree 9.5 3.47 3.70 EBT - temperate Temperate evergreen broadleaf tree 1.5 2.58 2.61 DBT - tropical Tropical deciduous broadleaf tree 8.0 2.65 2.31 DBT - temperate Temperate deciduous broadleaf tree 3.1 1.78 1.74 DBT - boreal Boreal deciduous broadleaf tree 1.3 1.23 1.08 C_3 arctic grass 3.1 0.81 0.66 C_3 grass 8.8 1.48 1.60 C_4 grass 8.0 2.06 2.32 Crop Unmanaged rainfed C_3 crop 10 1.90 1.76~~

Table 1. ET and latent heat flux in-situ observations from various studies and the values in CLM - BASE and CLM - PLUS at the respective locations.

Study	Region	PFTs	Unit	Season	Obs.	CLM - BASE	CLM - PLUS
da Rocha et al. (2004)	Amazon	EBT	$\mathrm{mm\,day^{-1}}$	Dry	3.96	3.49	3.48
				Wet	3.18	3.57	3.37
				All	3.51	3.54	3.40
von Randow et al. (2004)	Amazon	EBT	$\mathrm{W\,m^{-2}}$	Dry	108.6	82.9	90.8
				Wet	104.5	113.9	108.9
		Grass		Dry	63.9	81.2	64.7
				Wet	83.0	113.9	100.1
Liu et al. (2005)	Alaska	Grass	$\mathrm{W\,m^{-2}}$	All	16.1	16.4	16.8
		DBT		All	22.5	13.7	14.1
		ENF		All	23.9	18.0	18.4

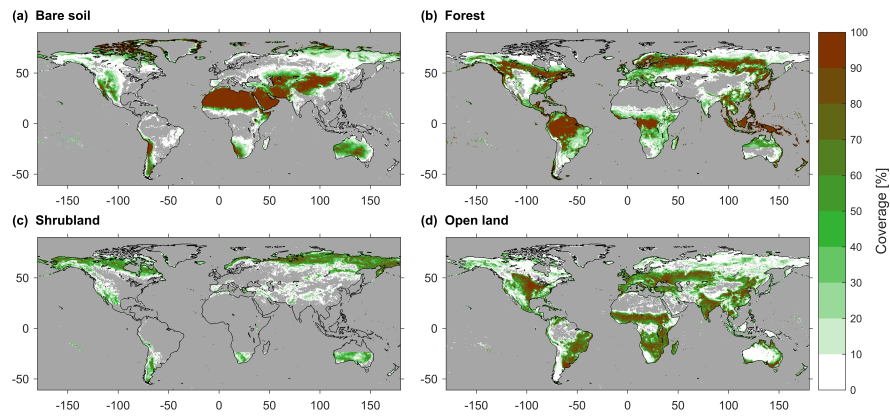**Figure A1.** The fraction of the CLM4.5 grid cells covered by (a) bare soil, (b) forest, (c) shrubland, and (d) open land.
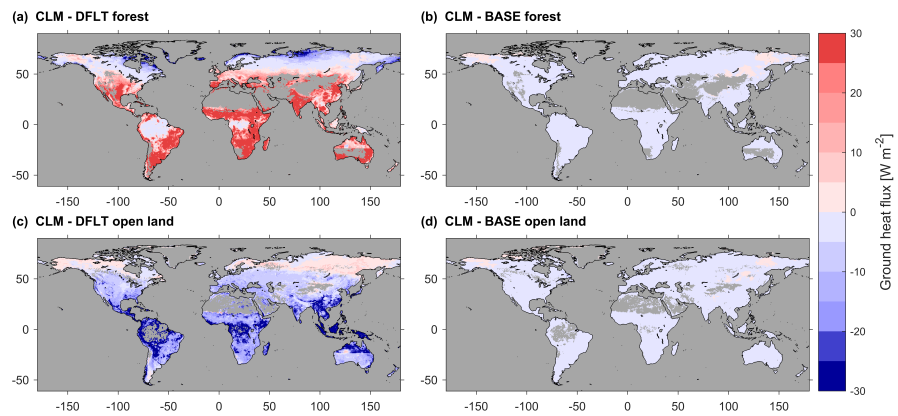
Figure A2. ~~GH flux~~ GHF for forests (a and b) and open land (c and d) in CLM - DFLT (a and c) and CLM - BASE (b and d). Positive values correspond to a heat flux out of the soil.

Figure A3. Difference in vertically-averaged annual mean soil temperature of forest minus open land in CLM - BASE.
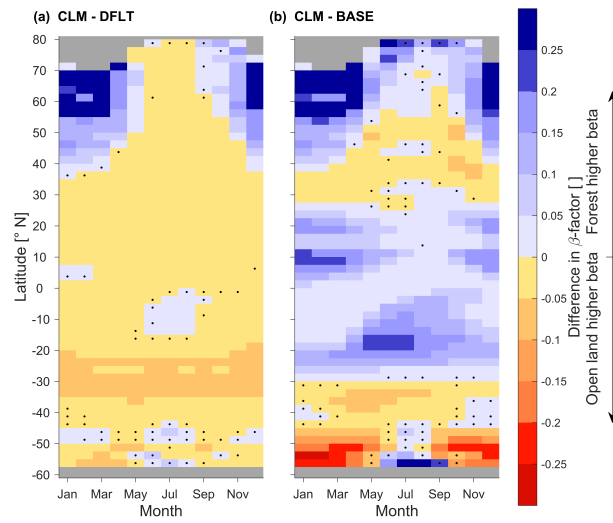
Figure A4. Seasonal and latitudinal variations of β_t-factor differences of forest minus open land in (a) CLM‑DFLT and (b) CLM‑BASE. Points with a mean which is insignificantly different from zero in a two-sided t-test at 95 % confidence level are marked with a black dot.
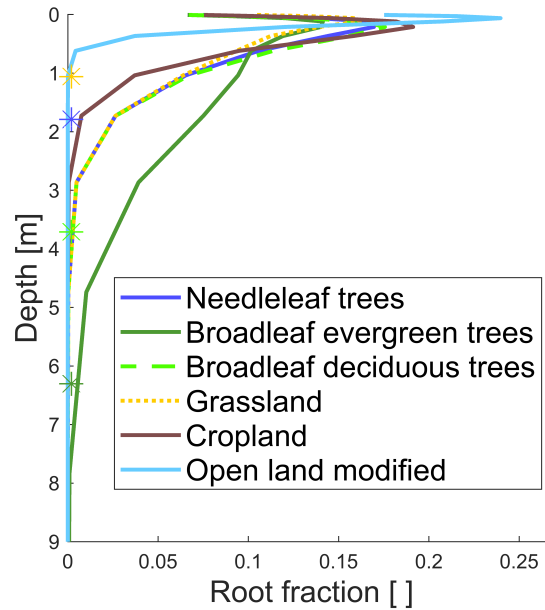
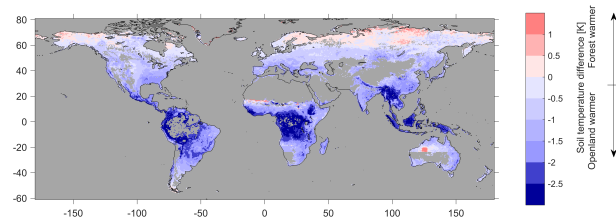**Figure A5.** Vertical root fraction distribution of the different PFTs in the default version of CLM4.5 and in light blue the modified root fraction distribution of open land PFTs used in CLM‑PLUS. The asterisks mark the reported maximum rooting depths of Fan et al. (2017) for annual grass (yellow), evergreen needleleaf trees (dark blue), deciduous broadleaf trees (ligth green), and evergreen broadleaf trees (dark green).
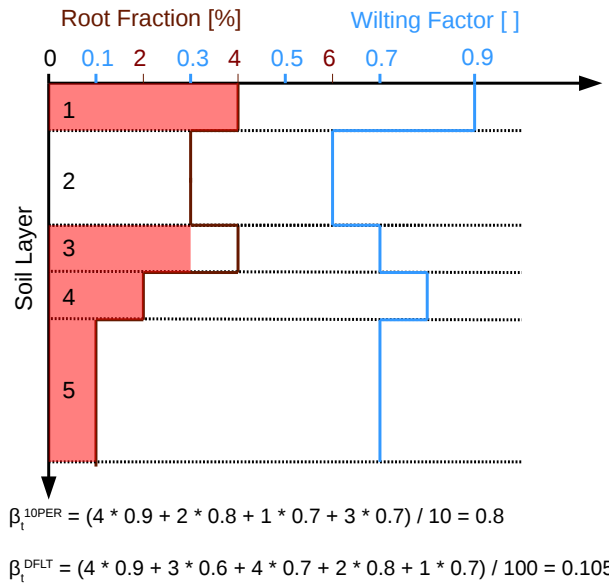
Figure A6. Example of the calculation of the β_t-factor with the 10 % modification. Shown are five soil layers with the fraction of the roots in these layers in brown and the wilting factor in blue. On the bottom the calculation of β_t for this particular example with the 10 % modification (β_t^{10PER}) and the default calculation in CLM4.5 (β_t^{DFLT}), assuming the roots not shown have a wilting factor of zero. The root fractions eventually used to calculate β_t^{10PER} are shaded in red.
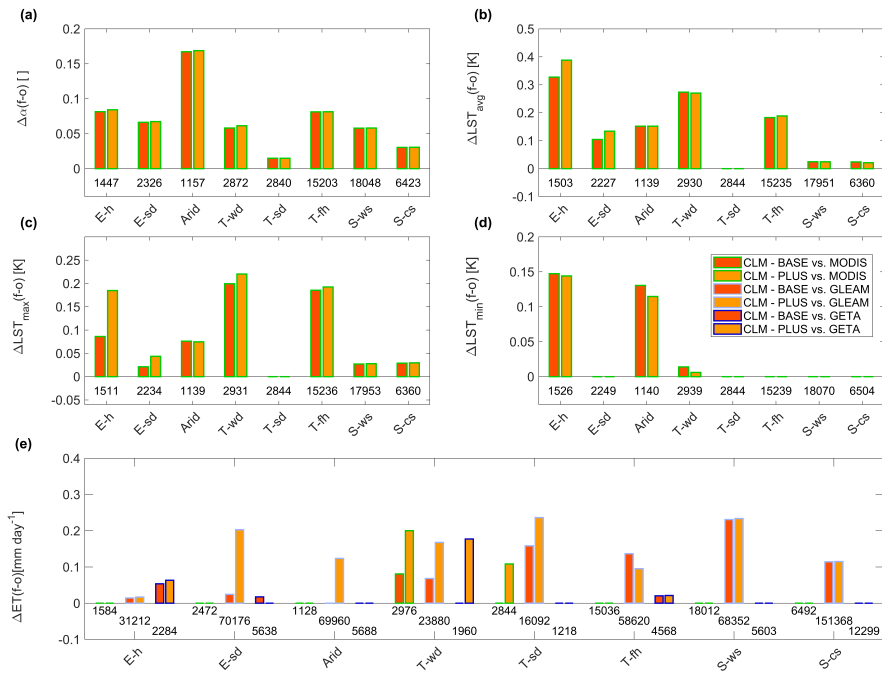
Figure A7. IA (Duveiller et al., 2016) of CLM‑BASE (red), and CLM‑PLUS (orange) with MODIS observations over Köppen-Geiger climate zones (Kottek et al., 2006, ; Fig. 1) for monthly (a) $\Delta\alpha$(f-o), (b) ΔLST_{avg}(f-o), (c) ΔLST_{max}(f-o), and (d) ΔLST_{min}(f-o). Panel (e) shows the IA over the Köppen-Geiger climate zone for ΔET(f-o) of CLM‑BASE (red), and CLM‑PLUS (orange) with MODIS (green edge), GLEAM (light blue edge), GETA (dark blue edge). The numbers indicate the number of data samples used for the calculation of IA.
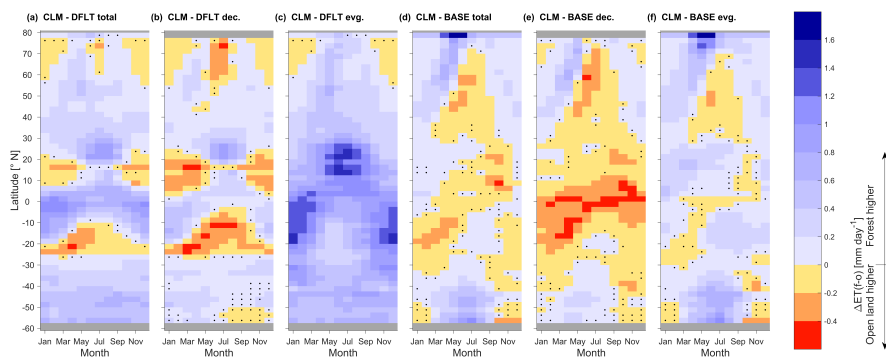
Figure A8. Seasonal and latitudinal variations of $\Delta ET(f\text{-}o)$ in ~~(a) MODIS, (b) GLEAM and, in~~ CLM - ~~BASE~~ DFLT for (~~e~~a) all tree PFTs minus open land, (~~d~~b) deciduous tree PFTs only minus open land, (~~e~~c) evergreen tree PFTs only minus open land. Points with a mean which is insignificantly different from zero in a two-sided t-test at 95 % confidence level are marked with a black dot. All data from the 2002-2010 analysis period corresponding to a given latitude and a given month are pooled to derive the sample set for the test. The same for CLM - BASE in panels (d), (e), and (f).

As Fig A8 but for CLM - DFLT.

Table A1. The PFT-specific values of V_{cmax} [$\mu mol\,m^{-2}\,s^{-1}$], r_a, and r_b [ ] in default of CLM4.5 and in CLM‑PLUS.

PFT name	Default			CLM‑PLUS		
	r_a	r_b	V_{cmax}	r_a	r_b	V_{cmax}
NET‑temperate	7.0	2.0	62.5	default		50
NET‑boreal	7.0	2.0	62.6	default		default
NDT‑boreal	7.0	2.0	39.1	default		default
EBT‑tropical	7.0	1.0	55.0	default		35
EBT‑temperate	7.0	1.0	61.5	default		50
DBT‑tropical	6.0	2.0	41.0	default		65
DBT‑temperate	6.0	2.0	57.7	default		default
DBT‑boreal	6.0	2.0	57.7	default		70
C_3 arctic grass	11.0	2.0	78.2	11.0	11.0	90
C_3 grass	11.0	2.0	78.2	11.0	11.0	60
C_4 grass	11.0	2.0	51.6	11.0	11.0	default
Crop	6.0	3.0	100.7	11.0	11.0	90

~~As Fig. 5 but for the CLM‑PLUS.~~

5 ~~Seasonal and latitudinal variations of ΔLST_{max}(f-o) in (a) the MODIS observations and in (c) CLM‑BASE. Panel (b) shows the difference in daily maximum T2M in CLM‑BASE.~~

~~Seasonal and latitudinal variations of ΔLST_{avg}(f-o) in (a) the MODIS observations, (b) CLM‑DFLT, and (c) CLM‑BASE.~~
930 ~~Points with a mean which is insignificantly different from zero in a two-sided t-test at 95 % confidence level are marked with a black dot. Only grid cells containing valid data in the MODIS observations were considered for the analysis of CLM‑DFLT and CLM‑BASE. The same for ΔLST_{max}(f-o) in panels (d),(e),(f) and for ΔLST_{min}(f-o) in panels (g),(h),(i).~~

Table A2. The default PFT classification in CLM4.5.

No.	Abbr.	Full name	Class
1	Bare soil		–
2	NET - temperate	Temperate evergreen needleleaf tree	~~Tree~~ Forest
3	NET - boreal	Boreal evergreen needleleaf tree	~~Tree~~ Forest
4	NDT - boreal	Boreal deciduous needleleaf tree	~~Tree~~ Forest
5	BET - tropical	Tropical evergreen broadleaf tree	~~Tree~~ Forest
6	BET - temperate	Temperate evergreen broadleaf tree	~~Tree~~ Forest
7	BDT - tropical	Tropical deciduous broadleaf tree	~~Tree~~ Forest
8	BDT - temperate	Temperate deciduous broadleaf tree	~~Tree~~ Forest
9	BDT - boreal	Boreal deciduous broadleaf tree	~~Tree~~ Forest
10	BES - temperate	Temperate evergreen broadleaf shrub	–
11	BDS - temperate	Temperate deciduous broadleaf shrub	–
12	BDS - boreal	Boreal deciduous broadleaf shrub	–
13	C_3 arctic grass		Open land
14	C_3 grass		Open land
15	C_4 grass		Open land
16	Crop	Unmanaged rainfed C_3 crop	Open land

Table A3. The land cover types from Ambrose and Sterling (2014) (GETA) used in this study and the number of respective PFTs in CLM4.5 applied to the different land cover types (Table A2).

Abbr. GETA	Full name GETA	PFTs of CLM4.5
ENF	Evergreen needleleaf forest	2, 3
DNF	Deciduous needleleaf forest	4
EBF	Evergreen broadleaf forest	5, 6
DBF	Deciduous broadleaf forest	7, 8, 9
GRS	Grassland	13, 14, 15
CRN	Non-irrigated cropland	16

Table A4. Overview of the different modifications of CLM4.5 incorporated in the simulations presented this study.

Run	SeSCs	Shallow roots	10 %	Light limitation	V_{cmax}
CLM - DFLT	–	–	–	–	–
CLM - BASE	✓	–	–	–	–
CLM - ROOT	✓	✓	–	–	–
CLM - 10PER	✓	✓	✓	–	–
CLM - LIGHT	✓	✓	✓	✓	–
CLM - PLUS	✓	✓	✓	✓	✓

Table A5. Area-weighted annual mean ET for each PFT analyzed in this study according to the GETA data and in the different configurations of CLM4.5 and fraction of the land surface covered by the different PFTs. On the bottom is listed the global integral of annual ET.

Abbr.	Full name	Frac. [%]	ET [mm day^{-1}]					
			GETA	BASE	ROOT	10PER	LIGHT	PLUS
NET - temperate	Needleleaf evergreen tree - temperate	3.2	1.74	1.78	1.78	1.81	1.84	1.75
NET - boreal	Needleleaf evergreen tree - boreal	6.9	1.00	0.97	0.97	0.98	1.00	1.00
NDT - boreal	Needleleaf deciduous tree - boreal	1.0	0.72	0.72	0.72	0.72	0.73	0.73
EBT - tropical	Broadleaf evergreen tree - tropical	9.5	3.47	3.70	3.70	3.78	3.87	3.52
EBT - temperate	Broadleaf evergreen tree - temperate	1.5	2.58	2.61	2.61	2.66	2.70	2.60
DBT - tropical	Broadleaf deciduous tree - tropical	8.0	2.65	2.31	2.31	2.42	2.44	2.62
DBT - temperate	Broadleaf deciduous tree - temperate	3.1	1.78	1.74	1.74	1.76	1.79	1.79
DBT - boreal	Broadleaf deciduous tree - boreal	1.3	1.23	1.08	1.08	1.08	1.10	1.13
C_3 arctic grass		3.1	0.81	0.66	0.65	0.65	0.66	0.67
C_3 grass		8.8	1.48	1.60	1.53	1.56	1.57	1.53
C_4 grass		8.0	2.06	2.32	2.18	2.22	2.12	2.04
Crop	C_3 unmanaged rainfed crop	10	1.90	1.76	1.70	1.73	1.74	1.73
Total ET [km^3 yr^{-1}]				70223	69059	70322	70649	69023

Table A6. ET and latent heat flux in-situ observations from various studies and the values of the different CLM4.5 sensitivity tests at the respective locations.

Study	Region	PFTs	Unit	Season	Obs.	BASE	ROOT	10PER	LIGHT	PLUS
da Rocha et al. (2004)	Amazon	EBT	mm day^{-1}	Dry	3.96	3.49	3.49	3.90	4.06	3.48
				Wet	3.18	3.57	3.57	3.57	3.64	3.37
				All	3.51	3.54	3.54	3.68	3.79	3.40
von Randow et al. (2004)	Amazon	EBT	W m^{-2}	Dry	108.6	82.9	82.9	100.6	105.3	90.8
				Wet	104.5	113.9	113.9	113.8	116.2	108.9
		Grass		Dry	63.9	81.2	56.0	60.2	62.7	64.7
				Wet	83.0	113.9	113.9	113.9	106.1	100.1
Liu et al. (2005)	Alaska	Grass	W m^{-2}	All	16.1	16.4	16.8	16.8	16.8	16.8
		DBT		All	22.5	13.7	13.7	14.0	14.0	14.1
		ENF		All	23.9	18.0	18.0	18.4	18.4	18.4